

А. С. Нинул
**ОПТИМИЗАЦИЯ
ЦЕЛЕВЫХ ФУНКЦИЙ**

- основные аналитические, вычислительные и планово-экспериментальные аспекты процесса оптимизации
- экстремумы разных уровней для скалярных функций от независимой и от зависимой скалярной переменной
- экстремум скалярной функции от вектор-переменной
- условные экстремумы, методы их задания и поиска
- двухступенчатый параметрический метод
- метод множителей Лагранжа и метод условных тензор-производных в их параллельном сопоставлении
- предельные методы с малым и большим параметрами
- безусловный и условный экстремумы вещественных функций от разнообразных комплексных переменных
- оригинальные примеры решения задач на экстремум
- генеральное неравенство для средних величин
- полные требования к коэффициентам алгебраического уравнения для положительности всех его корней
- условный экстремум скалярной функции 2-го порядка от вектор-аргумента при линейном ограничении
- нормальное решение линейного уравнения
- максимум и минимум функции отношения Релея
- метод наименьших квадратов и линейная регрессия
- численные методы одно- и многомерной оптимизации
- планово-вычислительные методы оптимизации
- планово-экспериментальные методы оптимизации
- анализ вкладов случайной и систематической ошибок
- критерий адекватности полиномиальных моделей

А. С. Нинул
ОПТИМИЗАЦИЯ ЦЕЛЕВЫХ ФУНКЦИЙ

А. С. Нинул

ОПТИМИЗАЦИЯ ЦЕЛЕВЫХ ФУНКЦИЙ

АНАЛИТИКА
ЧИСЛЕННЫЕ МЕТОДЫ
ПЛАНИРОВАНИЕ ЭКСПЕРИМЕНТА



$$2 \times 2 = 11$$

ISBN 9785-94052-175-4



Издательство «ФИЗМАТЛИТ»

А.С. Нинул

ОПТИМИЗАЦИЯ ЦЕЛЕВЫХ ФУНКЦИЙ

**АНАЛИТИКА
ЧИСЛЕННЫЕ МЕТОДЫ
ПЛАНИРОВАНИЕ ЭКСПЕРИМЕНТА**



Москва
ФИЗМАТЛИТ
2009

ББК 22.161.5
Н 60
УДК 517.5/519.6/512

Р е ц е н з е н т:

доктор физико-математических наук, профессор,
заслуженный деятель науки Российской Федерации **О. В. Мантуров**

НИНУЛ А. С. Оптимизация целевых функций:

Аналитика. Численные методы. Планирование эксперимента. —
М.: Издательство Физико-математической литературы, 2009, 336с.
ISBN 9785-94052-175-4

В монографии рассмотрены основные аналитические, численные, планово-вычислительные и планово-экспериментальные методы для поиска и идентификации экстремумов целевых функций от одной или от нескольких скалярных переменных. Столь обширный охват методов оптимизации обусловлен стремлением автора отобразить в одной книге проблему в целом. Даны характерные примеры, в том числе из общей и линейной алгебры, аппроксимационного и регрессионного анализа.

Для специалистов в области анализа и решений экстремальных задач, а также научных работников, преподавателей, аспирантов и студентов физико-математических и технических специальностей.

К читателям

В области теории функций, начиная с исторических работ Ферма, Ньютона и Лейбница, всегда особо пристальное внимание математиков привлекали постановка и разрешение экстремальных проблем. Впервые они произошли из ряда геометрических задач, которые, по сути, были продиктованы какими-либо жизненными ситуациями.

За прошедшие три века развития строгого математического анализа сформировались главные направления и основные методы оптимизации целевых функций, т. е. выявления и идентификации их экстремумов.

Как приоритетные в книге определяются следующие направления:

- аналитическая оптимизация целевых функций от одной независимой или от одной зависимой скалярной переменной;
- аналитическая безусловная оптимизация целевых функций от одной независимой векторной переменной;
- аналитическая условная оптимизация целевых функций от зависимой или от ограниченной векторной переменной;
- численная оптимизация целевых функций от скалярной переменной;
- численная безусловная и условная оптимизация целевых функций от нескольких скалярных переменных;
- планово-вычислительная оптимизация целевых функций;
- планово-экспериментальная оптимизация целевых функций отклика.

Главная цель данной книги состоит в том, чтобы в одном издании последовательно и в логически естественном порядке охватить всю проблему оптимизации целевых функций (вплоть до математического программирования) с заполнением имеющихся пробелов.

В конце книги, как приложение, приведена физико-математическая кунсткамера с рядом проблемных вопросов и задач.

Автор будет признателен тем читателям, кто выскажет свои отзывы, замечания или какие-либо полезные предложения по этой монографии на интернет-сайте: [«http://www.ninulas.narod.ru/»](http://www.ninulas.narod.ru/).

Ninul A. S. Optimization of Objective Functions.

Analysis. Numerical methods. Design of experiments.

Appendix. Physical-Mathematical Kunstcamera.

Publishing House «Fizmatlit», Moscow, 2009.

Resume

The main aim of this monograph is to consider the basic optimization methods for objective functions (up to mathematical programming) in logic order which emphasizes their genesis, and also to fill up the existing here “blind spots”.

In the first Chapter the analytical aspects of finding the unconditional extremum for objective functions of a scalar or of a vector variable are considered. The solutions of special tasks are discussed, for example, the proof of full hierarchy of all average positive values. (Specific illustrative meaning to decision of the last task is given in Chapters 1, 3 and 4.)

In the second Chapter the analytical aspects of finding the conditional extremum for objective functions of a constrained vector variable – either depended on some parameters or bounded by some connection equations, are considered. In addition, analytical bases of limit optimization methods are discussed. The geometrical interrelation between all three directions of conditional optimization is demonstrated with the use of eigen functional projectors in two symmetric matrix forms. A secular equation in stationary point for “conditional eigenvalues” of the Hesse matrix is derived.

In the third Chapter a formal analysis for nonholomorphic functions of complex variables is developed (without, as is usually done, doubling their dimension). With the use of this formal complex analysis the methods of unconditional and conditional optimization for objective real functions of one or several pairs complex conjugate or of mixed variables are proposed.

In the fourth Chapter important examples of solutions of extremal problems in general and linear algebra are given. As one of the results it is possible to note the theorem about complete requirements to the coefficients of a real algebraic equation that provide reality and positivity of its roots.

In the fifth Chapter the main numerical methods of the orders 0, 1 and 2 of searching the extremum for objective functions of one or of several scalar variables are considered. Methods of searching the conditional extremum are considered separately in the two cases of the vector variable – depended or bounded as indicated earlier.

In the sixth Chapter the main planned-calculational methods of searching the extremum for objective functions of several scalar variables on the basis of their difference models of the 1st, incomplete and complete 2nd orders are considered.

In the seventh Chapter the main planned-experimental methods of searching the extremum for objective functions of response of several factors on the basis of their linear regression models, identical by form to difference models of the incomplete and complete 2nd orders, are considered.

Web-site for the communication: www.ninulas.narod.ru/english.html

*«Мы любим всё – и жар холодных числ,
И дар божественных видений,
Нам внятны всё – и острый галльский смысл,
И сумрачный германский гений.»*
Александр Блок «Скифы»

Введение

Широко известно крылатое изречение великого Леонарда Эйлера: *«Так как здание всего мира совершенно и возведено премудрым Творцом, то в мире не происходит ничего, в чём бы не был виден смысл какого-то максимума или минимума»*. И, действительно, в качестве примера возьмём, казалось бы, весьма далёкую от математики музыкальную классику. Музыкальная мысль в процессе своего развития оказывает максимальное эмоциональное воздействие именно на самых высоких нотах, но и вызывает состояние наибольшего умиротворения, наоборот, на самых низких нотах. Для иллюстрации можно указать хотя бы на начальные темы из волшебной 19-й сонаты Бетховена. Или вспомним в том же аспекте патетические прелюдии Рахманинова *cis-moll* и *g-moll*. Это, конечно, только ярчайшие примеры из нашей духовной сферы.

Если из указанных заоблачных вершин высочайшего искусства мы перенесёмся в сугубо прикладные направления деятельности человека, то здесь практически во всех её сферах и на каждом шагу столкнёмся с разнообразными примерами поиска или уже состоявшейся реализации экстремальных решений, т. е. либо максимизации, либо минимизации, либо же их комбинации — для каких-то количественных величин.

Как хорошо известно, первой в истории человечества поставленной задачей на экстремум считается Задача Дидоны (Финикийская царица, IX век до н.э., согласно древнегреческой поэме Вергилия «Энеида»), или *изопериметрическая задача*. В наше время она формулируется так: «На евклидовой плоскости среди замкнутых плоских кривых, имеющих заданную длину, найти кривую, охватывающую максимум площади.» Интуитивно эта древнейшая задача, согласно поэме, была практически разрешена самой Дидоной, указавшей охватить участок земли по кругу. В той же Древней Греции была поставлена и разрешена гораздо более сложная *изопифанная задача*, заключающаяся в нахождении замкнутой поверхности с заданной её площадью, охватывающей в пространстве максимум объёма. Историки науки не смогли установить, кому же из древнегреческих мыслителей первым удалось дать научно приемлемые доказательства свойств максимальной вместимости круга и шара. Но известно, что Аристотель (IV век до н.э.) пользовался в своих трудах этими геометрическими фактами, как уже вполне доказанными [34].

Неустанное стремление математиков давать всё более строгие и изящные решения разнообразных экстремальных задач способствовало развитию математической науки и её фундаментальных основ. На этом пути получили значительные результаты многие корифеи математики — от Архимеда в III веке до н.э. до Штейнера и Вейерштрасса в XIX веке.

Однако тут следует заметить, что практически до конца XVIII века в европейской математике превалировало геометрическое направление в постановке и разрешении самых различных экстремальных проблем. Даже математики зачастую узко именовались геометрами, хотя это и являлось тогда как бы высшей оценкой их деятельности. Но очевидно то, что *аналитическим образом* экстремальные задачи могут разрешаться в самом общем виде единственно тогда, когда эти задачи изложены на языке абстрактных математических понятий и операций с ними.

Подобную сверхзадачу в новейшей истории развития мировой науки впервые поставил и корректно разрешил в 1629 г. Пьер Ферма для класса степенных функций. Он изложил свой аналитический метод в письме к Жиллю Робервалю (1638 г.) — известному математику того времени и одному из провозвестников зарождавшегося тогда дифференциального исчисления [58]. Затем Готфрид Лейбниц дал её решение для общего класса не менее чем дважды дифференцируемых функций, в том числе для иррациональных и трансцендентных, в своей знаменитой статье «Новый метод максимумов и минимумов ...» (1684 г.) [61]. В ней же фундаментальным образом были заложены и применены практически первоосновы дифференциального исчисления в принятых и поныне стройных обозначениях. Несколько ранее (1671 г.), как известно, основы дифференциального исчисления заложил Исаак Ньютон в собственной оригинальной форме флюксий. Опубликовать этот исторический труд удалось лишь через 65 лет посмертно, когда авторитет Ньютона стал непререкаемым. В рассматриваемом аспекте весьма интересно и то, что Ньютон основал ещё одно новое направление — численное решение скалярных уравнений, разработав классический метод касательных, в том числе и тех дифференциальных уравнений, которые возникают при решении экстремальных задач [63].

Как Ньютон, так и Лейбниц на отдельных частных примерах уже имели дело с приращениями функций от 2-х и более числа переменных. Но фундаментальным образом к анализу приращений и экстремумов для функций от нескольких переменных впервые подошёл в середине XVIII века Леонард Эйлер [97]. Им были получены общие выражения для дифференциалов таких скалярных функций — частных и полного, найдено необходимое условие для неособого экстремума — обнуление первых дифференциалов функции, начат анализ природы экстремума исходя из значений вторых дифференциалов.

В результате дальнейшего продвижения по данному направлению математики естественно пришли к понятиям дифференциальных форм и установлению их взаимосвязи с экстремальным поведением функций. В частности, фундаментальные результаты по анализу влияния второй дифференциальной формы на характер экстремума функции в точке её стационарности получил Джеймс Сильвестр.

Постановку и эффективное решение задачи на условный экстремум функции от нескольких переменных реализовал Жозеф Луи Лагранж в 1797 г. в фундаментальном труде «Аналитическая теория функций» [60]. Предложенный им и ставший сразу классическим *метод множителей* для решения задач на условный экстремум был навеян его более ранними результатами в области аналитической механики систем со связями.

Новый прорыв в решении задач на условный экстремум численным путём сделали Курант и его школа в середине XX века [10], разработав *метод штрафных функций* (с применением большого параметра).

Вообще численные методы решения экстремальных задач начали интенсивно развиваться только в XX веке в связи с появившимися тогда насущными практическими потребностями и в рамках общего процесса разработки разнообразных теорий оптимизации. Но фундаментальные истоки численных методов оптимизации для скалярных функций, или численных методов поиска экстремума находились в работах классиков математического анализа — Ньютона, Эйлера, Коши. Так, *градиентный метод* Огюстен Луи Коши изложил в отдельной статье в 1847 г. [56].

Общий метод Ньютона получил развитие и обоснование в работах Л.В. Канторовича [20]. Метод с малым параметром для решения ряда некорректных и почти вырожденных задач безусловной оптимизации скалярных функций при неточных значениях их исходных элементов, известный как *метод регуляризации*, предложил А.Н.Тихонов [36, 37].

Кроме того, в середине XX века с появлением основополагающей публикации Дж. Бокса и К. Уилсона [54] широкий ряд исследователей начал тоже дерзать в области оптимизации, максимально приближённой к практической и даже к производственной сфере. Данное направление логично получило общепринятое название *планирование эксперимента при поиске оптимальных условий*. Дж. Бокс также впервые опубликовал статью [55], направленную на применение планирования эксперимента непосредственно в сфере промышленного производства с практической целью повышения его производительности. Следует, однако, заметить, что математики-профессионалы высокого уровня в XX веке не проявили какого-либо заметного интереса к новому направлению в оптимизации, видимо, из-за его слишком прозаического характера. Это идёт вразрез с подходом отцов-основателей математического анализа. Отсюда в ряде публикаций в данной области имеется много явных несообразностей.

Главная цель данной монографии состоит в том, чтобы изложить в естественном порядке все основные направления оптимизации целевых функций, т. е. поиска и идентификации их максимумов или минимумов, и при необходимости в этом шаге заполнить имеющиеся «белые пятна». Как видно из содержания монографии, автор вовсе не придерживается компиляционного подхода, свойственного чисто учебной литературе.

Однако за рамками этого изложения остаются геометрические методы решения экстремальных задач, так как все они имеют ярко выраженный эвристический характер и прямо не связаны с аналитическим подходом. В имеющейся литературе, например [5, 28, 34], геометрические методы оптимизации рассматриваются, как правило, отдельно и представлены на весьма фундаментальном уровне.

Монография состоит из семи глав, охватывающих в максимально полном виде исходные аналитические основы методов оптимизации и далее производимые по иерархии численные, планово-вычислительные и планово-экспериментальные процедуры поиска экстремума целевой скалярной функции от одной и от нескольких переменных.

В 1-й главе излагаются аналитические аспекты решения задач на безусловный экстремум целевой функции сначала от одной и затем от нескольких скалярных переменных (от одной векторной переменной). На характерных примерах иллюстрируются экстремальные свойства целевых функций. Рассматриваются решения конкретных задач, в том числе задачи на доказательство иерархии всех средних величин, которой в первых 4-х главах книги придаётся особое иллюстративное значение.

Во 2-й главе излагаются аналитические аспекты решения задач на условный экстремум целевой функции от векторной переменной — либо зависимой от какого-нибудь аргумента (параметра), либо ограниченной каким-нибудь уравнением связи. Такой двоякий подход к наложению условий на исходную переменную отвечает принятым двум способам задания связного многообразия, вложенного в пространство координат. Кроме того, в этой же главе рассматриваются аналитические аспекты предельных методов решения задач на условный экстремум. Наглядно показана геометрическая взаимосвязь всех трёх направлений условной оптимизации (включая и классический метод множителей Лагранжа) с применением характеристических проекторов в 2-х матричных формах. Выведено характеристическое (вековое) алгебраическое уравнение для условных собственных значений матрицы Гессе.

В 3-й главе развит формальный анализ целевых скалярных функций от комплексных переменных. С применением этого анализа разработаны методы для безусловной и условной оптимизации целевых функций от одной или нескольких комплексных или от смешанных переменных — без увеличения размерности задачи как обычно вдвое.

В 4-й главе, с учётом научных пристрастий автора, даны в качестве приложений важные примеры решения экстремальных задач из общей и линейной алгебры. Как один из результатов отметим теорему о полных требованиях к коэффициентам алгебраического уравнения степени n для вещественности и положительности его корней. Она ставит точку в решении алгебраической проблемы, исследуемой ещё Рене Декартом.

В 5-й главе рассматриваются основные численные методы поиска экстремума целевой функции 0-го, 1-го и 2-го порядка от одной или от нескольких скалярных переменных. Отдельно изложены методы поиска условного экстремума в 2-х вышеуказанных вариантах переменных.

В 6-й главе рассмотрены планово-вычислительные методы поиска экстремума целевой функции от нескольких скалярных переменных на основе её разностных моделей 1-го, неполного и полного 2-го порядка.

В 7-й главе рассмотрены планово-экспериментальные методы поиска экстремума целевой функции от нескольких скалярных переменных на основе её линейных регрессионных моделей, тождественных по форме разностным моделям неполного и полного 2-го порядка.

В 6-й и 7-й главах применяются специальные планы расположения точек в координатном пространстве для разностных или регрессионных оценок расчётных характеристик — градиента и матрицы Гессе.

Методы оптимизации излагаются последовательно в естественном порядке, подчёркивающим их генетическую взаимосвязь.

Особо также отметим то, что в данной монографии для аналитических преобразований и формул автор придерживается, по сути, классических обозначений Лейбница. Разумеется, делаются минимальные поправки при операциях с векторными переменными, например, для обозначения тензор-производных. (Преобразования осуществляются по правилам линейной алгебры.) Подобные обозначения весьма наглядны и дают возможность существенно сократить объём, занимаемый формулами.

Используемый в монографии фундаментальный материал изложен в литературе учебного и справочного характера [1, 3, 4, 7, 8, 12, 14, 15, 17, 19, 21 - 25, 29 - 31, 33, 38, 39, 41, 43, 44, 46, 47, 49, 50, 52, 53, 59, 62]. В коллаже на обложке книги иллюстративно используется изображение с произведения «Мыслитель» Огюста Родена. В порядке исторической ретроспективы дан ряд ссылок на первопроходческие труды классиков, содержащие основополагающие идеи в области оптимизации.

Книга адресуется как математикам-профессионалам, так и широкой аудитории читателей, интересующихся фундаментальной и прикладной математикой. Автор надеется, что читатели почерпнут из публикуемого труда для себя много нового и интересного в областях теории и практики оптимизации целевых функций, а также в применении математических методов оптимизации при решении конкретных практических задач.

Используемые обозначения

1. Обозначения арифметических пространств

$\mathcal{R}$ — вещественная числовая ось,
 $\mathcal{A}^n$ — вещественное аффинное пространство размерности n ,
 $\mathcal{E}^n$ — вещественное евклидово пространство размерности n ,
 Φ^K — аффинное факторное пространство размерности K ,
 Φ^{K+1} — расширенное факторное пространство размерности $K+1$,
 $\mathcal{D}$ — область определения целевой функции,
 $\mathcal{I}$ — отрезок (закрытый интервал) на $\mathcal{R}$,
 $\mathcal{I}^n$ — прямоугольная арифметическая область в $\mathcal{A}^n$, $\mathcal{E}^n$ или Φ^K ,
 $\mathcal{C}$ — комплексная числовая ось,
 $\mathcal{E}^n$ — комплексное аффинное пространство размерности n ,
 $\mathcal{H}^n$ — комплексное эрмитово пространство размерности n .

2. Некоторые общие обозначения

C_n^t — биномиальные коэффициенты,
 $\text{rang } \dots$ — ранг матрицы $\dots$,
 p — уровень экстремума или перегиба (стационарности),
 s и $\mathbf{s}$ — точки стационарности на $\mathcal{R}$ и на $\mathcal{A}^n$ или $\mathcal{E}^n$,
 $s^\bullet$ и $\mathbf{s}^\bullet$ — точки экстремума (максимума или минимума) здесь же,
 s^+ и $\mathbf{s}^+$ — точки максимума здесь же,
 s^- и $\mathbf{s}^-$ — точки минимума здесь же,
 $s^\pm$ — точка стационарного перегиба на $\mathcal{R}$,
 $\mathbf{s}^\pm$ — точка стационарной седловины на $\mathcal{A}^n$ или $\mathcal{E}^n$,
 $\lambda, \boldsymbol{\lambda}$ и Λ — скалярный, векторный и матричный множители Лагранжа,
 $\mathcal{N} \rightarrow \infty$ и $\alpha \rightarrow 0$ — большой и малый параметры в предельных методах.

3. Обозначение матриц и матричных характеристик

A — прямоугольная либо $n \times q$, либо $m \times n$ -матрица,

A' — транспонированная матрица,

A^+ — квазиобратная от A матрица Мура — Пенроуза размера либо $q \times n$, либо $n \times m$, $(A')^+ = (A^+)'$;

$\langle \text{im } A \rangle$, $\langle \ker A \rangle$ и $\langle \text{im } A' \rangle$, $\langle \ker A' \rangle$ — образ и ядро матриц A и A' ;

для исходной $n \times q$ -матрицы A ($n > q$):

AA' и $A'A$ — внешняя и внутренняя гомомультимпликация для A ,

$\overleftarrow{AA'} = AA^+$ — проектор в $\mathcal{A}^n$ на $\langle \text{im } A \rangle$ параллельно

$\langle \ker A' \rangle \equiv \langle \ker A \rangle_L$ или ортопроектор в $\mathcal{E}^n$ на $\langle \text{im } A \rangle$,

$\overrightarrow{AA'} = [I - AA^+]$ — проектор в $\mathcal{A}^n$ на $\langle \ker A' \rangle \equiv \langle \ker A \rangle_L$ параллельно

$\langle \text{im } A \rangle$ или ортопроектор в $\mathcal{E}^n$ на $\langle \ker A' \rangle \equiv \langle \ker A \rangle_L$, $(\overleftarrow{AA'} + \overrightarrow{AA'} = I)$;

для исходной $m \times n$ -матрицы A ($n > m$):

$A'A$ и AA' — внешняя и внутренняя гомомультимпликация для A ,

$\overleftarrow{A'A} = A^+A$ — проектор в $\mathcal{A}^n$ на $\langle \text{im } A' \rangle \equiv \langle \text{im } A \rangle_D$ параллельно $\langle \ker A \rangle$
или ортопроектор в $\mathcal{E}^n$ на $\langle \text{im } A' \rangle \equiv \langle \text{im } A \rangle_D$,

$\overrightarrow{A'A} = [I - A^+A]$ — проектор в $\mathcal{A}^n$ на $\langle \ker A \rangle$ параллельно

$\langle \text{im } A' \rangle \equiv \langle \text{im } A \rangle_D$ или ортопроектор в $\mathcal{E}^n$ на $\langle \ker A \rangle$, $(\overleftarrow{A'A} + \overrightarrow{A'A} = I)$;

B — квадратная $n \times n$ -матрица,

C — несингулярная клеточная матрица ($\det C \neq 0$),

D — диагональная матрица, D_m — масштабная матрица ($\det D \neq 0$),

$\tilde{E}$ — матрица единичного базиса,

$G = G'$ — матрица Гессе,

H — эрмитова матрица, Q — косозермитова матрица,

I — единичная матрица, It — тотально-единичная матрица,

N — нормальная матрица,

P — простая матрица,

R — $n \times n$ -матрица ортогонального преобразования в $\mathcal{E}^n$,

S — симметричная матрица, K — кососимметричная матрица,

V — матрица линейного модального преобразования в $\mathcal{A}^n$,

Z — нулевая матрица.

4. Обозначения переменных и функций

x и $\mathbf{x}$ — независимые скалярная и $n \times 1$ -векторная переменные,
 $\tilde{x} : \tilde{x} = x(u)$ или $u = u(\tilde{x})$ — зависимая скалярная переменная,
 $\tilde{\mathbf{x}} : \tilde{\mathbf{x}} = \mathbf{x}(u)$ — зависимая $n \times 1$ -векторная переменная,
 $\tilde{\mathbf{x}} = \mathbf{L} = \mathbf{L}(u)$ — линейная зависимая $n \times 1$ -векторная переменная,
 где u — $q \times 1$ -векторная независимая переменная,
 $\tilde{\mathbf{x}} : \mathbf{h} = \mathbf{h}(\tilde{\mathbf{x}}) = \mathbf{0}$ — ограниченная $n \times 1$ -векторная переменная,
 $\tilde{\mathbf{x}} = \mathbf{L} : \mathbf{h} = \mathbf{h}(\mathbf{L}) = \mathbf{0}$ — линейная ограниченная $n \times 1$ -векторная переменная,
 где $\mathbf{h}$ — $m \times 1$ -векторная функциональная переменная,
 $y(x)$ и $y(\mathbf{x})$ — целевые вещественные скалярные функции от x и от $\mathbf{x}$,
 $\mathcal{L}(\mathbf{x})$ — целевая вещественная скалярная функция 1-го порядка от $\mathbf{x}$,
 $q(\mathbf{x})$ — вещественная скалярная функция 2-го порядка от $\mathbf{x}$,
 $k(\mathbf{x})$ — вещественная скалярная функция неполного 2-го порядка от $\mathbf{x}$,
 $\mathbf{z}(\mathbf{x})$ — комплексная функция от комплексной переменной $\mathbf{x}$,
 $\mathbf{w} \equiv (\mathbf{x}, \bar{\mathbf{x}})$ — $2n \times 1$ бинарная комплексная переменная,
 в частности, для одномерного аргумента $n = 1$,
 $\Phi_{\mathcal{M}}(\mathbf{x})$ и $\mathbf{b}_{\mathcal{M}}(\mathbf{x})$ — композиционная функция с большим параметром $\mathcal{M}$
 и её градиент,
 $\Phi_{\alpha}(\mathbf{x})$ и $\mathbf{b}_{\alpha}(\mathbf{x})$ — композиционная функция с малым параметром α
 и её градиент,
 $K_{\mathcal{M}}(\mathbf{x})$ — композиционная функция Куранта,
 $T_{\alpha}(\mathbf{x})$ — композиционная функция Тихонова,
 $\theta(\mathbf{x})$ — стабилизатор-функция в методе регуляризации Тихонова.

5. Обозначения дифференциалов, приращений, производных и интегралов, в том числе формальных

$dx = \Delta x$ — дифференциал, или приращение независимой переменной x ,
 dy, d^2y, d^py — 1, 2, p -е дифференциалы целевой функции y ,
 $\Delta y, \Delta^2y, \Delta^py$ — 1, 2, p -е приращения целевой функции y ,
 $d\tilde{x}, d^2\tilde{x}, d^q\tilde{x}$ — 1, 2, q -е дифференциалы зависимой переменной $\tilde{x}$,
 $\Delta\tilde{x}, \Delta^2\tilde{x}, \Delta^q\tilde{x}$ — 1, 2, q -е приращения зависимой переменной $\tilde{x}$,
 $d\mathbf{x}$ и ∂x_i — полный и частный дифференциалы независимой переменной $\mathbf{x}$
 или свободный полный и частный дифференциалы переменной $\tilde{\mathbf{x}}$,
 $d\tilde{\mathbf{x}}, d^2\tilde{\mathbf{x}}$ и $\partial \tilde{x}_i, \partial^2 \tilde{x}_i$ — условные полные и частные дифференциалы
 зависимой или ограниченной переменной $\tilde{\mathbf{x}}$,

$dy(\mathbf{x}=\mathbf{c})$, $d^2y(\mathbf{x}=\mathbf{c})$ — дифференциалы y , отсчитываемые от точки $\mathbf{c}$,

$\frac{dy}{d\mathbf{x}} = \mathbf{g}(\mathbf{x})$ — градиент функции $y(\mathbf{x})$, или 1-я тензор-производная, т. е. $1 \times n$ -вектор 1-ых частных производных,

$\frac{d^2y}{d\mathbf{x}'d\mathbf{x}} = \mathbf{G}(\mathbf{x}) = \frac{d\mathbf{g}}{d\mathbf{x}'} = \frac{d\mathbf{g}'}{d\mathbf{x}} = [\mathbf{G}(\mathbf{x})]'$ — матрица Гессе функции $y(\mathbf{x})$, или 2-я тензор-производная, т. е. $n \times n$ -матрица 2-ых частных производных,

$[dy]$ и $[d^2y]$ — условные 1-й и 2-й дифференциалы функции $y(\mathbf{x})$,

$\left[\frac{dy}{d\mathbf{x}} \right]$ и $\left[\frac{d^2y}{d\mathbf{x}'d\mathbf{x}} \right]$ — условные градиент и матрица Гессе для $y(\mathbf{x})$,

$\frac{d\tilde{\mathbf{x}}}{d\mathbf{u}}$, $\frac{d\mathbf{u}}{d\tilde{\mathbf{x}}}$ и $\frac{d\mathbf{h}}{d\tilde{\mathbf{x}}}$ — $n \times q$, $q \times n$ и $m \times n$ -матрицы Якоби векторных функций

$\tilde{\mathbf{x}} = \mathbf{x}(\mathbf{u})$, $\mathbf{u} = \mathbf{u}(\tilde{\mathbf{x}})$ и $\mathbf{h} = \mathbf{h}(\tilde{\mathbf{x}})$, или их 1-ые тензор-производные, т. е. двумерные матрицы частных производных,

$\frac{d^2\tilde{\mathbf{x}}}{d\mathbf{u}'d\mathbf{u}}$, $\frac{d^2\mathbf{u}}{d\tilde{\mathbf{x}}'d\tilde{\mathbf{x}}}$ и $\frac{d^2\mathbf{h}}{d\tilde{\mathbf{x}}'d\tilde{\mathbf{x}}}$ — 2-ые тензор-производные функций $\tilde{\mathbf{x}} = \mathbf{x}(\mathbf{u})$, $\mathbf{u} = \mathbf{u}(\tilde{\mathbf{x}})$ и $\mathbf{h} = \mathbf{h}(\tilde{\mathbf{x}})$, т. е. трёхмерные матрицы частных производных,

$\mathcal{R}e(\mathbf{x})$ — отношение Релея от векторного параметра $\mathbf{x}$,

$\mathcal{R}e(\mathbf{x}, \mathbf{y})$ — бинарное отношение Релея от векторных параметров $\mathbf{x}$ и $\mathbf{y}$,

$\mathbf{D}(\mathbf{x})$ — диагональная $n \times n$ -матрица однородных частных производных,

$\mathbf{p}(\mathbf{x}) = [\mathbf{G}(\mathbf{x})]^{-1} \cdot \mathbf{g}(\mathbf{x})$ — директивный вектор 2-го порядка,

$\mathbf{j}(\mathbf{x}) = [\mathbf{D}(\mathbf{x})]^{-1} \cdot \mathbf{g}(\mathbf{x})$ — директивный вектор неполного 2-го порядка,

$\mathbf{p}_\alpha(\mathbf{x})$ и $\mathbf{j}_\alpha(\mathbf{x})$ — регуляризованные по Тихонову те же векторы;

для функций от комплексных переменных $y(\mathbf{x}, \bar{\mathbf{x}}) \equiv y(\mathbf{w})$:

$d\mathbf{x}$ и $d\bar{\mathbf{x}}$ — формальные дифференциалы одномерных комплексных переменных $\mathbf{x}$ и $\bar{\mathbf{x}}$, $d\mathbf{w} \equiv (d\mathbf{x}, d\bar{\mathbf{x}})$;

$d\mathbf{x}$, ∂x_i и $d\bar{\mathbf{x}}$, $\partial \bar{x}_i$ — полные и частные формальные дифференциалы многомерных комплексных переменных $\mathbf{x}$ и $\bar{\mathbf{x}}$, $d\mathbf{w} \equiv (d\mathbf{x}, d\bar{\mathbf{x}})$;

$d\ldots$ и $\partial\ldots$ — формальные полный и частный дифференциалы функции $\ldots$,

для функций от комплексных переменных $p(\mathbf{x}, \bar{\mathbf{x}})$ и $\mathbf{p}(\mathbf{x}, \bar{\mathbf{x}})$:

$\int \ldots d\mathbf{x}$ и $\int \ldots d\bar{\mathbf{x}}$ — формальные интегралы от функций $\ldots$ по одномерным комплексным переменным $\mathbf{x}$ и $\bar{\mathbf{x}}$,

$\int \ldots d\mathbf{x}$ и $\int \ldots d\bar{\mathbf{x}}$ — формальные интегралы от функций $\ldots$ по многомерным комплексным переменным $\mathbf{x}$ и $\bar{\mathbf{x}}$.

6. Обозначения в плано-вычислительных и в плано-экспериментальных методах оптимизации

$y = y(\mathbf{x}) = y(x_1, x_2, \dots, x_n)$ — оптимизируемая целевая функция от $\mathbf{x}$,
 $y = \varphi(\mathbf{t}) = \varphi(t_1, t_2, \dots, t_n)$ — эта же функция от нормированной $\mathbf{t}(\mathbf{x})$,
 K — количество частных факторов u_i в модели линейной регрессии $\hat{y}$,
 N — количество точек в плане вычисления или эксперимента,
 квадрантный n -симплекс — план, в котором одиночные точки находятся
 на всех ортах в 1-м квадранте и одна в центре, $N = n + 1$,
 осевой n -крест — план, в котором точки находятся в вершинах креста,
 совмещённого с координатными осями, и одна в центре, $N = 2n + 1$,
 n -СКП — симметричный композиционный план, состоящий из осевого
 и координатно плоскостного n -креста, $N = 2n^2 + 1$,
 n -куб — план, в котором точки находятся в вершинах центрального куба,
 все грани которого перпендикулярны осям координат, $N = 2^n$,
 n -ЦКП — центральный композиционный план, состоящий из n -куба
 (или его дробной m -реплики) и осевого n -креста, $N = 2^m + 2n + 1$,
 ε — нормирующее отношение плеч осевого n -креста и сторон n -куба,
 m — количество опытов в отдельных точках плана,
 M — общее количество опытов во всех точках плана,
 $\tilde{y}(\mathbf{x})$ или $\tilde{y}(x_1, x_2, \dots, x_n)$ — оценка значения целевой функции отклика,
 имеющая только случайную ошибку измерения в плановых методах,
 $\hat{y}(\mathbf{x})$ или $\hat{y}(x_1, x_2, \dots, x_n)$ — разностная модель целевой функции и её
 оценка на области плана по данной модели,
 $\hat{g}, \hat{g}_i, \hat{g}_{ij}, \hat{g}_{ij}$ — оценки скалярных коэффициентов разностной модели;
 $\dot{y}(\mathbf{x})$ или $\dot{y}(x_1, x_2, \dots, x_n)$ — регрессионная модель целевой функции и
 её оценка на области плана по данной модели,
 $\dot{g}, \dot{g}_i, \dot{g}_{ij}, \dot{g}_{ij}$ — оценки скалярных коэффициентов модели регрессии;
 $\hat{\mathbf{g}}, \hat{G}, \hat{D}, \hat{\mathbf{p}}, \hat{\mathbf{j}}$ — разностные оценки данных характеристик,
 $\dot{\mathbf{g}}, \dot{G}, \dot{D}, \dot{\mathbf{p}}, \dot{\mathbf{j}}$ — регрессионные оценки данных характеристик,
 $u = u(\alpha \cdot \mathbf{r})$ — профиль функции при движении по вектору $\mathbf{r}$ в $\mathcal{A}^n$,
 u_i или v_i — факторы, образуемые из x_i , при линейной регрессии $\hat{y}$,
 $\mathbf{u} = \mathbf{u}(u_1, u_2, \dots, u_K)$, $\mathbf{v} = (1, \mathbf{u})$, причём $u_i = v_i$ при $i \neq 0$, $u_0 = 0$, $v_0 = 1$,
 $\hat{\mathbf{b}} = (\hat{b}_0, \hat{b}_1, \hat{b}_2, \dots, \hat{b}_K)$ — коэффициенты линейной регрессии $\hat{y}|\mathbf{u}$,
 $\hat{\mathbf{h}} = (\hat{h}_0, \hat{h}_1, \hat{h}_2, \dots, \hat{h}_K)$ — коэффициенты линейной регрессии $\hat{y}|\mathbf{v}$.

7. Обозначения вероятностных и статистических характеристик

$\bar{\tilde{y}}$ — выборочное среднее арифметическое для случайных величин $\tilde{y}$,

$\bar{X}$ — среднее арифметическое для величин X ,

$\mathcal{P}\langle \dots \rangle$ — доверительная вероятность события $\langle \dots \rangle$,

μ и $\mu\{\dots\}$ — математическое ожидание, в т. ч. случайной величины $\{\dots\}$,

σ^2 и $\sigma^2\{\dots\}$ — квадратичная дисперсия, в т. ч. случайной величины $\{\dots\}$,

T — распределение Стьюдента,

T_c — критическое значение T ($T_{c(1)}$ и $T_{c(2)}$ — одно- и двустороннее),

χ^2 — распределение Пирсона,

F — распределение Фишера — Снедекора, F_c — критическое значение F ,

s^{*2} и s^2 — выборочная дисперсия, смещённая и несмещённая,

s — выборочное среднее квадратичное отклонение,

$s_r^2(\dot{y})$ — выборочная дисперсия воспроизводимости модели регрессии,

$s_{ad}^2(\dot{y}, \tilde{y})$ — выборочная дисперсия адекватности модели регрессии,

k — число степеней свободы данной статистики (случайной величины),

$v(y|x) = \mu\{v(\tilde{y}|x)\}$ — плановая ковариация $\tilde{y}$ и x при регрессии $\tilde{y}|x$,

$v(\tilde{y}|x)$ — выборочная ковариация $\tilde{y}$ и x при регрессии $\tilde{y}|x$,

$\Delta^2\langle x \rangle = \mu\{\Delta^2\langle x \rangle\}$ — плановая дисперсия x при регрессии $\tilde{y}|x$,

$\Delta^2\langle \tilde{y} \rangle = \mu\{s^2(\tilde{y}|x)\}$ — плановая дисперсия $\tilde{y}$ при регрессии $\tilde{y}|x$,

$s^2(\tilde{y}|x)$ — выборочная дисперсия $\tilde{y}$ при регрессии $\tilde{y}|x$,

$k_{(x, \tilde{y})}$ — выборочный коэффициент линейной корреляции $\tilde{y}$ и x ,

U или V — матрица планирования, W — информационная матрица,

$C = W^{-1} \sigma^2\{\tilde{y}\}$ — дисперсионно-ковариационная матрица,

$\sigma^2(\dot{h}_p)$ — плановая дисперсия коэффициента $\dot{h}_p$,

$v(\dot{h}_p, \dot{h}_s)$ — плановая ковариация коэффициентов $\dot{h}_p$ и $\dot{h}_s$,

$\lambda(\dot{h}_p, \dot{h}_s)$ — плановый коэффициент линейной корреляции $\dot{h}_p$ и $\dot{h}_s$,

$d[\dot{y}(x)]$ — общая ошибка модели в плане эксперимента,

$\Delta[\dot{y}(x)]$ — систематическая ошибка модели (плановое смещение),

$\delta[\dot{y}(x)]$ — случайная ошибка модели (случайное отклонение),

$\tilde{\gamma}$ и $\hat{\gamma}$ — относительные вклады случайной и систематической ошибок,

ξ — относительная ошибка модели целевой функции, %.

8. Используемые символы

- $'$ — знак простого транспонирования,
 $*$ — знак эрмитового транспонирования,
 $\subset$ — множество ... принадлежит множеству ...,
 $\subseteq$ — множество ... принадлежит или тождественно множеству ...,
 $\in$ — элемент ... принадлежит множеству ...,
 $\notin$ — элемент ... не принадлежит множеству ...,
 $\cup$ — знак объединения множеств,
 $\equiv$ — знак тождества множеств,
 $\langle \dots \rangle$ — множество элементов ...,
 $\wedge$ — предикат «... и ...», $\vee$ — предикат «... или ...»,
 $\Rightarrow$ или $\rightarrow$ — предикат «из ... следует ...»,
 $\Leftarrow$ или $\leftarrow$ — предикат «... следует из ...»,
 $\Leftrightarrow$ или $\leftrightarrow$ — предикат «... равносильно ...»,
 $>$ — больше, $\geq$ — больше или равно,
 $<$ — меньше, $\leq$ — меньше или равно,
 $\forall$... — любой элемент ...,
 $[...]$ — целая часть числа ...,
 $|...|$ — абсолютное значение (модуль) числа ...,
 $\|... \|^$ — модуль вектора ...,
 $\begin{vmatrix} \dots \end{vmatrix}$ — вектор или матрица из элементов ...,
 Σ — знак алгебраического суммирования,
 $\oplus$ — знак прямого суммирования,
 $\div$ — знак задаваемого интервала,
*** — знак окончания доказательства.

Глава 1. Аналитическая безусловная оптимизация

§ 1.1. Экстремумы целочисленных уровней $p \geq 2$ для функций от независимой скалярной переменной

Пусть исходно в явном виде задана некая числовая вещественная скалярная функция $y = y(x)$. Причём, пока иное не оговорено особо, x есть *независимая* скалярная вещественная переменная, или аргумент. Для заданной функции $y(x)$ её аргумент может принимать множество допустимых значений на вещественной числовой оси $(-\infty, +\infty) \equiv \mathcal{R}$, на котором и определяется сама функция. Если это множество на оси $\mathcal{R}$ непрерывно-связное, то оно называется *областью определения* $\mathcal{D} \subseteq \mathcal{R}$ данной функции. Например, для степенной функции $y = x^2$ это есть сама числовая ось $(-\infty, +\infty)$, или множество всех действительных чисел; для корневой функции $y = \sqrt{x}$ это есть числовая полуось $[0, +\infty)$; но для функции $y = \arcsin x$ это есть отрезок $[-1, +1]$, а для корневой функции $y = \sqrt{-x^2}$ это есть изолированная точка 0 и так далее. В свою очередь, отображение $y(\mathcal{D})$ есть полное множество значений данной функции y от аргумента x . Для некоторой функции полное множество её значений может быть ограниченным, т. е. иметь верхнюю или/и нижнюю конечную грань. Подобные функции являются, как правило, основным предметом анализа в задачах оптимизации. Кроме того, в этом же аспекте особый интерес вызывают такие скалярные функции $y(x)$, которые на области своего определения $\mathcal{D} \subseteq \mathcal{R}$ не только ограниченные, но, вместе с тем, непрерывные и непрерывно дифференцируемые. Функции с подобными свойствами называются *эволюционными*. На области определения они и их первые производные не подвержены каким-либо скачкообразным изменениям при бесконечно малом изменении аргумента. Непрерывные функции на компакте (отрезке, или закрытом интервале) обязательно *равномерно непрерывные* (однако на открытом интервале переменной подобное утверждение не всегда верно). В свою очередь, равномерная непрерывность — весьма полезное свойство для разнообразных оценок при анализе поведения функции на выбираемой области аргумента.

Зачастую для наглядности или большей конкретности (например, при использовании численных методов) поведение функции изучают на некотором компактном подмножестве из области её определения. Это может быть некий закрытый интервал $[a, b] \equiv \mathcal{T} \subset \mathcal{R}$. Важно то, что на компакте, согласно известной **теореме Вейерштрасса** [17, 23], непрерывная функция всегда имеет свою верхнюю и нижнюю грани.

Пусть функция $y(x)$ ограничена на $\mathcal{D}$ либо сверху, либо снизу, т. е. или $y(x) \leq M$ или $y(x) \geq M$, где M — конечное число. Тогда $y(x)$ на некотором $\mathcal{T} \subset \mathcal{R}$ в некоторой точке $s^* \in \mathcal{T}$ в первом случае принимает максимальное значение $y(s^*)$ (*максимум*) и во втором случае принимает минимальное значение $y(s^*)$ (*минимум*). В их общем определении $y(s^*)$ есть экстремальное значение функции на интервале $\mathcal{T}$ (*экстремум*). Дальнейший интерес будет представлять только такой нетривиальный вариант экстремума, когда s^* не является крайней точкой интервала $\mathcal{T}$, а находится внутри его (т. е. именно *внутренний* экстремум). В данной монографии такие экстремумы и их поиск для изначально заданных эволюционных функций $y(x)$ представляют главный интерес. Поэтому для них будет применяться как основной термин *целевая функция* — для того чтобы, по сути, отличать таковую от других скалярных функций. Процедура поиска экстремума есть *оптимизация целевой функции*.

На переломном рубеже развития математической науки совершенно естественным образом возникла классическая задача об *аналитическом* методе определения и поиска экстремума скалярной функции. Такого рода задачу впервые чётко поставил и корректно разрешил Пьер Ферма для класса целостепенных скалярных функций (1629 г.) [58]. Позднее Готфрид Лейбниц дал аналитическое решение экстремальной задачи для дифференцируемых функций, включая трансцендентные функции (1684 г.) [61]. Так была основана *аналитическая оптимизация*.

Из курса анализа известно (см., например, [23, т. 1, с. 223]): если s^* является точкой внутреннего экстремума непрерывной функции $y(x)$, определённой хотя бы на некоторой её окрестности, то при этом 1-я производная $y'(s^*)$ либо равна нулю, либо не существует. Это суть общие необходимые условия существования такого рода экстремума для непрерывной функции $y(x)$. Однако именно первое из этих двух альтернативных необходимых условий в аспекте рассматриваемой задачи применяется к эволюционным целевым функциям, для которых возможны только *неособые* (т. е. плавные) экстремумы. В анализе эта закономерность известна как **лемма Ферма** о необходимом условии реализации внутреннего экстремума непрерывно дифференцируемой скалярной функции от одной *вещественной* скалярной переменной.

В более широком смысле то же самое условие для эволюционных функций $y(x)$, но как уравнение $dy/dx = 0$, позволяет в результате его решения выявить точку s (или подмножество точек) *стационарности* $y(x)$ на области её определения $\mathcal{D} \subseteq \mathcal{R}$. В анализе это обосновывает **теорема Ферма** о необходимом и достаточном условии существования стационарности непрерывно дифференцируемой скалярной функции от вещественной скалярной переменной. Оно выражается аналитически любым из указанных ниже тождественных уравнений — либо через 1-й дифференциал, либо через 1-ю производную функции в точке $s \in \mathcal{R}$:

$$dy(x=s) = 0 \quad \leftrightarrow \quad \frac{dy}{dx}(s) = 0. \quad (1)$$

В этой точке целевая функция имеет *стационарное значение* $y(s)$. Стационарность эволюционной функции $y(x)$ по форме её графика в окрестности точки s может представлять собой либо искомый *неособый экстремум* (максимум или минимум), либо *стационарный перегиб*.

Кроме того, эти же тождественные уравнения (1), как указывалось, согласно лемме Ферма, задают аналитически и необходимое условие существования в точке $s \in \mathcal{T}$ неособого экстремума для эволюционной целевой функции $y(x)$. Причём *строгий* экстремум $y(s^*)$ определяется тем, что в точке s^* эта целевая функция имеет экстремальное значение, а в любой достаточно малой окрестности точки s^* — не имеет.

На рис. 1(1)—(4) приведены наглядно характерные абстрактные примеры непрерывных скалярных функций $y(x)$ — как эволюционных, так и нет, с разнообразными экстремумами и перегибами на интервале $\mathcal{T} \subset \mathcal{R}$. В табл. 1 приведена классификация экстремумов и перегибов на основе наиболее ныне распространённой терминологии. Вариант (1) тут отвечает именно эволюционным функциям $y(x)$, представляющим наибольший интерес в прикладном отношении. Отметим при этом, что для многоэкстремального (полимодального) варианта (2) в отличие от одноэкстремального (унимодального) варианта (1) для кривых $y(x)$ любые используемые аналитические методы и базирующиеся на них численные методы нахождения экстремума эволюционной функции позволяют выявить и далее идентифицировать отдельным образом лишь какой-либо её *локальный* экстремум или последовательным образом — все таковые. Среди последних выбирают соответственно наибольший или наименьший (как *глобальный* экстремум). Очевидно, что максимум в варианте (2) на рис. 1(2) есть глобальный экстремум функции $y(x)$ на заданном интервале $\mathcal{T} \subset \mathcal{R}$ аргумента x или даже, быть может, на всей области её определения, что, однако, требует отдельного доказательства.

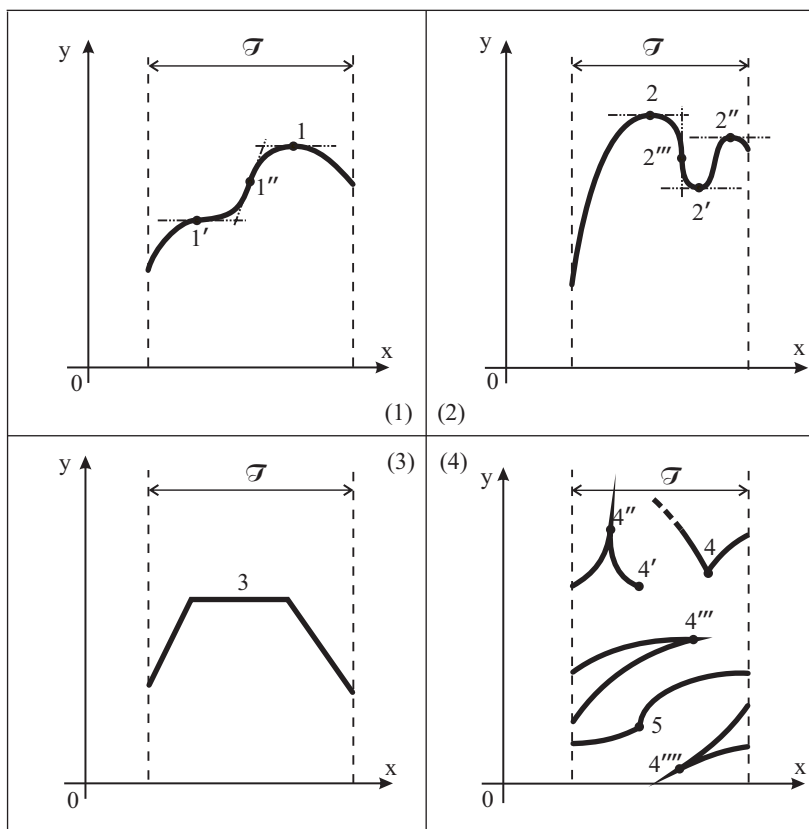


Рис. 1. Абстрактные варианты непрерывных функций от скалярного аргумента $y(x)$ с разнообразными типами экстремумов и перегибов на числовом интервале $\mathcal{T} \subset \mathcal{R}$:

- (1), (2) — функции с *неособыми* экстремумами (1, 2, 2', 2'') и перегибами (1', 1'', 2''', 2'');
 (3), (4) — функции с *особыми* экстремумами (3, 4, 4', 4'', 4''', 4''') и перегибами (5).

Кроме этого, на рис. 1 можно наглядно проследить различие между гладкими и негладкими кривыми (с одной стороны), эволюционными и неэволюционными функциями (с другой стороны). Само же понятие *гладкость* для вложенного многообразия, в том числе и для кривой, является абсолютным по отношению к выбору координатной системы, в данном случае на координатной плоскости $\langle x, y \rangle$. Обе верхние кривые на рис. 1(1) и (2) суть гладкие, кривые на рис. 1(3) и (4) — негладкие.

Далее пусть целевая функция $y(x)$ — непрерывная на заданном интервале $[a, b] \equiv \mathcal{T} \subset \mathcal{R}$ и при этом она на нём, по крайней мере, дважды непрерывно дифференцируемая. При таких ещё более сильных тут допущениях необходимые и достаточные условия существования в некоей внутренней точке этого интервала $s^* \in \mathcal{T}$ неособого строгого экстремума целевой функции $y(x)$ 2-го уровня формулируются в виде классических **правил Лейбница** отдельно для случаев её максимума и минимума того же *целочисленного* 2-го уровня:

$$\frac{dy}{dx}(s^+) = 0, \quad \frac{d^2y}{dx^2}(s^+) < 0; \quad (2)$$

$$\frac{dy}{dx}(s^-) = 0, \quad \frac{d^2y}{dx^2}(s^-) > 0. \quad (3)$$

Точки экстремума s^* обозначаются как s^+ для максимума и как s^- для минимума, чтобы их можно было достаточно легко различать. Точки же перегибов обозначаются как $s^\pm$ для g -образного и как $s^\mp$ для s -образного *стационарного перегиба*. Понятие *уровень* экстремума или стационарного перегиба (т. е. стационарности вообще) в общем случае определяет тот же порядок касательной или соприкасающейся в точке s собственной параболы от аргумента x . Последняя здесь и в дальнейшем определяется как *аппроксимирующая* парабола для $y(x)$, что затем будет отдельно пояснено. (В рассматриваемых пока случаях её порядок и уровень самой стационарности *целочисленный*, т. е. 2.) Для экстремумов целевой скалярной функции 2-го уровня аппроксимирующая парабола имеет тот же порядок 2. Графически она есть касательная в точке s параболическая кривая 2-го порядка — либо u -образная при $d^2y/dx^2 > 0$, либо n -образная при $d^2y/dx^2 < 0$:

$$w(x) = \frac{1}{2} \cdot \frac{d^2y}{dx^2}(s) \cdot (x - s)^2 + y(s).$$

Графически подобные варианты экстремумов в абстрактной форме изображены на рис. 2(1) и (2). Максимум и минимум 2-го уровня, определяемые уравнениями (1) и (2), суть *простейшие* типы неособых экстремумов для целевой функции от одной скалярной переменной и при этом они же наиболее значимые в прикладном смысле.

В свою очередь, для *нестрогих* экстремумов точка s заменяется подмножеством $\langle s \rangle$, а в правых уравнениях в (2) и (3) знаки неравенств типа $<$ и $>$ заменяются на знаки неравенств типа $\leq$ и $\geq$.

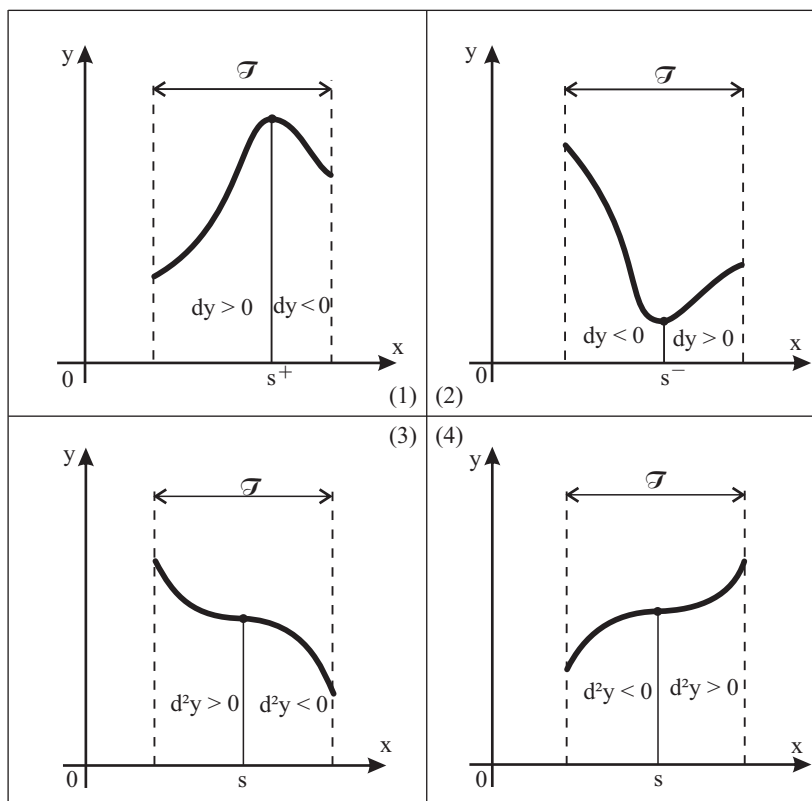


Рис. 2. Характерные варианты строгой стационарности непрерывно дифференцируемых функций $y(x)$ с экстремумами и стационарными перегибами на числовом интервале $\mathcal{J} \subset \mathcal{R}$:

- (1) — максимум 2-го уровня,
- (2) — минимум 2-го уровня,
- (3) — s -образный стационарный перегиб 3-го уровня,
- (4) — z -образный стационарный перегиб 3-го уровня.

Следующие возможные здесь по логике варианты — *более пологий* стационарности функции $y(x)$ и *целочисленного* уровня $p = 3$ задаются аналогичными уравнениям (1) дифференциальными уравнениями, но с повышением их порядка и степени вырожденности до двух:

$$dy(x = s) = d^2y(x = s) = 0 \quad \leftrightarrow \quad \frac{dy}{dx}(s) = \frac{d^2y}{dx^2}(s) = 0.$$

Такие варианты для идентификации стационарности требуют уже, по крайней мере, трёхкратной непрерывной дифференцируемости $y(x)$.

Отсюда имеем очередные два варианта стационарности:

$$\frac{dy}{dx}(s^{\pm}) = \frac{d^2y}{dx^2}(s^{\pm}) = 0, \quad \frac{d^3y}{dx^3}(s^{\pm}) < 0 \quad (4)$$

для случая *s*-образного стационарного перегиба 3-го уровня — рис. 2(3);

$$\frac{dy}{dx}(s^{\mp}) = \frac{d^2y}{dx^2}(s^{\mp}) = 0, \quad \frac{d^3y}{dx^3}(s^{\mp}) > 0 \quad (5)$$

для случая *z*-образного стационарного перегиба 3-го уровня — рис. 2(4).

Следующие затем два возможных варианта — *ещё более пологий* стационарности $y(x)$ и *целочисленного* уровня $p = 4$ требуют, по крайней мере, четырёхкратной непрерывной дифференцируемости $y(x)$ и задаются поэтому дифференциальными уравнениями со степенью вырожденности 3:

$$\frac{dy}{dx}(s^+) = \frac{d^2y}{dx^2}(s^+) = \frac{d^3y}{dx^3}(s^+) = 0, \quad \frac{d^4y}{dx^4}(s^+) < 0 \quad (6)$$

для случая максимума 4-го уровня;

$$\frac{dy}{dx}(s^-) = \frac{d^2y}{dx^2}(s^-) = \frac{d^3y}{dx^3}(s^-) = 0, \quad \frac{d^4y}{dx^4}(s^-) > 0 \quad (7)$$

для случая минимума 4-го уровня.

Данная цепь вырожденных дифференциальных уравнений может, в принципе, продолжаться и далее опять-таки каждый раз до первой ненулевой производной — либо чётного, либо нечётного порядка p с использованием аналогичных принятых подходов к идентификации стационарности целочисленного уровня p . Соответственно повышается вышеуказанный целочисленный уровень выявляемых экстремумов или стационарных перегибов. При всём этом здесь пока принимается, что производные в указанных правых неравенствах по величине *конечные*! Они суть *значимые*, т. е. ненулевые и ограниченные по величине.

Изложенные правила ступенчатой идентификации стационарности *конечного* уровня p действуют для целевых функций $y(x)$, по природе не менее чем p -кратно непрерывно дифференцируемых (хотя бы на некоторой окрестности искомой точки стационарности s). Эти хорошо известные правила выводятся достаточно строго через разложение таковых целевых функций по формуле Тейлора в виде степенного ряда с остаточным членом в форме Пеано. Но вначале с целью упрощения и для сравнения обратимся, хотя и к гораздо менее общему, но при этом более простому представлению *аналитических* целевых функций $y(x)$ полным степенным рядом Тейлора.

В окрестности какой-либо конкретной внутренней точки заданного интервала $c \in \mathcal{T}$ имеем разложение аналитической скалярной функции $y = y(x)$ по общей формуле Тейлора в виде полного степенного ряда, например, от дифференциала её аргумента dx , а также в итоге по порядковым дифференциалам или по порядковым приращениям самой скалярной функции:

$$\begin{aligned} y(c+dx) &= y(c) + \Delta y(x=c) = y(c) + \frac{dy}{dx}(c)dx + \\ &+ \frac{d^2y}{dx^2}(c)\frac{dx^2}{2!} + \frac{d^3y}{dx^3}(c)\frac{dx^3}{3!} + \frac{d^4y}{dx^4}(c)\frac{dx^4}{4!} + \dots = \\ &= y(c) + dy(x=c) + \frac{d^2y(x=c)}{2!} + \frac{d^3y(x=c)}{3!} + \frac{d^4y(x=c)}{4!} + \dots = \\ &= y(c) + \Delta^1 y(x=c) + \Delta^2 y(x=c) + \Delta^3 y(x=c) + \Delta^4 y(x=c) + \dots \end{aligned} \quad (8)$$

Здесь используется тот факт, что для *независимой* переменной имеет место $\Delta x = dx$, $d^2x = d^3x = d^4x = \dots = 0$, так как полное приращение аргумента, по определению, есть дифференциал dx . В частности, в (8) dx отсчитывается от точки c . Все же дифференциалы и приращения функции (полные и порядковые) отсчитываются от значения $y(c)$.

По Вейерштрассу $y(x)$ есть *аналитическая* функция в точке $c \in \mathcal{T}$, когда она, во-первых, имеет в ней производные любого порядка и, во-вторых, на некоторой её окрестности представима полным собственным сходящимся степенным рядом Тейлора порядка N [17, 23]:

$$y(x) = y(c) + \sum_{q=1}^N \frac{1}{q!} \cdot \frac{d^q y}{dx^q}(c) \cdot (x-c)^q, \quad (9)$$

где, в частности, $N \rightarrow \infty$. Понятие *аналитичность* для функции $y(x)$ далее естественным образом распространяется на заданный интервал $\mathcal{T}$ для аргумента x и вообще на всю область её определения. Очевидно, что для некоего функционального многочлена от x с целочисленными показателями его степеней число N равно максимальной степени x .

Более общо $y(x)$ есть, по крайней мере, *k раз дифференцируемая (непрерывно)* функция в точке $c \in (a, b)$, когда она имеет в ней все свои производные порядка от 1 до k и при этом на некоторой её окрестности представима собственным степенным рядом Тейлора порядка k именно с остаточным членом ряда в форме Пеано порядка малости более k (*определение k-кратно непрерывной дифференцируемости функции*):

$$y(x) = y(c) + \sum_{q=1}^k \frac{1}{q!} \cdot \frac{d^q y}{dx^q}(c) \cdot (x-c)^q + o_k(x-c), \quad (10)$$

$$\text{где } \lim_{x \rightarrow c} \frac{o_k(x-c)}{(x-c)^k} = 0.$$

(В частности, при $k = 1$ это есть известное определение непрерывной дифференцируемости, или непрерывности производной функции $y(x)$ по Фреше.) Свойство непрерывной дифференцируемости порядка k для функции $y(x)$ естественным образом распространяется и на заданный интервал для аргумента x и вообще на всю область её определения.

Если же функция $y(x)$ имеет ещё и производную порядка $k + 1$ в некоторой окрестности точки c , то остаточный член в (10) можно выразить более конкретно, например, в форме Лагранжа:

$$o_k(x-c) = \frac{1}{(k+1)!} \cdot \frac{d^{k+1} y}{dx^{k+1}}[\xi(x)] \cdot (x-c)^{k+1}, \quad x < \xi(x) < c. \quad (11)$$

(Для аналитической функции $y(x)$ это всегда имеет место, так как она, по определению, бесконечное число раз дифференцируемая.)

Возвратимся теперь к вопросу обоснования ранее уже изложенной ступенчатой идентификации стационарности $y(x)$. Пусть функция $y(x)$ не менее чем p -кратно непрерывно дифференцируемая, хотя бы в некоторой окрестности точки стационарности s . Причём в этой точке все её производные вплоть до порядка $q = p - 1$ нулевые, а производная порядка p ненулевая и конечная по величине (т. е. значимая). Отсюда в окрестности точки s , согласно (10), имеем представление $y(x)$ как

$$y(s+dx) = y(s) + \frac{1}{p!} \cdot \frac{d^p y}{dx^p}(s) dx^p + o_p(dx) = a_0 + a \cdot (dx)^p + o_p(dx). \quad (12)$$

Выделим *соприкасающуюся* с кривой $y(x)$ в точке стационарности s параболу порядка p :

$$w(x) = w(s+dx) = y(s) + \frac{1}{p!} \cdot \frac{d^p y}{dx^p}(s) dx^p = a_0 + a \cdot (dx)^p. \quad (13)$$

В достаточно малой окрестности точки s эта парабола, вместе с тем, является и *аппроксимирующей* для $y(x) \equiv y(s+dx)$, так как разность между ней и целевой функцией $y(s+dx) - y(s) = o_p(dx)$ есть функция порядка малости более p относительно $dx \rightarrow 0$.

Поведение графиков функций $y(s + dx)$ и $w(s + dx)$ в достаточно малой окрестности s исходя из локальной формы обеих кривых — одно и то же. Следовательно, о локальной форме графика $y(x) = y(s + dx)$ в окрестности точки её стационарности s можно судить по локальной форме в ней же аппроксимирующей параболы (13). Она может быть либо $\cap$ -образной (для максимума) или $\cup$ -образной (для минимума), либо z -образной или s -образной (для обоих перегибов). Таким образом, *правила ступенчатой идентификации стационарности доказаны*.

Например, в простейших случаях нижеуказанный функциональный степенной многочлен (с целочисленными показателями степеней)

$$y(x) = a_0 + a \cdot (x - s)^p + \sum_{(i > p)} a_i \cdot (x - s)^i \quad (14)$$

при $p = 2t$ и $a < 0$ ($a > 0$) в точке s имеет максимум (минимум) уровня p ; но при $p = 2t + 1$ и $a < 0$ ($a > 0$) в точке s имеет стационарный перегиб уровня p — либо s -образный, либо z -образный.

Очевидно, что любая функция $y(x)$, представляемая *конечным* рядом Тейлора, есть степенной многочлен от аргумента x с целочисленными показателями степеней и обратно.

Из вышеизложенного хорошо видно, что для p раз непрерывно дифференцируемой целевой функции $y(x)$, у которой в точке её стационарности производные порядка от 1 до $p - 1$ все нулевые, а производная порядка p ненулевая и причём конечная по величине, экстремумы и перегибы имеют всегда *целочисленный* уровень p — *чётный* для экстремумов и *нечётный* для стационарных перегибов. При этом тот же самый степенной параметр p имеют отвечающие им аппроксимирующие параболы. Но для эволюционных функций вообще возможны неособые экстремумы целочисленного *нечётного* уровня и стационарные перегибы целочисленного *чётного* уровня, а также — неособые экстремумы и стационарные перегибы *нецелочисленного* уровня (см. далее). Им соответствуют аппроксимирующие параболы тех же самых аномальных порядков. Вызвано всё это тем, что для эволюционных функций (по их определению) важны существование и непрерывность только 1-й производной!

Так, степенная функция вида

$$y(x) = a \cdot |x^{2t+1}| \quad (15)$$

при значении $a > 0$ имеет неособый строгий минимум уровня $p = 2t + 1$ в точке $s = 0$ и в ней же все нулевые производные вплоть до порядка $2t$ и высшую производную порядка $2t + 1$ с разрывом « $\mp a \cdot (2t + 1)!$ ».

С другой стороны, степенная функция вида

$$y(x) = a \cdot |x^{2t-1}| \cdot x \quad (16)$$

при значении $a > 0$ имеет неособый z -образный стационарный перегиб уровня $p = 2t$ в точке стационарности $s = 0$ и в ней же все нулевые производные вплоть до порядка $2t-1$ и высшую производную порядка $2t$ с разрывом « $\mp a \cdot (2t)!$ ».

Очевидно, что в рассмотренных выше двух простейших примерах аппроксимирующие в точке s параболы просто тождественны заданным степенным функциям. В частности, при $t = 1$ в первом примере имеем кубическую параболу с минимумом в s , а во втором примере имеем квадратичную параболу с z -образным стационарным перегибом в s .

* * *

Далее приведём примеры одномерной аналитической оптимизации.
(Много интересных примеров имеется в сборнике задач [1].)

Пример 1. Найти в интервале $(0, +\pi/2)$ максимум функции

$$y = a \cdot \sin \varphi + b \cdot \cos \varphi \quad (a > 0, b > 0).$$

Имеем: $y'(\varphi) = a \cdot \cos \varphi - b \cdot \sin \varphi = 0 \rightarrow \varphi^* = \arctg a/b$
(поскольку $b > 0$, то $\cos \varphi^* \neq 0$),

$$\begin{aligned} y''(\varphi^*) &= -(a \cdot \sin \varphi^* + b \cdot \cos \varphi^*) = -y(\varphi^*) = \\ &= -(a \cdot \tg \varphi^* + b) / \sqrt{1 + \tg^2 \varphi^*} = -\sqrt{a^2 + b^2} < 0. \end{aligned}$$

То есть в точке φ^* здесь имеем максимум $y(\varphi)$ со значением $\sqrt{a^2 + b^2}$.
В частности, при $a = b = 1$ имеем: $\varphi^* = \pi/4$, $y(\varphi^*) = \sqrt{2}$.

Пример 2. Найти для $x > 0$ минимум функции

$$y(x) = (x^n + x^{-n}) / 2.$$

Имеем: $y'(x) = n \cdot (x^{(n-1)} - x^{-(n+1)}) / 2 = 0 \rightarrow x^* = +1$,

$$y''(x^*) = n \cdot \left[(n-1) \cdot x^{(n-2)} + (n+1) \cdot x^{-(n+2)} \right] / 2 = n^2 > 0.$$

То есть минимум достигается в точке $x^* = 1$ со значением $y(x^*) = 1$.

Пример 3. Найти в интервале $(0, 2\pi)$ точку экстремума функции

$$y(x) = \int_0^x \frac{\sin x}{x} dx.$$

$$\text{Имеем: } y'(x) = \frac{\sin x}{x} = 0 \rightarrow x^\bullet = \pi, \quad y''(x^\bullet) = \frac{x^\bullet \cdot \cos x^\bullet - \sin x^\bullet}{x^{\bullet 2}} = -1/\pi < 0.$$

В точке $x^\bullet = \pi$ имеем максимум функции $y(x)$ со значением:

$$y(x^\bullet) = \int_0^\pi \frac{\sin x}{x} dx.$$

Здесь это рационально объясняется тем, что в данной точке $y'(x)$ под знаком интеграла попросту меняет знак. Отсюда в ней экстремум $y(x)$.

Пример 4. Найти и идентифицировать экстремум функции

$$y(x) = \frac{a_1 + a_2 + \dots + a_n + x}{n+1} - n+1 \sqrt[n]{a_1 \cdot a_2 \cdot \dots \cdot a_n \cdot x},$$

где $a_1, a_2, \dots, a_n, x > 0$.

$$\text{Преобразуем функцию к виду: } y(x) = \frac{n \cdot \alpha + x}{n+1} - n+1 \sqrt[n]{\gamma^n \cdot x},$$

где α и γ — средние арифметическое и геометрическое чисел $a_i > 0$.

Имеем:

$$y'(x) = 1/(n+1) \cdot \left[1 - \left(\gamma^n \cdot x \right)^{-n/n+1} \cdot \gamma^n \right] = 0 \rightarrow x^\bullet = \gamma.$$

При $\alpha > \gamma$ имеем $y(\gamma) > 0$, при $\alpha = \gamma$ имеем $y(\gamma) = 0$; $y''(x = \gamma) > 0$.
То есть при $x = \gamma$ в 1-м квадранте имеется глобальный минимум $y(x)$.

Пример 5. Найти и идентифицировать экстремум функции

$$y(x) = \frac{a_1 + a_2 + \dots + a_n + x}{n+1} / n+1 \sqrt[n]{a_1 \cdot a_2 \cdot \dots \cdot a_n \cdot x},$$

где $a_1, a_2, \dots, a_n, x > 0$.

$$\text{Преобразуем функцию к виду: } y(x) = \frac{n \cdot \alpha + x}{n+1} / n+1 \sqrt[n]{\gamma^n \cdot x}. \text{ Имеем:}$$

$$y'(x) = \gamma^{-n/n+1} / (n+1) \cdot \left[\left(x - \frac{n \cdot \alpha + x}{n+1} \right) / n+1 \sqrt[n]{x^{n+2}} \right] = 0 \rightarrow x^\bullet = \alpha.$$

При $\alpha > \gamma$ имеем $y(\alpha) > 1$, при $\alpha = \gamma$ имеем $y(\alpha) = 1$; $y''(x = \alpha) > 0$.
То есть при $x = \alpha$ в 1-м квадранте имеется глобальный минимум $y(x)$.

Полезность последних двух примеров состоит в том, что в каждом из них методом математической индукции далее доказывается известное алгебраическое неравенство Коши: $\alpha \geq \gamma$, где равенство имеет место тогда и только тогда, когда $a_1 = a_2 = \dots = a_n$.

Пример 6. Проанализировать на возможность наличия экстремумов или перегибов кривую 3-го порядка вида

$$y(x) = x^3 - k_1 \cdot x^2 + k_2 \cdot x - k_3, \text{ где } k_1, k_2, k_3 > 0, x \in \mathcal{R}.$$

$$\text{Имеем: } y'(x) = 3x^2 - 2k_1 \cdot x + k_2 = 0 \rightarrow x_{1,2} = \left(k_1 \pm \sqrt{k_1^2 - 3k_2} \right) / 3.$$

Соответственно имеем три варианта исходя из значения дискриминанта первой производной:

$$1) \Delta = k_1^2 - 3k_2 = 0 \rightarrow k_2 = k_1^2 / 3,$$

$x_1 = x_2 = s = k_1 / 3$, $y(s) = (k_1 / 3)^3 - k_3$, $y'(s) = y''(s) = 0$, $y'''(s) = 6 > 0$, что отвечает точке $\mathcal{Z}$ -образного стационарного перегиба.

Данная кривая с аргументом x на вещественной числовой оси $\mathcal{R}$ представляет собой $\mathcal{Z}$ -образную кубическую параболу, смещённую от центра координат вправо на s и вверх на $y(s)$.

$$2) \Delta = k_1^2 - 3k_2 < 0 \rightarrow k_2 > k_1^2 / 3,$$

$$x_{1,2} = \left(k_1 \pm i \cdot \sqrt{-\Delta} \right) / 3; \quad y'(x_{1,2}) = 0, \quad y''(x_{1,2}) = \pm 2i \cdot \sqrt{-\Delta}.$$

Данная кривая на вещественной числовой оси $\mathcal{R}$ не имеет точки стационарности, но имеет точку наклонного перегиба при $x = k_1 / 3$.

$$3) \Delta = k_1^2 - 3k_2 > 0 \rightarrow k_2 < k_1^2 / 3, \quad x_{1,2} = \left(k_1 \pm \sqrt{\Delta} \right) / 3 = s_{1,2};$$

$$x_{1,2} = \left(k_1 \pm \sqrt{\Delta} \right) / 3 = s_{1,2}; \quad y'(s_{1,2}) = 0, \quad y''(s_1) = +2\sqrt{\Delta} > 0, \quad y''(s_2) = -2\sqrt{\Delta} < 0.$$

Данная кривая с аргументом x на вещественной числовой оси $\mathcal{R}$ имеет здесь максимум 2-го уровня в точке $s_1 = \left(k_1 + \sqrt{k_1^2 - 3k_2} \right) / 3$ с аппроксимирующей параболой 2-го порядка $w(x) = y(s_1) + \sqrt{\Delta} \cdot (x - s_1)^2$ и имеет здесь минимум 2-го уровня в точке $s_2 = \left(k_1 - \sqrt{k_1^2 - 3k_2} \right) / 3$ с аппроксимирующей параболой 2-го порядка $w(x) = y(s_2) - \sqrt{\Delta} \cdot (x - s_2)^2$; где $s_2 < s_1$.

На данном примере (вариант 2) весьма наглядно интерпретируется тот факт, что при решении задач на экстремум скалярной функции $u(x)$ в их обычном понимании обнуление 1-й производной по лемме Ферма допускается только при *вещественных* значениях её аргумента. Хотя впоследствии в гл. 3 будут рассматриваться и задачи оптимизации на комплексных областях сопряжённых аргументов, где требование к вещественности аргумента в общем случае не будет выполняться.

Кроме того, вышерассмотренные примеры, как видно, имеют чисто математическое содержание, т. е. заданы как бы абстрактно. Подобные задачи, например, могут возникать при теоретическом анализе каких-то математических проблем. Развёрнутые примеры этого рода будут даны в гл. 4. К реальной действительности, разумеется, более близки задачи, возникающие из практики и сформулированные в конкретных областях науки: в физике, химии, экономике и т. д. Их аналитическое разрешение иногда приводило к открытию *новых* закономерностей в природе.

Пример 7. В заключение приведём исторически первый классический пример успешного аналитического разрешения задачи на экстремум, осуществлённого первопроходцем в данной области математики Пьером Ферма. Изложим доказательство в той строгой аналитической форме, которую ей придал Лейбниц в своей знаменитой статье [61]. Здесь он впервые успешно справлялся с дифференцированием функций с радикалами (подкоренными выражениями). Этот пример, к тому же, убедительно показывает, как аналитическая постановка и разрешение подобных экстремальных задач в итоге способствует выдвижению и обоснованию важнейших физических принципов природы.

Как известно, в 1621 г. Снеллиус экспериментальным путём открыл закон преломления света на границе раздела двух прозрачных сред (изначально воздуха и некоего прозрачного диэлектрика):

$$\sin \alpha_1 / \sin \alpha_2 = n = \text{const} (n > 1),$$

где α_1 — угол падения луча света в воздушной среде, α_2 — угол преломления света в другой оптически прозрачной среде. Однако теоретического обоснования данный закон природы до этой физико-математической работы Ферма не имел. С целью же аналитического обоснования этого закона Ферма выдвинул физический *экстремальный принцип*, что луч света в оптически прозрачной, но неоднородной среде при своём распространении из точки 1 в точку 2 (рис. 3) выбирает такую траекторию, чтобы затраченное время его движения было всегда минимальным. При этом Ферма интуитивно предположил, что скорость света при переходе его из воздуха в иную оптически прозрачную среду уменьшается, т. е. $v < c$.

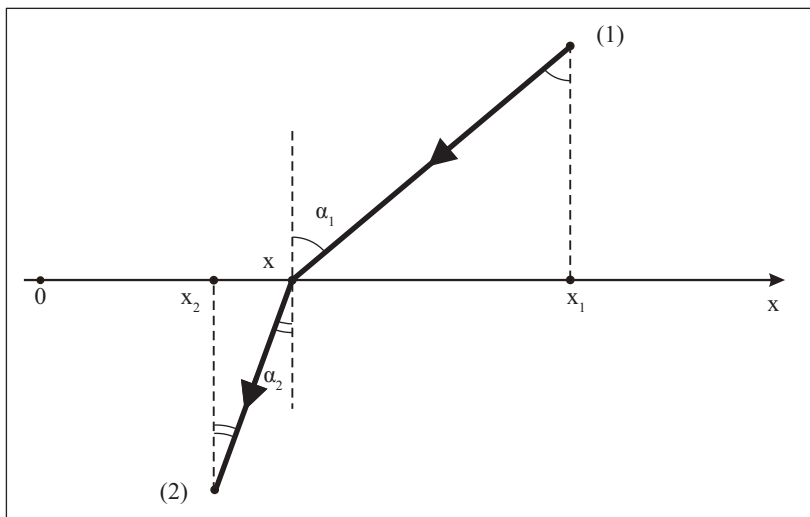


Рис. 3. К аналитическому обоснованию закона преломления света на границе оптически прозрачных сред Снеллиуса (Ферма, Лейбниц):

- (1) — исходная точка, α_1 — угол падения луча света;
 (2) — конечная точка, α_2 — угол преломления луча света.

Обратимся к рис. 3. Из него здесь непосредственно видно, что время движения света из точки 1 в точку 2 составляет:

$$t(x) = \frac{\sqrt{(x_1 - x)^2 + d_1^2}}{c} + \frac{\sqrt{(x - x_2)^2 + d_2^2}}{v}.$$

Применим далее теорему Ферма при поиске минимума функции $t(x)$ на интервале (x_2, x_1) :

$$\frac{dt}{dx} = -\frac{x_1 - x}{c \cdot \sqrt{(x_1 - x)^2 + d_1^2}} + \frac{x - x_2}{v \cdot \sqrt{(x - x_2)^2 + d_2^2}} = 0,$$

$$\text{или } \frac{\sin \alpha_1}{c} - \frac{\sin \alpha_2}{v} = 0, \text{ или } \frac{\sin \alpha_1}{\sin \alpha_2} = \frac{c}{v} = n > 1.$$

Таким образом, коэффициент преломления в законе Снеллиуса и сам закон получили полное обстоятельное аналитическое и физическое объяснение. Что особенно важно, в теоретической физике со временем укоренился один из первых фундаментальных законов природы — **экстремальный принцип Ферма**.

§ 1.2. Экстремумы нецелочисленных уровней $p > 1$

Рассмотрим функциональный степенной многочлен вида:

$$\begin{aligned} y(x) &= a_0 + a \cdot |x-s|^p + \sum_{(i>p)} a_i \cdot |x-s|^i \equiv \\ &\equiv y(s+dx) = a_0 + a \cdot |dx|^p + o_p(dx), \end{aligned} \quad (17)$$

где $p > 1$ — нецелочисленный показатель степени.

Пусть $[p]$ — его целая часть. Тогда $y(x)$ есть $[p]$ раз непрерывно дифференцируемая на множестве $\mathcal{R}$ функция от аргумента x . В точке s функция обязательно имеет стационарность, так как $dy/dx(s) = 0$. Более того, в этой точке все производные вплоть до порядка $[p]$ нулевые, но производная порядка $[p+1]$ бесконечная. Последняя же в точке s равна либо « $-\infty$ » при $a < 0$, либо « $+\infty$ » при $a > 0$. В (17) фигурирует остаточный член в форме Пеано, что и в ряде (10), порядка малости относительно $dx \rightarrow 0$ более p :

$$\begin{aligned} o_p|x-s| &= \sum_{(i>p)} a_i \cdot |x-s|^i \equiv \sum_{(i>p)} a_i \cdot |dx|^i, \\ \lim_{x \rightarrow s} \frac{o_p|x-s|}{|x-s|^p} &= \lim_{dx \rightarrow 0} \frac{o_p|dx|}{|dx|^p} = 0. \end{aligned}$$

Выделим касательную в точке s параболу той же нецелочисленной степени p :

$$w|s+dx| = a_0 + a \cdot |dx|^p, \quad (18)$$

где $p > 1$. В достаточно малой окрестности точки s эта парабола является аппроксимирующей для $y(x)$. Следовательно, по её локальной форме в окрестности точки стационарности и в данном случае также можно судить о локальной форме в ней же кривой (17). Она может быть либо $\cap$ -образной — для максимума $y(x)$, либо $\cup$ -образной — для минимума $y(x)$. Формально в данном случае тоже выполняются общие правила ступенчатой идентификации стационарности. Но тут вслед за цепью из $[p]$ нулевых производных сразу же идёт бесконечная производная порядка $[p+1]$, а именно « $-\infty$ » для максимума и « $+\infty$ » для минимума. В варианте степенного многочлена типа (17) теперь вполне очевидно, что в точке s его экстремум имеет нецелочисленный уровень p .

В свою очередь, вещественный степенной многочлен типа (14) при нецелочисленном уровне p и, возможно, нецелочисленных i требует определённой корректности в значениях этих показателей степени.

Во-первых, они должны быть *рациональными* числами, например, $p = m/n$, где $m > n$. Во-вторых, m и n в рациональных дробях должны быть взаимно простыми и n должен быть нечётным. При этих условиях $y(x)$ обязательно будет вещественной скалярной степенной функцией, определенной на *всей* оси $\mathcal{R}$. Для такой степенной функции возможны и экстремумы, и стационарные перегибы обоих типов. Идентификация типа стационарности нецелочисленного уровня $p > 1$ для такого вида функционального многочлена осуществляется опять-таки по общим правилам и с той же вышеуказанной особенностью: за цепью из $[p]$ нулевых производных сразу же идёт бесконечная производная порядка $[p + 1]$. При этом для эволюционной целевой функции, естественно, снимается требование по *непрерывной* дифференцируемости порядка именно p , так как последнее — целое число.

§ 1.3. Маргинальные ($1 < p < 2$) и особые ($p \leq 1$) экстремумы

Особые экстремумы, как известно [1], реализуются в таких точках — рис.1(4), в которых некая функция $y = y(x)$, принимая экстремальное значение, терпит излом в форме ряда разнообразных пиков: острых или неострых, вертикальных или наклонных или даже горизонтальных, односторонних или двусторонних (относительно общей касательной) — рис. 1(4). Так, последние два варианта реализуются в точках возврата (заострения) 2-го и 1-го рода. Однако такие экзотические экстремумы представляют больше теоретический интерес, нежели практический. Они рассматриваются в рамках теории особых точек плоских кривых.

Вообще в теории экстремумов целевых скалярных функций от одной вещественной переменной используется наглядная аналогия графика целевой функции $y(x)$ в *декартовых* координатах с профилем горной географической местности с разнообразными вершинами и впадинами. Указанные объекты природы суть либо округлые (т. е. как неособые экстремумы), либо пикообразные (т. е. как особые экстремумы).

Кривизна и радиус касательной окружности для кривой $y(x)$ в точке экстремума 2-го уровня s^* в декартовых координатах выражается весьма простой формулой:

$$K(s^*) = 1 / R(s^*) = \frac{d^2 y}{dx^2}(s^*). \quad (19)$$

Знак кривизны определяется знаком 2-й производной (или характером экстремума 2-го уровня). Для экстремумов более высокого уровня, т. е. при $p > 2$, имеем $K = 0$ ($R = +\infty$ или $-\infty$). Для неособых экстремумов более низкого уровня ($1 < p < 2$) имеем $K = +\infty$ или $-\infty$ ($R = 0$).

Определим неособые экстремумы уровня в интервале $1 < p < 2$ как *маргинальные*, поскольку они занимают крайнее нижнее положение в их иерархии (по значениям уровня p) и соответственно примыкают к множеству особых экстремумов.

Пусть $u(x)$ есть *однократно* непрерывно дифференцируемая в точке $s \in \mathcal{T}$ целевая функция, но при этом в ней её 1-я производная нулевая, а 2-я производная равна или « $-\infty$ » или « $+\infty$ ». Соответственно для этой функции снимается требование по *непрерывной* дифференцируемости. Тогда, согласно правилам Лейбница (1, 2), имеем в первом случае маргинальный максимум, во втором случае маргинальный минимум в точке стационарности s . Например, степенные параболы общего вида

$u(x) = a \cdot \sqrt[n]{|x|}^m$, где n и m взаимно простые числа, удовлетворяющие неравенству $n < m < 2n$ (т. е. иначе при $1 < p < 2$), суть простейшие алгебраические кривые, иллюстрирующие маргинальные экстремумы указанного уровня p . Ввиду того, что эти кривые разделяют множества *обычных* гладких ($p \geq 2$) и негладких ($p \leq 1$) парабол, то они называются тут и далее аналогичным образом как *маргинальные параболы* (дробно-рациональной степени $1 < p < 2$). Эти параболы не теряют гладкость 1-го порядка, но в точке экстремума просто имеют более крутой изгиб. Разумеется, более общее понятие — *маргинальная стационарность*.

* * *

Несколько простейших примеров функций $u(x)$ с маргинальными и особыми экстремумами проиллюстрированы конкретными примерами, графически отображёнными на рис. 4.

Пример 1. Прямая парабола Нейля. Она же — гладкая полукубическая парабола ($a = +1$, $n = 2$, $m = 3$) — рис. 4(1):

$$u = \sqrt{|x|^3}, \quad u'(x) = \pm 3/2 \cdot \sqrt{|x|}, \quad u''(x) = +3/4 \cdot \sqrt{|x|},$$

где для $u'(x)$ выбирают знак «+» при $x > 0$ и знак «−» при $x < 0$.

Указанная функция в точке $s = 0$ имеет маргинальный минимум уровня $p = 3/2$, так как $u'(0) = 0$, $u''(0) = +\infty$.

Все маргинальные параболы $y = \sqrt[n]{|x|^m}$ с минимумом в точке $s = 0$ расположены между параболой $y = x^2$ и линией $y = |x|$ — рис. 4(1). Они имеют 3 общие точки: $\{-1, +1\}$, $\{0, 0\}$ и $\{+1, +1\}$. Граничная линия $y = |x|$ имеет в $s = 0$ особый минимум уровня $p = 1$. Маргинальные экстремумы ($1 < p < 2$) в сравнении с обычными ($p \geq 2$) имеют большее искривление графика $y(x)$ именно в окрестности точки стационарности. В самой точке s кривизна кривой по величине бесконечная. При $p = m/n \rightarrow 1$ маргинальная парабола неуклонно приближается к линии $y = |x|$.

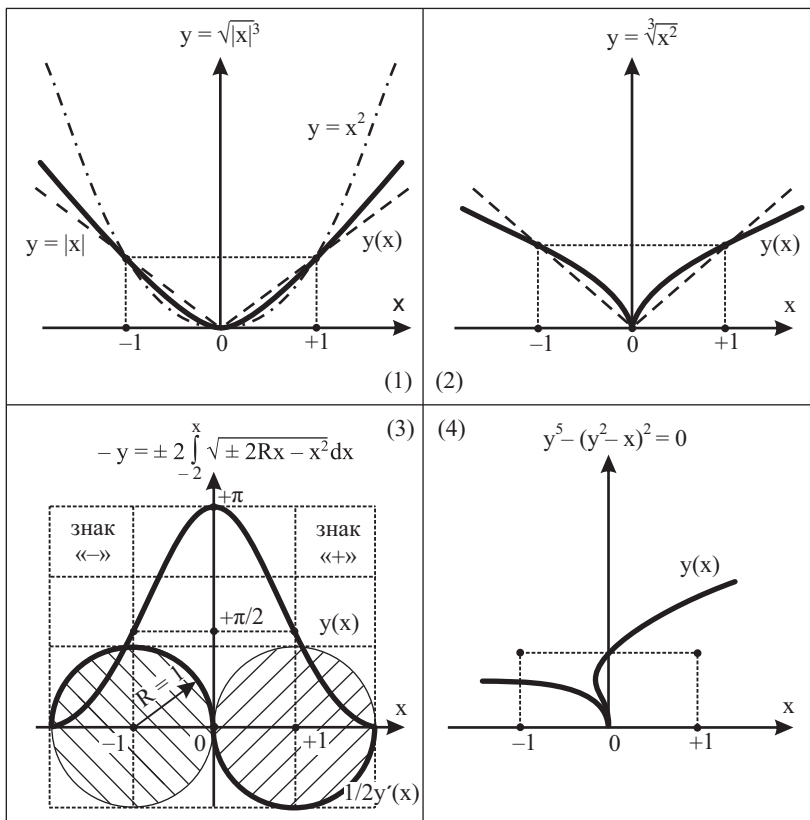


Рис. 4. Простейшие варианты маргинальных и особых экстремумов:

- (1) — гладкая *прямая* парабола Нейля с маргинальным минимумом,
- (2) — негладкая *обратная* парабола Нейля с особым минимумом,
- (3) — кривая *колокол* $R = 1$ с маргинальным максимумом в точке $s = 0$,
- (4) — кривая 5-го порядка с особым минимумом в точке $x = 0$.

Пример 2. Обратная парабола Нейля ($a = +1$, $n = 3$, $m = 2$) — рис. 4(2):

$$y = \sqrt[3]{x^2}, \quad y'(x) = 2/3\sqrt[3]{x}, \quad y''(x) = -2/9\sqrt[3]{x^4}.$$

Функция имеет особый (неострый, вертикальный, двусторонний) минимум в точке возврата 1-го рода $\{0, 0\}$ — рис. 4(2). Отдельные ветви обеих парабол Нейля по форме в точности совпадают друг с другом — рис. 4(1) и (2). В данном случае меняется только их расположение на координатной плоскости $\langle x, y \rangle$.

Пример 3. Гладкая кривая *колокол* — рис. 4(3):

$$y = \pi R^2 - R^2 \cdot \arccos \frac{R - (\pm x)}{R} \pm [R - (\pm x)] \cdot \sqrt{2R \cdot (\pm x) - x^2}.$$

$$y'(x) = - \left[\pm 2\sqrt{2R \cdot (\pm x) - x^2} \right], \quad y''(x) = -2 \cdot [R - (\pm x)] / \sqrt{2R \cdot (\pm x) - x^2},$$

где выбирают знак «+» при $x > 0$ и знак «-» при $x < 0$. Функция $y(x)$ задаётся на $\mathcal{T} \equiv (-2R, +2R)$ и получается поэтапным интегрированием площадей двух кругов с учётом знака — рис. 4(3), положительного слева и отрицательного справа от оси y . Кривая $y(x)$ при $R = 1$ имеет форму колокола. Функция имеет маргинальный максимум с уровнем $p = 3/2$ в точке $s^+ = 0$, так как $y'(0) = 0$, $y''(0) = -\infty$. Очевидно, что $y(0) = +\pi$. Уравнение аппроксимирующей маргинальной параболы здесь порядка $p = 3/2$ имеет вид: $w(x) = \pi R^2 - 4/3 \cdot \sqrt{2R \cdot |x|^3}$. Это есть касательная в точке максимума $\{0, \pi R^2\}$ обращённая вниз прямая парабола Нейля. Кроме того, в точках экстремумов полуокружностей $1/2 \cdot y'(x)$, т. е. при $x = \pm R$, кривая *колокол* имеет наклонные перегибы.

Пример 4. Алгебраическая кривая 5-го порядка, заданная неявно в виде: $y^5 = (y^2 - x)^2$. Она имеет особый (острый, вертикальный, односторонний) минимум в точке возврата 2-го рода $\{0, 0\}$ — рис. 4(4).

Разумеется, стационарные перегибы для эволюционной функции $y(x)$, в принципе, также могут иметь уровни в интервале $1 < p < 2$, т. е. быть маргинальными. Например, если в уравнении для кривой *колокол*, рис. 4(3), перед интегралом применять знак «+», т. е. попросту считать площади обоих интегрируемых кругов положительными, то приходим к уравнению этой кривой с маргинальным $\mathcal{Z}$ -образным перегибом уровня $p = 3/2$ в точке $s = 0$ и с аппроксимирующей маргинальной параболой $w(x) = \pi R^2 + 4/3 \cdot x \cdot \sqrt{2R \cdot |x|}$. В ней $y''(x)$ терпит разрыв $-\infty \div +\infty$.

Кроме данных примеров маргинального и особого экстремального поведения скалярных функций, покажем существование аналитических функций $y(x)$, хотя и не постоянных на области определения, но при этом с нулевыми значениями *всех* производных в точке экстремума! Впервые их рассмотрел Коши на приводимом ниже примере.

Пример 5. Пусть дана трансцендентная функция вида

$$\left. \begin{aligned} y(x) &= \exp(-1/x^2), \quad x \neq 0; \\ y(x) &= 0, \quad x = 0. \end{aligned} \right\}$$

В любой точке оси абсцисс она непрерывная, имеет также непрерывные производные любого наперед заданного порядка t :

$$y^{(t)}(x) = \mathcal{P}_{3t}(1/x) \cdot y(x),$$

где $\mathcal{P}_{3t}(1/x)$ — многочлен от аргумента $1/x$ степени $3t$. Общая формула при конечном параметре t доказывается по индукции. В частности,

$$y'(x) = [+2/x^3] \cdot y(x), \quad y''(x) = [-6/x^4 + 4/x^6] \cdot y(x).$$

В точке $x = 0$ функция и все её производные нулевые, в силу известного предела $\lim_{u \rightarrow \infty} u^m / e^u = 0$. Отсюда при $x \rightarrow 0$ и конечном t имеем:

$$y(0) = y^{(1)}(0) = \dots = y^{(t)}(0) = y^{(t+1)}(0) = 0.$$

Эта функция Коши отображается графически *чашеобразной* кривой, симметричной относительно оси ординат y , поскольку $y(x) = y(-x)$. Слева и справа она стремится асимптотически к прямой $x = 1$ (снизу). Здесь совершенно очевидно, что функция имеет минимум $y(0) = 0$, т. е. в начале координат. Но формально он не обосновывается общими ступенчатыми правилами. Однако применим *рациональный критерий*, основанный на изучении знаков 1-й производной $y'(x)$ слева и справа от точки стационарности. Имеем: $y'(+\alpha) > 0$ и $y'(-\alpha) < 0$. При значениях аргумента $x_{1,2} = \pm\sqrt{2/3}$ кривая $y(x)$ имеет наклонные перегибы.

Разумеется, можно сконструировать и другие примеры аналогичных скалярных функций с бесконечно большим уровнем стационарности, в том числе с различными характерами поведения в окрестности точки стационарности. При этом формально они остаются в том же самом классе эволюционных функций.

§ 1.4. Экстремумы для функций от дискретной переменной

Пусть аргумент целевой функции $y(x)$ может принимать только *дискретные* значения, относящиеся к некоторому полному дискретному множеству $\langle x \rangle \subset \mathcal{D} \subset \mathcal{R}$. Тогда отображение $y(\langle x \rangle)$ есть производное полное дискретное множество значений данной целевой функции y от аргумента x . Однако отображение $y(\mathcal{D})$, по-прежнему, есть полное множество значений этой же целевой функции от как бы *непрерывного* аргумента x . Очевидно, что в последнем варианте для определения и вычисления точек стационарности, в том числе экстремумов, функции $y(x)$ можно вполне использовать аналитические приёмы, изложенные выше. Точки экстремумов функции $y(x)$ при непрерывном x могут тогда и не принадлежать дискретному множеству корректных значений $\langle x \rangle$.

Например, пусть $s^* \notin \langle x \rangle$ — точка максимума выпуклой функции $y(x)$ на области $\mathcal{D}$ при *непрерывном* аргументе x ; x_1 и x_2 — ближайшие к s^* на числовой оси $\mathcal{R}$ снизу и сверху значения *дискретного* аргумента x , принадлежащие области $\mathcal{D}$. Тогда, очевидно, что $\max \langle y(x_1), y(x_2) \rangle$ есть максимум на $\mathcal{D} \subset \mathcal{R}$ функции $y(x)$ именно при дискретном аргументе. Заметим, что в данном варианте для функции $y(x)$ от непрерывного x можно принять (разумеется, с возможной соответствующей проверкой) и более слабое требование, а именно: чтобы она монотонно убывала слева и справа от точки s^* .

Задаваемые дискретные множества аргумента $\langle x \rangle$ могут быть весьма разнообразными. Например, это может быть полуоткрытое множество целых положительных чисел $\langle m \rangle$, где $m = 1, 2, \dots, +\infty$.

Пример. Найти целое основание для наиболее экономичной системы счисления. (Здесь под *экономичностью* системы счисления понимается её способность к записи максимального количества чисел при заданном количестве исходных цифр в пределах имеющегося основания.)

С целью упрощения допустим, что количество исходных цифр n прямо пропорционально количеству k разрядов, а последние — полные. Тогда $n = k \cdot m$, где m — целочисленное положительное основание некоей системы счисления. С учётом этого, количество записываемых m цифрами целых чисел есть степенное выражение $y(m) = m^k = m^{n/m}$. Подобная функция $y(m)$ в случае *непрерывности положительного аргумента* принимает строгий максимум при нецелом $m = e \approx 2,718$, где e есть основание для натуральных логарифмов (число Эйлера). Действительно,

$$\frac{dy}{dm} = \frac{n}{m} \cdot m^{\frac{n}{m}-1} - \frac{n}{m^2} \cdot m^{\frac{n}{m}} \cdot \ln m = n \cdot m^{\frac{n}{m}-2} \cdot (1 - \ln m) = 0 \rightarrow$$

$$\rightarrow \ln m = 1 \rightarrow m = e.$$

Причём слева и справа от $m = e$ функция $y(m)$ монотонно убывает, что отвечает её прохождению через максимум. Ближайшие слева и справа к этой точке дискретные значения аргумента: $m = 2$ и $m = 3$. Соответственно, $y(2) = 2^{n/2}$, $y(3) = 3^{n/3}$, $y(e) = e^{n/e}$. Здесь имеет место неравенство: $2^{n/2} < 3^{n/3} < e^{n/e}$. Следовательно, наиболее экономичная система счисления имеет своим основанием число $m = 3$. (Практически этот результат означает, что весьма актуальным остаётся реализация *троичной* системы счисления в средствах вычислительной техники.)

§ 1.5. Экстремумы 1-й и 2-й степени для функций от зависимой скалярной переменной типа $\tilde{x}$: $\tilde{x} = x(u)$

При вычислении дифференциалов и приращений целевой функции $y = y(x)$ порядка более 1 природа переменной x — *независимая* или *зависимая* имеет существенное значение. В данном параграфе она рассматривается как зависимая переменная и далее обозначается со знаком тильды сверху буквы, т. е. как $\tilde{x} = x(u)$, где u — независимая переменная, или аргумент. Причём по аналогии: $\tilde{x} \in [\tilde{a}, \tilde{b}] \equiv \tilde{\mathcal{T}} \subset \mathcal{R}$; $u \in [a, b] \equiv \mathcal{T} \subset \mathcal{R}$. В данном случае имеет место простейший, т. е. *двухступенчатый* вариант сложной функции: $y = y(\tilde{x}) = y[x(u)] = f(u)$. Для аналитической на отрезке $\tilde{\mathcal{T}}$ функции $\tilde{x} = x(u)$ аналогично (8) имеем *приращение зависимой переменной*:

$$\Delta \tilde{x} = d\tilde{x} + \frac{d^2 \tilde{x}}{2!} + \frac{d^3 \tilde{x}}{3!} + \frac{d^4 \tilde{x}}{4!} + \dots =$$

$$= \Delta^1 \tilde{x} + \Delta^2 \tilde{x} + \Delta^3 \tilde{x} + \Delta^4 \tilde{x} + \dots \quad (20)$$

Общее и порядковые приращения и дифференциалы переменной $\tilde{x}$ отсчитываются от какой-либо конкретной точки на числовой оси $\mathcal{R}$, например, от $\tilde{c}$. Для однозначной $y(\tilde{x})$ на $\tilde{\mathcal{T}}$ однозначность сложной функции $f(u) = y[x(u)]$ на эквивалентном отрезке $\mathcal{T}$ обеспечивается, например, для аналитического отображения $x(u)$. В таком случае она, как преобразование, осуществляет только деформацию растяжения-сжатия кривой $y(\tilde{x})$ вдоль оси $\mathcal{R}$.

Абстрактные характерные примеры разнообразных по природе однозначных (прямых) функциональных отображений типа $\tilde{x} = x(u)$ приведены на рис. 5.

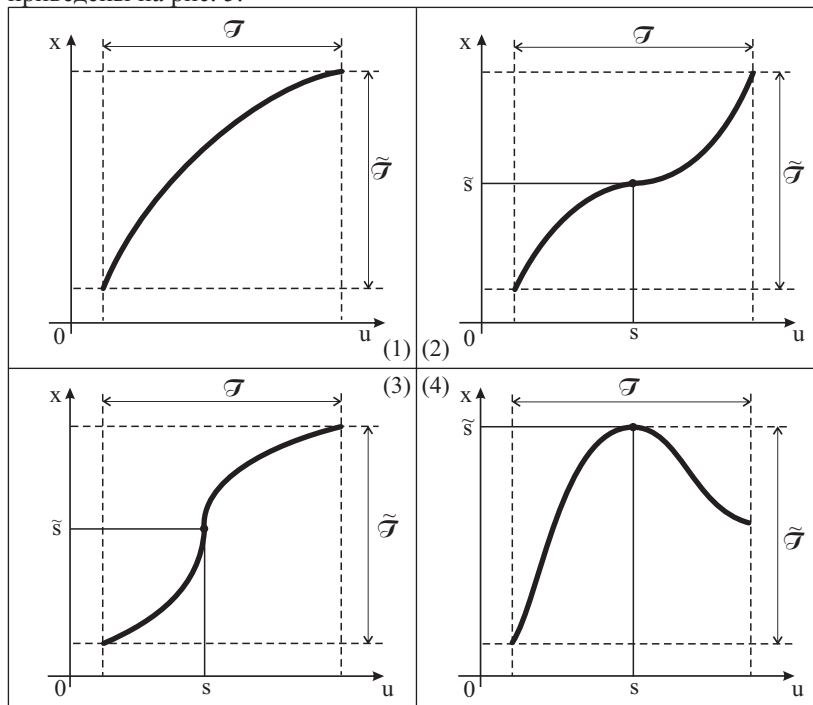


Рис. 5. Некоторые характерные варианты для скалярных гладких и негладких, регулярных и нерегулярных функциональных отображений типа $x(u) \leftrightarrow u(x)$ на эквивалентных числовых отрезках $\mathcal{T}$ и $\tilde{\mathcal{T}}$:

- (1) — $x(u)$ и $u(x)$, как взаимно гладкие и регулярные отображения;
- (2) — $x(u)$ со стационарным перегибом в точке s , как негладкое и регулярное прямое отображение (с нарушением гладкости в точке s), и $u(x)$ с крутым перегибом в точке s , как гладкое и нерегулярное обратное отображение (с нарушением непрерывной дифференцируемости в точке s);
- (3) — $x(u)$ с крутым перегибом в точке s , как гладкое и нерегулярное прямое отображение (с нарушением непрерывной дифференцируемости в точке s), и $u(x)$ со стационарным перегибом в точке s , как негладкое и регулярное обратное отображение (с нарушением гладкости в точке s);
- (4) — $x(u)$ с экстремумом в точке s , как негладкое и регулярное прямое отображение (с нарушением гладкости в точке s), и $u(x)$ с овальностью в точке s , как гладкое и нерегулярное обратное отображение (с нарушением непрерывной дифференцируемости в точке s).

Отметим, что двухступенчатую сложную аналитическую функцию $y = y(\tilde{x}) = y[x(u)] = f(u)$, не нарушая общности, можно представить собственным рядом Тейлора в двух вариантах переменной:

$$\begin{aligned}
 y &= y(\tilde{c} + \Delta\tilde{x}) = y(\tilde{c}) + \Delta y(x = \tilde{c}) = y(\tilde{c}) + \frac{dy}{d\tilde{x}}(\tilde{c}) \cdot \Delta\tilde{x} + \\
 &+ \frac{d^2y}{d\tilde{x}^2}(\tilde{c}) \cdot \frac{\Delta\tilde{x}^2}{2!} + \frac{d^3y}{d\tilde{x}^3}(\tilde{c}) \cdot \frac{\Delta\tilde{x}^3}{3!} + \frac{d^4y}{d\tilde{x}^4}(\tilde{c}) \cdot \frac{\Delta\tilde{x}^4}{4!} + \dots = \\
 &\equiv f(c + du) = f(c) + \Delta y(u = c) = f(c) + \frac{dy}{du}(c)du + \\
 &+ \frac{d^2y}{du^2}(c) \frac{du^2}{2!} + \frac{d^3y}{du^3}(c) \frac{du^3}{3!} + \frac{d^4y}{du^4}(c) \frac{du^4}{4!} + \dots
 \end{aligned} \tag{21}$$

В силу эквивалентности точек $\tilde{c}$ и c , здесь $y(\tilde{c}) = f(c)$. С учётом разложения (20), группируя слагаемые одного и того же порядка, последовательно получаем систему (22) из дифференциалов функции:

$$\left. \begin{aligned}
 dy(u = c) &= \Delta^1 y(u = c) = \frac{dy}{du}(c)du = dy(\tilde{x} = \tilde{c}) = \frac{dy}{d\tilde{x}}(\tilde{c})d\tilde{x}, \\
 d^2y(u = c) &= 2! \Delta^2 y(u = c) = \frac{d^2y}{du^2}(c)du^2 = d^2y(\tilde{x} = \tilde{c}) = \\
 &= \frac{d^2y}{d\tilde{x}^2}(\tilde{c})d\tilde{x}^2 + \frac{dy}{d\tilde{x}}(\tilde{c})d^2\tilde{x}, \\
 d^3y(u = c) &= 3! \Delta^3 y(u = c) = \frac{d^3y}{du^3}(c)du^3 = d^3y(\tilde{x} = \tilde{c}) = \frac{d^3y}{d\tilde{x}^3}(\tilde{c})d\tilde{x}^3 + \\
 &+ 3 \cdot \frac{d^2y}{d\tilde{x}^2}(\tilde{c})d\tilde{x}d^2\tilde{x} + \frac{dy}{d\tilde{x}}(\tilde{c})d^3\tilde{x}, \\
 d^4y(u = c) &= 4! \Delta^4 y(u = c) = \frac{d^4y}{du^4}(c)du^4 = d^4y(\tilde{x} = \tilde{c}) = \frac{d^4y}{d\tilde{x}^4}(\tilde{c})d\tilde{x}^4 + \\
 &+ 6 \cdot \frac{d^3y}{d\tilde{x}^3}(\tilde{c})d\tilde{x}^2d^2\tilde{x} + 4 \cdot \frac{d^2y}{d\tilde{x}^2}(\tilde{c})d\tilde{x}d^3\tilde{x} + 3 \cdot \frac{d^2y}{d\tilde{x}^2}(\tilde{c})(d^2\tilde{x})^2 + \frac{dy}{d\tilde{x}}(\tilde{c})d^4\tilde{x}, \\
 &\dots
 \end{aligned} \right\} \tag{22}$$

Дифференциалы и приращения целевой функции в такой форме записи (т. е. с их отсчётом от значения функции в указываемой точке) являются многочлен-функциями от дифференциалов зависимой переменной $\tilde{x}$. Причём нетрудно видеть, что 1-й дифференциал-приращение любой последовательно непрерывно дифференцируемой сложной функции (двухступенчатой и многоступенчатой) имеет инвариантную форму выражения относительно выбора переменной — см., например, [23]. Из системы (22) теперь последовательно находим производные целевой функции любого порядка p , согласно новой системе (23):

$$\left. \begin{aligned} \frac{dy}{du}(c) &= \frac{dy}{d\tilde{x}}(\tilde{c}) \cdot \frac{d\tilde{x}}{du}(c), \\ \frac{d^2y}{du^2}(c) &= \frac{d^2y}{d\tilde{x}^2}(\tilde{c}) \cdot \left(\frac{d\tilde{x}}{du}(c) \right)^2 + \frac{dy}{d\tilde{x}}(\tilde{c}) \cdot \frac{d^2\tilde{x}}{du^2}(c), \\ \frac{d^3y}{du^3}(c) &= \frac{d^3y}{d\tilde{x}^3}(\tilde{c}) \cdot \left(\frac{d\tilde{x}}{du}(c) \right)^3 + 3 \cdot \frac{d^2y}{d\tilde{x}^2}(\tilde{c}) \cdot \frac{d^2\tilde{x}}{du^2}(c) \cdot \frac{d\tilde{x}}{du}(c) + \frac{dy}{d\tilde{x}}(\tilde{c}) \cdot \frac{d^3\tilde{x}}{du^3}(c), \\ \frac{d^4y}{du^4}(c) &= \frac{d^4y}{d\tilde{x}^4}(\tilde{c}) \cdot \left(\frac{d\tilde{x}}{du}(c) \right)^4 + 6 \cdot \frac{d^3y}{d\tilde{x}^3}(\tilde{c}) \cdot \frac{d^2\tilde{x}}{du^2}(c) \cdot \left(\frac{d\tilde{x}}{du}(c) \right)^2 + \\ &+ 4 \cdot \frac{d^2y}{d\tilde{x}^2}(\tilde{c}) \cdot \frac{d^3\tilde{x}}{du^3}(c) \cdot \frac{d\tilde{x}}{du}(c) + 3 \cdot \frac{d^2y}{d\tilde{x}^2}(\tilde{c}) \cdot \left(\frac{d^2\tilde{x}}{du^2}(c) \right)^2 + \frac{dy}{d\tilde{x}}(\tilde{c}) \cdot \frac{d^4\tilde{x}}{du^4}(c), \\ &\dots \end{aligned} \right\} \quad (23)$$

Причём эти формулы для итоговых производных используются далее при нахождении и идентификациях стационарностей сложной целевой функции $y = y(\tilde{x}) = y[x(u)]$ любого имеющегося уровня p посредством вышеуказанных правил (1)–(7).

Пусть $y = y[x(u)] = f(u)$ имеет стационарность в точке s . Согласно правилу Ферма (см. § 1.1) и первым соотношениям из систем (22), (23), имеем определяющие стационарность целевой функции аналитические формулировки, выражаемые нижеуказанными дифференциальными уравнениями:

$$dy(u=s) = \frac{dy}{du}(s)du = \frac{dy}{d\tilde{x}}(\tilde{s}) \cdot \frac{d\tilde{x}}{du}(s)du = \frac{dy}{d\tilde{x}}(\tilde{s})d\tilde{x} = 0 \leftrightarrow \quad (24)$$

$$\leftrightarrow \frac{dy}{du}(s) = \frac{dy}{d\tilde{x}}(\tilde{s}) \cdot \frac{d\tilde{x}}{du}(s) = 0. \quad (25)$$

В самом общем случае идентификация возможного экстремума 2-го уровня для целевой двухступенчатой сложной функции выполняется через анализ либо её 2-го дифференциала, либо её 2-й производной в точке стационарности s , взятых из второго соотношения системы (22) или системы (23):

$$d^2y(u=s) = \frac{d^2y}{du^2}(s)du^2 = d^2y(\tilde{x}=\tilde{s}) = \frac{d^2y}{d\tilde{x}^2}(\tilde{s})d\tilde{x}^2 + \frac{dy}{d\tilde{x}}(\tilde{s})d^2\tilde{x} \leftrightarrow \quad (26)$$

$$\leftrightarrow \frac{d^2y}{du^2}(s) = \frac{d^2y}{d\tilde{x}^2}(\tilde{s}) \cdot \left[\frac{d\tilde{x}}{du}(s) \right]^2 + \frac{dy}{d\tilde{x}}(\tilde{s}) \cdot \frac{d^2\tilde{x}}{du^2}(s). \quad (27)$$

Причём тут возможно использование предельных соотношений по **правилам анализа неопределённостей Лопиталья**. Однако, когда все эти дифференциальные характеристики существуют, тогда могут иметь место два *непредельных* варианта стационарности целевой функции, а, следовательно, и её экстремума.

1) Либо $dy/d\tilde{x}(\tilde{s}) = 0$, при этом $d\tilde{x}/du$ существует. Тогда отображение $x(u)$ регулярное, причём, по крайней мере, в окрестности точки s . Это соответствует стационарности целевой функции в $\tilde{s}$ 1-й ступени.

2) Либо $dy/d\tilde{x}(\tilde{s}) \neq 0$ и существует (т. е. *значимая*), но $d\tilde{x}/du(s) = 0$. Тогда отображение $\tilde{x} = x(u)$ в самой точке s негладкое, но регулярное. Это соответствует стационарности целевой функции в $\tilde{s}$ 2-й ступени. Такая стационарность в данном случае реализуется в *особых* точках негладкой кривой $\tilde{x} = x(u)$, как, например, на рис. 5(2), (3), (4). В этих точках при изменении параметра отображения последнее не меняется. (Более наглядно и практично особые точки проявляются при изучении *векторных* отображений типа $\tilde{\mathbf{x}} = \mathbf{x}(\mathbf{u})$ — см. об этом в § 2.1.)

Пусть в первом варианте стационарности $d^2\tilde{x}/du^2(s)$ здесь также существует (как и 1-я производная), т. е. отображение $x(u)$ регулярное порядка не менее 2-х. Тогда, согласно (25), имеем:

$$\frac{d^2y}{du^2}(s) = \frac{d^2y}{d\tilde{x}^2}(\tilde{s}) \cdot \left[\frac{d\tilde{x}}{du}(s) \right]^2.$$

Если данная 2-я производная ненулевая, то тогда по её знаку идентифицируют характер экстремума 1-й ступени и 2-го уровня, используя правила Лейбница (2) и (3). Обе 2-ые производные в (26) имеют одинаковый знак и могут обнуляться только вместе. Характер возможного стационарного перегиба функции 3-го уровня в точке s идентифицируют при необходимости в этом с привлечением третьего соотношения из системы (23) и т. д.

Пусть во втором варианте стационарности $d^2y/d\tilde{x}^2(\tilde{s})$ здесь также существует (как и 1-я производная). Тогда, согласно (25), имеем:

$$\frac{d^2y}{du^2}(s) = \frac{dy}{d\tilde{x}}(\tilde{s}) \cdot \frac{d^2\tilde{x}}{du^2}(s).$$

Если данная 2-я производная ненулевая, то тогда по её знаку идентифицируют характер экстремума 2-й степени и 2-го уровня, используя правила Лейбница (2) и (3). В этом варианте 2-ые производные $d^2y/du^2(s)$ и $d^2\tilde{x}/du^2(s)$ могут обнуляться лишь только вместе. Характер возможного стационарного перегиба функции 3-го уровня в точке s также идентифицируют при необходимости в этом с привлечением третьего соотношения из (23) и т. д.

Заметим, что для целевой сложной функции $y = y(\tilde{x}) = y[x(u)]$ может реализовываться внешний (краевой) условный экстремум тогда, когда отображение $\tilde{x} = x(u)$ налагает на зависимую переменную $\tilde{x}$ отвечающее этому случаю ограничение — см. табл. 1.

§ 1.6. Экстремумы 1-й и 2-й степени для функций от зависимой скалярной переменной типа $\tilde{x}$: $u = u(\tilde{x})$

Предшествующая задача разрешалась аналитически естественным путём через вычисление и анализ дифференциалов или производных двухступенчатой сложной функции. Данную новую задачу, в которой зависимая переменная $\tilde{x}$ выражается только через *обратную* функцию, можно решать, в принципе, двумя способами.

Либо сначала каким-то аналитическим образом выразить нужные производные целевой неявной функции $\tilde{x} = x(u)$ через производные явной функции $u = u(\tilde{x})$ и подставить их в систему (23).

Либо гораздо более естественным путём, излагаемым ниже, сразу же выразить все необходимые для идентификации стационарности целевой функции её производные.

В силу принятого ранее (см. рис. 5) характера монотонности или хотя бы кусочной монотонности отображений $\tilde{x} = x(u)$ и $u = u(\tilde{x})$ на эквивалентных интервалах $\mathcal{T}$ и $\tilde{\mathcal{T}} \subset \mathcal{R}$ целевую функцию y можно представить однозначно или двузначно в виде замкнутого цикла:

$$y = y(\tilde{x}) = y[x(u)] = f(u) = f[u(\tilde{x})] = y(\tilde{x}) = y.$$

На основании указанного циклического отображения y поменяем местами в системах (22) и (23) обе переменные $\tilde{x}$ и u . При этом справедливость всех соотношений в системах, разумеется, сохраняется.

Например, из модифицированной таким образом исходной системы (23) выражаем последовательно все итоговые производные целевой функции в виде новой системы (28):

$$\left. \begin{aligned} \frac{dy}{du}(c) &= \frac{dy}{d\tilde{x}}(\tilde{c}) \cdot \frac{d\tilde{x}}{du}(c), \\ \frac{dy^2}{du^2}(c) &= \left[\frac{dy^2}{d\tilde{x}^2}(\tilde{c}) - \frac{dy}{du}(c) \cdot \frac{du^2}{d\tilde{x}^2}(\tilde{c}) \right] \cdot \left[\frac{d\tilde{x}}{du}(c) \right]^2, \\ \frac{d^3y}{du^3}(c) &= \left[\frac{d^3y}{d\tilde{x}^3}(\tilde{c}) - 3 \cdot \frac{d^2y}{du^2}(c) \cdot \frac{d^2u}{d\tilde{x}^2}(\tilde{c}) \cdot \frac{du}{d\tilde{x}}(\tilde{c}) - \frac{dy}{du}(c) \cdot \frac{d^3u}{d\tilde{x}^3}(\tilde{c}) \right] \cdot \left[\frac{d\tilde{x}}{du}(c) \right]^3, \\ \frac{d^4y}{du^4}(c) &= \left\{ \frac{d^4y}{d\tilde{x}^4}(\tilde{c}) - 6 \cdot \frac{d^3y}{du^3}(c) \cdot \frac{d^2u}{d\tilde{x}^2}(\tilde{c}) \cdot \left[\frac{du}{d\tilde{x}}(\tilde{c}) \right]^2 - 4 \frac{d^2y}{du^2}(c) \cdot \frac{d^3u}{d\tilde{x}^3}(\tilde{c}) \cdot \frac{du}{d\tilde{x}}(\tilde{c}) - \right. \\ &\quad \left. - 3 \cdot \frac{d^2y}{du^2}(c) \cdot \left[\frac{d^2u}{d\tilde{x}^2}(\tilde{c}) \right]^2 - \frac{dy}{du}(c) \cdot \frac{d^4u}{d\tilde{x}^4}(\tilde{c}) \right\} \cdot \left[\frac{d\tilde{x}}{du}(c) \right]^4, \\ &\dots \dots \dots \end{aligned} \right\} \quad (28)$$

По этим *рекуррентным* соотношениям для итоговых производных скалярной функции можно так же, как и в предыдущем параграфе, находить и идентифицировать её экстремум 1-й и 2-й степени. Так, из второго соотношения системы (28) получаем тождественные формулам (26, 27) выражения для дифференциальных характеристик 2-го порядка в точке стационарности целевой функции:

$$d^2y(u=s) = \frac{d^2y}{du^2}(s) du^2 = \frac{d^2y}{d\tilde{x}^2}(\tilde{s}) \cdot \left[\frac{d\tilde{x}}{du}(s) \right]^2 du^2 - \frac{dy}{d\tilde{x}}(\tilde{s}) \cdot \frac{d\tilde{x}}{du}(\tilde{s}) d^2u \leftrightarrow \quad (29)$$

$$\leftrightarrow \frac{d^2y}{du^2}(s) = \left[\frac{d^2y}{d\tilde{x}^2}(\tilde{s}) - \frac{dy}{d\tilde{x}}(\tilde{s}) \cdot \frac{d\tilde{x}}{du}(\tilde{s}) \cdot \frac{d^2u}{d\tilde{x}^2}(\tilde{s}) \right] \cdot \left[\frac{d\tilde{x}}{du}(\tilde{s}) \right]^2. \quad (30)$$

Тождественность формул (27) и (30) обеспечивает, например, известная взаимосвязь вторых производных *прямой* и *обратной* функции:

$$\begin{aligned} \frac{d^2\tilde{x}}{du^2} &= -\frac{d^2u}{d\tilde{x}^2} \cdot \left(\frac{d\tilde{x}}{du} \right)^3 \leftrightarrow \frac{d^2u}{d\tilde{x}^2} = -\frac{d^2\tilde{x}}{du^2} \cdot \left(\frac{du}{d\tilde{x}} \right)^3 \\ \left(d^2\tilde{x} = -\frac{d\tilde{x}}{du} d^2u \neq 0 \leftrightarrow d^2u = -\frac{du}{d\tilde{x}} d^2\tilde{x} \neq 0. \right) \end{aligned} \quad (31)$$

Дальнейший анализ вполне аналогичен приведённому в § 1.5.

§ 1.6.1. Изопараметрические многочлены

Метод группировки, применённый выше в § 1.5 для формирования системы (22), позволяет в итоге сразу же находить все приращения, дифференциалы и производные любого нужного порядка для сложной скалярной функции $y = y(\tilde{x}) = y[x(u)] = f(u)$. (Аналогичный подход может применяться и в случае *многоступенчатой* сложной скалярной функции.) В *двухступенчатом* варианте общие формулы выводятся вполне логичным и естественным вышеописанным путём — сначала для приращений, а затем для дифференциалов и производных:

$$\Delta^p f(u = c) = \sum_{q=1}^p \left\{ \frac{d^q y}{d\tilde{x}^q}(\tilde{c}) \cdot \left[\sum_{Q(p,q)} \frac{\Delta^i \tilde{x}^j}{j!} \right] \right\},$$

$$d^p f(u = c) = \frac{d^p f}{du^p}(c) dy^p = p! \Delta^p f(u = c) =$$

$$= \sum_{q=1}^p \left\{ \frac{d^q y}{d\tilde{x}^q}(\tilde{c}) \cdot \left[\sum_{Q(p,q)} \frac{p!}{j_1! \cdot j_2! \cdot \dots \cdot j_m!} \left(\frac{d^{i_1} \tilde{x}}{i_1!} \right)^{j_1} \left(\frac{d^{i_2} \tilde{x}}{i_2!} \right)^{j_2} \cdot \dots \cdot \left(\frac{d^{i_m} \tilde{x}}{i_m!} \right)^{j_m} \right] \right\}. \quad (32)$$

Причём в двухступенчатом варианте применяются индексы двух типов i и j , подчиняющиеся соотношениям:

$$\left. \begin{aligned} j_1 + j_2 + \dots + j_m &= q \leq p, \\ i_1 \cdot j_1 + i_2 \cdot j_2 + \dots + i_m \cdot j_m &= p. \end{aligned} \right\} (p = \text{const}) \quad (33)$$

Поэтому степенные выражения формы (32) от дифференциалов $\tilde{x}$ определяются тут как *изопараметрические многочлены валентности 2*. В этом определении *валентность* многочлена есть количество звеньев в сложной функции, а наиболее общо это есть количество независимых индексов в элементах. Причём в системе (33) p есть главный индекс, а q есть ограниченный им второстепенный индекс.

При валентности 2 общее количество разнообразных комбинаций для произведений степеней в любом изопараметрическом многочлене (в частности, здесь это есть количество всевозможных комбинаций для произведений степеней дифференциалов $d^i \tilde{x}$) при исходно задаваемом значении p есть некоторая арифметическая функция $\mathcal{A}(p)$. Количество тех же самых комбинаций при задаваемых значениях p и q есть некая арифметическая функция $\mathcal{Q}(p, q) < \mathcal{A}(p)$. Обе эти дискретные функции при валентности 2 получаются из требований к индексам типа (33).

В числовом виде обе указанные арифметические функции приведены в таблице 2 до значения главного индекса (аргумента) $p = 10$.

Таблица 2. Арифметические функции $\mathcal{P}(p)$ и $\mathcal{Q}(p, q)$ для двухвалентных изопараметрических многочленов

p \ q	$\mathcal{Q}(p, q)$										$\mathcal{P}(p)$
	1	2	3	4	5	6	7	8	9	10	
1	1	—	—	—	—	—	—	—	—	—	1
2	1	1	—	—	—	—	—	—	—	—	2
3	1	1	1	—	—	—	—	—	—	—	3
4	1	2	1	1	—	—	—	—	—	—	5
5	1	2	2	1	1	—	—	—	—	—	7
6	1	3	3	2	1	1	—	—	—	—	11
7	1	3	4	3	2	1	1	—	—	—	15
8	1	4	5	5	3	2	1	1			22
9	1	4	7	6	5	3	2	1	1	—	30
10	1	5	8	9	6	4	2	1	1	1	38

Если левую и правую части формулы (32) поделить на du^p , то в итоге получается общая формула для производных двухступенчатой сложной функции $y = y(\tilde{x}) = y[x(u)] = f(u)$. Она также, согласно системе (23), имеет структуру двухвалентного изопараметрического многочлена от производных $d^q \tilde{x}/du^q$ (при постоянном параметре p).

Обратим тут внимание на то, как в системе (22) и в формуле (32) для каждого дифференциала $d^p y(u=c)$ последовательно образуются суммы произведений. Во-первых, порядок q производной от y по $\tilde{x}$ пробегает значения от 1 до p . Во-вторых, сумма степеней j дифференциалов $\tilde{x}$ (суммарная размерность) при производной $d^q y/dx^q$ всегда равна $q \leq p$. В-третьих, сумма произведений порядков i дифференциалов $\tilde{x}$ на их степени j равна главному параметру p . Это отвечает соотношениям индексов в (33). Таким образом, формируется своеобразный многочлен (в рассматриваемом случае от дифференциалов $\tilde{x}$ указанных порядков i и степеней j), содержащий в своей самой полной форме для заданного основного брутто-параметра p всегда строго обусловленные количества слагаемых-произведений $\mathcal{P}(p)$ и $\mathcal{Q}(p, q)$ — см. табл. 2.

В данной монографии ещё не раз повстречаются *двухвалентные изопараметрические многочлены* с разнообразными элементами или аргументами в них, а также с различными числовыми коэффициентами при частных слагаемых. Например, в многочленах (32) от аргументов

$d^i \tilde{x}$ в степенях j у отдельных слагаемых имеются полиномиальные коэффициенты типа:

$$N_{j_1, j_2, \dots, j_m}^{p, q} = \frac{p!}{\left[j_1! (i_1!)^{j_1} \right] \cdot \left[j_2! (i_2!)^{j_2} \right] \cdot \dots \cdot \left[j_m! (i_m!)^{j_m} \right]}.$$

Они отличаются от обычных полиномиальных коэффициентов тем, что содержат здесь не один («р»), а два («р» и «q») брутто-параметра и поэтому два ряда индексов i и j , подчиняющихся соотношениям (33).

Для k -ступенчатой сложной функции в этих изопараметрических многочленах валентности $k \geq 2$ от аргументов $d^i \tilde{x}$ применяются полиномиальные коэффициенты с k брутто-параметрами и с k рядами индексов, подчиняющихся собственным иерархическим соотношениям, обобщающим (33). Для каждого из многочленов валентности k главный параметр p постоянен, т. е. $p = \text{const}$. Именно поэтому все они по своей природе суть изопараметрические многочлены.

В рассмотренном случае числовые коэффициенты при конкретных слагаемых — положительные величины, имеющие полиномиальное происхождение. Они тождественны таковым, например, для полинома

$$\text{вида: } \left(1 + x_1 + \frac{x_2}{2!} + \dots + \frac{x_p}{p!} \right)^p.$$

Абстрагируясь от конкретных числовых значений коэффициентов, приведём первые четыре примера двухвалентных изопараметрических многочленов вообще без коэффициентов:

1. $p = 1, m_1 = x_1^1$ ($q = 1$);
2. $p = 2, m_2 = x_2^1 + x_1^2$ ($q = 1$ и 2);
3. $p = 3, m_3 = x_3^1 + x_2^1 \cdot x_1^1 + x_1^3$ ($q = 1, 2$ и 3);
4. $p = 4, m_4 = x_4^1 + (x_3^1 \cdot x_1^1 + x_2^2) + x_2^1 \cdot x_1^2 + x_1^4$ ($q = 1, 2, 3$ и 4).

§ 1.7. Экстремумы для функции от независимой скалярной переменной, заданной через обратную функцию

Пусть целевая функция y от независимой переменной x задана через обратную функцию $x = x(y)$ и в явном виде y тут не выражается, т. е. как $y = y(x)$. Для того чтобы здесь найти и идентифицировать

экстремумы или стационарные перегибы прямой целевой функции $y(x)$, необходимо каким-либо образом выразить её дифференциалы и производные через те же характеристики для обратной функции $x(y)$. Разумеется, в этих преобразованиях должна учитываться возможность многозначности отображений $y(x)$ и $x(y)$. Отметим, что ещё Ньютон вычислял методом флюксий самые первые коэффициенты обратного степенного ряда через коэффициенты прямого степенного ряда [63].

Для производных 1-го порядка имеем весьма очевидную простую формулу:

$$\frac{dy}{dx}(x) = 1 / \frac{dx}{dy}(y), \text{ или } dy(x) = \left[1 / \frac{dx}{dy}(y) \right] dx.$$

В обоих вариантах этой формулы аргументы x и y должны отвечать заданной функции $x = x(y)$ и быть взаимно однозначными парами!

Для вычисления дифференциалов и производных целевой функции любого необходимого порядка p , например, скалярной функции $y(x)$, непрерывно дифференцируемой не менее чем p раз, обратимся вновь к системе (22). Представим целевую функцию как отображение на самое себя — через *формально* зависимую в данном случае переменную $x(y)$, а именно в виде $y = y[x(y)]$. Как первичная переменная, т. е., по сути, аргумент, она в данном случае имеет 1-й *ненулевой* дифференциал и следующие порядка более 1 *нулевые* дифференциалы $d^2y, d^3y, d^4y, \dots$. Подставляя в систему (22) зависимую переменную с заменой $\tilde{x} \rightarrow x$, получаем новую систему уравнений для дифференциалов и производных обратной и прямой функций:

$$\left. \begin{aligned} dy &= \frac{dy}{dx} dx, \\ 0 &= \frac{d^2y}{dx^2} dx^2 + \frac{dy}{dx} d^2x, \\ 0 &= \frac{d^3y}{dx^3} dx^3 + 3 \cdot \frac{d^2y}{dx^2} dx d^2x + \frac{dy}{dx} d^3x, \\ 0 &= \frac{d^4y}{dx^4} dx^4 + 6 \cdot \frac{d^3y}{dx^3} dx^2 d^2x + 4 \cdot \frac{d^2y}{dx^2} dx d^3x + \\ &\quad + 3 \cdot \frac{d^2y}{dx^2} d^2x^2 + \frac{dy}{dx} d^4x, \\ \dots \end{aligned} \right\} \quad (34)$$

Отсюда логичным и вполне естественным образом выводятся общие рекуррентные формулы порядка p для производных целевой функции, т. е. $d^p y/dx^p$, и для её же дифференциалов, т. е. $d^p y(x) = (d^p y/dx^p)dx^p$. Рекуррентные формулы весьма удобны для нахождения связи между производными прямой и обратной функций порядка от 1 до p :

$$\left. \begin{aligned} \frac{dy}{dx} &= \left(\frac{dx}{dy} \right)^{-1}, \\ \frac{d^2 y}{dx^2} &= - \left(\frac{dy}{dx} \right)^3 \cdot \frac{d^2 x}{dy^2}, \\ \frac{d^3 y}{dx^3} &= - \left(\frac{dy}{dx} \right)^4 \cdot \frac{d^3 x}{dy^3} - 3 \cdot \left(\frac{dy}{dx} \right)^2 \cdot \frac{d^2 y}{dx^2} \cdot \frac{d^2 x}{dy^2}, \\ \frac{d^4 y}{dx^4} &= - \left(\frac{dy}{dx} \right)^5 \cdot \frac{d^4 x}{dy^4} - 4 \cdot \left(\frac{dy}{dx} \right)^3 \cdot \frac{d^2 y}{dx^2} \cdot \frac{d^3 x}{dy^3} - \\ &\quad - 3 \cdot \left(\frac{dy}{dx} \right)^4 \cdot \frac{d^2 y}{dx^2} \cdot \left(\frac{d^2 x}{dy^2} \right)^2 - 6 \cdot \left(\frac{dy}{dx} \right)^2 \cdot \frac{d^3 y}{dx^3} \cdot \frac{d^2 x}{dy^2}, \\ &\dots \dots \dots \end{aligned} \right\} \quad (35)$$

Вышеуказанные формулы содержат те же самые полиномиальные коэффициенты, что и фигурирующие в (22), (23), (28) и (32). Далее, придерживаясь стандартной процедуры, изложенной в § 1.1, находим и идентифицируем неособые экстремумы или стационарные перегибы целочисленных уровней $p \geq 2$. Отметим при этом, что стационарный перегиб кривой $y(x)$ есть крутой перегиб кривой $x(y)$ в одной и той же точке в сопутствующих базисах $\langle y, x \rangle$ и $\langle x, y \rangle$.

Полученная система рекуррентных формул (35) для поэтапного вычисления производных аналитических (или даже наиболее общо — p раз непрерывно дифференцируемых) обратных функций от скалярного аргумента в случае, когда они явно не выражаются, имеет, разумеется, гораздо более универсальное значение и применимость, нежели только в аспекте изучаемых в монографии решений задач на экстремум.

В свою очередь, из системы (35) далее довольно просто получаются *итоговые* общие формулы для связи производных прямой и обратной функций любого наперед заданного порядка p . Ниже они приводятся в mnemonic наиболее удобной форме:

$$\left. \begin{aligned}
 \frac{dy}{dx} &= \left(\frac{dx}{dy} \right)^{-1}, \\
 \frac{d^2y}{dx^2} &= - \left(\frac{dx}{dy} \right)^{-1} \cdot \left[\left(\frac{dx}{dy} \right)^{-2} \cdot \frac{d^2x}{dy^2} \right], \\
 \frac{d^3y}{dx^3} &= - \left(\frac{dx}{dy} \right)^{-1} \cdot \left[\left(\frac{dx}{dy} \right)^{-3} \cdot \frac{d^3x}{dy^3} - 3 \cdot \left(\frac{dx}{dy} \right)^{-4} \cdot \left(\frac{d^2x}{dy^2} \right)^2 \right], \\
 \frac{d^4y}{dx^4} &= - \left(\frac{dx}{dy} \right)^{-1} \cdot \left[\left(\frac{dx}{dy} \right)^{-4} \cdot \frac{d^4x}{dy^4} - 10 \cdot \left(\frac{dx}{dy} \right)^{-5} \cdot \frac{d^2x}{dy^2} \cdot \frac{d^3x}{dy^3} + \right. \\
 &\quad \left. + 15 \cdot \left(\frac{dx}{dy} \right)^{-6} \cdot \left(\frac{d^2x}{dy^2} \right)^3 \right], \\
 &\dots \dots \dots
 \end{aligned} \right\} \quad (36)$$

Обратим внимание на то, что хорошо известная вторая формула системы вытекает также из тождественных друг другу двух формул для кривизны касательной окружности к гладкой плоской кривой — абсолютной её характеристики:

$$K = \frac{\frac{d^2y}{dx^2}}{\left[1 + \left(\frac{dy}{dx} \right)^2 \right]^{3/2}} = - \frac{\frac{d^2x}{dy^2}}{\left[1 + \left(\frac{dx}{dy} \right)^2 \right]^{3/2}}.$$

Причём сумма обеих частей этой формулы, с учётом переноса, есть некий *нулевой* дифференциальный инвариант для *плоской* гладкой кривой, что далее получит свою интерпретацию и обобщение в § 1.7.1.

Таким образом, в данном параграфе был изложен аналитический подход к решению экстремальных задач в тех случаях, когда целевая функция задана формальным образом только через обратную функцию, т. е. в виде явной формулы она не выражается. Вместе с тем, выше в самом общем виде была установлена взаимосвязь между производными непрерывно дифференцируемых хотя бы p раз прямой и обратной функций порядка p . Отсюда, в частности, можно ещё дальше развить понятие об изопараметрических многочленах (§ 1.6.1).

§ 1.7.1. Зеркальные изопараметрические многочлены

Все формулы в системах (35) и (36) носят *зеркальный* характер по отношению к взаимному обмену переменной и функции $x \leftrightarrow y$. Кроме того, систему (36) можно преобразовать и дальше, а именно к форме системы *изопараметрических* многочленов валентности 2. При этом отмеченное свойство зеркальности сохраняется. Получаемая в итоге система приобретает зеркальный вид (без 1-го уравнения, так как оно используется при формировании этой новой системы):

$$\left. \begin{aligned} \frac{d^2 y}{dy} + \frac{d^2 x}{dx} &= 0, \\ \frac{d^3 y}{dy} + \frac{d^3 x}{dx} &= 3 \cdot \left(\frac{d^2 x}{dx} \right)^2 = 3 \cdot \left(\frac{d^2 y}{dy} \right)^2, \\ \frac{d^4 y}{dy} + \frac{d^4 x}{dx} &= 10 \cdot \frac{d^2 x}{dx} \cdot \frac{d^3 x}{dx} - 15 \cdot \left(\frac{d^2 x}{dx} \right)^3 = \\ &= 10 \cdot \frac{d^2 y}{dy} \cdot \frac{d^3 y}{dy} - 15 \cdot \left(\frac{d^2 y}{dy} \right)^3, \\ &\dots \end{aligned} \right\} \quad (37)$$

Поясним смысл этих уравнений. Если аргумент функции $y(x)$ имеет приращение dx , то функция имеет дифференциалы dy , d^2y , d^3y , $d^4y \dots$. Если аргумент функции $x(y)$ имеет приращение $dy = (dy/dx)dx$ (т. е. между dx и dy задана естественная здесь взаимосвязь), то обратная функция имеет также *ненулевые* дифференциалы dx , d^2x , d^3x , d^4x , $\dots$. Все они и фигурируют в системе уравнений (37).

Для того чтобы далее перейти к форме уравнений, отвечающей изопараметрическим многочленам валентности 2, примем следующие обозначения:

$$\frac{dy^{p+1}}{dy} = w_p, \quad \frac{dx^{p+1}}{dx} = v_p.$$

Применив данные обозначения к системе (37) в двух её вариантах, в итоге приходим к паре формально тождественных систем зеркальных изопараметрических многочленов:

$$\begin{cases} w_1^1 = -v_1^1, \\ w_2^1 = -v_2^1 + 3 \cdot v_1^2, \\ w_3^1 = -v_3^1 + 10 \cdot v_2^1 \cdot v_1^1 - 15 \cdot v_1^3, \\ w_4^1 = -v_4^1 + 15 \cdot v_3^1 \cdot v_1^1 + 10 \cdot v_2^2 - 105 \cdot v_2^1 \cdot v_1^2 + 105 \cdot v_1^4, \\ \dots \end{cases}$$

↓

$$\begin{cases} v_1^1 = -w_1^1, \\ v_2^1 = -w_2^1 + 3 \cdot w_1^2, \\ v_3^1 = -w_3^1 + 10 \cdot w_2^1 \cdot w_1^1 - 15 \cdot w_1^3, \\ v_4^1 = -w_4^1 + 15 \cdot w_3^1 \cdot w_1^1 + 10 \cdot w_2^2 - 105 \cdot w_2^1 \cdot w_1^2 + 105 \cdot w_1^4, \\ \dots \end{cases}$$

Числовые коэффициенты при слагаемых-произведениях здесь те же, что и в исходной системе (36) для производных обратной функции. Сумма числовых коэффициентов для каждого уравнения с параметром r равна $(-r)!$. Причём знаки этих коэффициентов даёт множитель $(-1)^r$. Количество слагаемых-произведений в каждом из уравнений задаёт числовая функция $\mathcal{A}(r)$, согласно, например, её значениям в табл. 2.

§ 1.8. Экстремумы для неявных функций от независимой скалярной переменной

Пусть целевая функция $y = y(x)$ задана вообще *неявным* образом уравнением $f(y, x) = 0$. Причём, по-прежнему, она обладает свойством *непрерывной дифференцируемости*, или эволюционности порядка хотя бы не менее 2-х. Предположим, что исходно заданная функция $f(y, x)$ также в достаточно малой окрестности кривой $y(x)$ на плоскости $\langle x, y \rangle$, по крайней мере, дважды непрерывно частно дифференцируема по x и по y . Полный (нулевой) дифференциал от f вдоль кривой $y(x)$ вычисляется по формуле:

$$df[x, y(x)] = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy = 0.$$

Поскольку $\partial f/\partial x$ и $\partial f/\partial y$ в окрестности $y(x)$ существуют и непрерывны, то из этого уравнения можно выразить 1-ю производную целевой функции как:

$$\frac{dy}{dx} = -\frac{\partial f}{\partial x} / \frac{\partial f}{\partial y} \quad (38)$$

($f(x, y) = 0$).

В свою очередь, 2-я и 3-я производные целевой функции при соответствующих ещё более сильных предположениях вычисляются дальнейшим последовательным дифференцированием (38) как:

$$\frac{d^2 y}{dx^2} = -\left[\frac{\partial^2 f}{\partial y^2} \cdot \left(\frac{dy}{dx} \right)^2 + 2 \cdot \frac{\partial^2 f}{\partial y \partial x} \cdot \frac{dy}{dx} + \frac{\partial^2 f}{\partial x^2} \right] / \frac{\partial f}{\partial y}, \quad (39)$$

$$\begin{aligned} \frac{d^3 y}{dx^3} = & -\left[3 \cdot \frac{\partial^2 f}{\partial y^2} \cdot \frac{dy}{dx} \cdot \frac{d^2 y}{dx^2} + 3 \cdot \frac{\partial^2 f}{\partial y \partial x} \cdot \frac{d^2 y}{dx^2} + 3 \cdot \frac{\partial^3 f}{\partial^2 y \partial x} \left(\frac{dy}{dx} \right)^2 + \right. \\ & \left. + 3 \cdot \frac{\partial^3 f}{\partial y \partial x^2} \cdot \frac{dy}{dx} + \frac{\partial^3 f}{\partial y^3} \cdot \left(\frac{dy}{dx} \right)^3 + \frac{\partial^3 f}{\partial x^3} \right] / \frac{\partial f}{\partial y}. \end{aligned} \quad (40)$$

Эти формулы для удобства представлены здесь в рекуррентной форме. Обратим внимание на то, что и в данном случае (по понятным соображениям) формулы (39) и (40) носят зеркальный характер по отношению к замене переменных $x \leftrightarrow y$.

В простейшем случае функция $y(x)$ стационарна в точке $\{x_0, y_0\}$, удовлетворяющей уравнению $f(y_0, x_0) = 0$, если в ней выполняются два частных условия

$$\frac{\partial f}{\partial x}(y_0, x_0) = 0 \quad \text{и} \quad \frac{\partial f}{\partial y}(y_0, x_0) \neq 0,$$

что вытекает из формулы (39). Чтобы тут в итоге выяснить характер стационарности $y(x)$ в этой точке, сначала обратимся к формуле (39):

$$\frac{d^2 y}{dx^2}(y_0, x_0) = -\frac{\partial^2 f}{\partial x^2}(y_0, x_0) / \frac{\partial f}{\partial y}(y_0, x_0).$$

При отрицательной 2-й производной $y''(x)$ имеем максимум, при положительной 2-й производной $y''(x)$ имеем минимум. Если же она нулевая, то тогда обратимся к формуле (40):

$$\frac{d^3 y}{dx^3}(y_0, x_0) = -\frac{\partial^3 f}{\partial x^3}(y_0, x_0) / \frac{\partial f}{\partial y}(y_0, x_0).$$

При нулевой 3-й производной имеем стационарный перегиб и т. д.

§ 1.9. Экстремумы уровня $p = 2$ для функций от независимой векторной переменной в аффинном пространстве $\mathcal{A}^n$

Пусть целевая скалярная функция задана как $y = y(x_1, x_2, \dots, x_n)$. Причём, пока иное не оговорено особо, $x_1, x_2, \dots, x_n$ суть n частных *независимых* скалярных вещественных переменных, или аргументов. Они могут принимать некие подмножества допустимых значений на собственных вещественных числовых осях — множествах $(-\infty, +\infty) \equiv \mathcal{R}_i$. Эти частные множества путём их прямого (декартова) произведения объединяются в некоторое общее множество. Если такое множество непрерывно-связное, то оно называется *областью определения* данной функции $\mathcal{D} \subseteq \mathcal{A}^n$. В свою очередь, отображение $y(\mathcal{D})$ есть полное множество значений целевой функции. Зачастую поведение функции рассматривают на некотором компактном подмножестве, образуемом подобным образом частными компактными подмножествами чисел x_i , например, отрезками типа $[a_i, b_i] \equiv \mathcal{T}_i \subset \mathcal{R}_i$. С целью более удобного представления математических операций с функциями $y(x_1, x_2, \dots, x_n)$ далее они рассматриваются как функции только от одной, но *векторной* переменной, т. е. как $y(\mathbf{x})$. В такой краткой форме представления $\mathbf{x}$ есть n -мерная независимая вещественная переменная, или вектор-аргумент. Он может принимать, в принципе, любые значения на области $\mathcal{D}$ определения целевой функции в n -мерном вещественном аффинном координатном пространстве $\mathcal{A}^n$. Однако для наглядности или для ещё большей конкретности целевую функцию $y(\mathbf{x})$ можно рассматривать на некотором заданном компактном подмножестве, например, на n -мерной закрытой области $\mathcal{T}^n \subset \mathcal{A}^n$. Причём заданное множество $\mathcal{T}^n$ есть некоторая прямоугольная n -мерная арифметическая область в $\mathcal{A}^n$, тождественно равная прямому (декартову) произведению множеств заданных частных отрезков $\mathcal{T}_i \subset \mathcal{R}_i$. Что важно отметить, на компакте, согласно **теореме Вейерштрасса** [17, 23], непрерывная функция $y(\mathbf{x})$ обязательно достигает своих локальных верхней и нижней граней. Традиционно (по крайней мере, с точки зрения *линейной алгебры*) считают, что в исходной форме векторный аргумент есть вектор-столбец, или $n \times 1$ -вектор с частными скалярными элементами $x_1, x_2, \dots, x_n$.

Обратим здесь особое внимание также на то, что исходное n -мерное координатное пространство для отображения любых или допустимых значений переменной $\mathbf{x}$ по своей природе, вообще, является *аффинным* пространством — с заданной в нём *аффинной* системой координат.

Ему исходно не присущи какие-либо меры или нормы для оценок длин, расстояний и углов. (Переход к метрическому или нормированному пространству в данной монографии всегда специально оговаривается и применяется лишь только тогда, когда это вызывается какой-либо необходимостью, например, при оценках скорости сходимости в ряде численных процедур оптимизации.)

Пусть целевая функция ограничена либо сверху, либо снизу, т. е. или $y(\mathbf{x}) \leq M$ или $y(\mathbf{x}) \geq M$, где M — конечное число. Тогда в первом случае $y(\mathbf{x})$ на некоторой области $\mathcal{T}^n$ принимает максимум и во втором случае $y(\mathbf{x})$ на некоторой области $\mathcal{T}^n$ принимает минимум в некоторой точке $\mathbf{s}^* \in \mathcal{T}^n$. Соответственно значение $y(\mathbf{s}^*)$ есть некий экстремум данной целевой функции. Причём дальнейший интерес представляет только такой нетривиальный вариант экстремума, для которого точка $\mathbf{s}^*$ не является какой-то граничной точкой области $\mathcal{D}$ или $\mathcal{T}^n$, а находится именно внутри её (т. е. *внутренний* экстремум).

Далее для функции $y(\mathbf{x})$ рассматривается *аналитический метод* решения задачи на экстремум целочисленного уровня $p = 2$ аналогично тому, как это было изложено в § 1.1 для одномерного варианта, но уже в многомерном варианте. Такую задачу для общего класса хотя бы дважды непрерывно дифференцируемых скалярных функций впервые достаточно результативно рассмотрел Леонард Эйлер (1730 г.) [57].

Примем, как и ранее, что целевая функция $y(\mathbf{x})$ по своей природе *эволюционная*, т. е. на области своего определения она ограничена по величине — либо сверху, либо снизу и, вместе с тем, непрерывная и непрерывно дифференцируемая. Для нахождения и идентификации имеющихся локальных экстремумов эволюционных функций $y(\mathbf{x})$ на вещественной области определения в $\mathcal{A}^n$ используют технику полных дифференциалов, развитую впервые также Леонардом Эйлером [57]. Предположим, что целевая функция $y(\mathbf{x})$ рассматривается на некоторой закрытой области $\mathcal{T}^n \subset \mathcal{A}^n$ и при этом она на ней, по крайней мере, дважды непрерывно дифференцируемая. При таких ещё более сильных допущениях необходимые и достаточные условия существования в некоторой внутренней точке данной области $\mathbf{s}^* \in \mathcal{T}^n$ неособого и при этом строгого экстремума скалярной функции $y = y(\mathbf{x})$ 2-го уровня формулируются в виде классических **правил Эйлера**. Они выражаются здесь через полные дифференциалы функции — отдельно для случаев её максимума и минимума того же 2-го уровня:

$$dy(\mathbf{x} = \mathbf{s}^+) = 0, d^2y(\mathbf{x} = \mathbf{s}^+) < 0; \quad (41)$$

$$dy(\mathbf{x} = \mathbf{s}^-) = 0, d^2y(\mathbf{x} = \mathbf{s}^-) > 0. \quad (42)$$

Очевидно, для *нестрогого* экстремума целевой функции в правых неравенствах применяют соответственно знаки $\leq$ и $\geq$. Важно отметить, что в рассматриваемом многомерном варианте в правых частях формул (41), (42) возможны также знакопеременные значения второго полного дифференциала функции, что отвечает её *стационарной седловине 2-го уровня* (смешанного типа). В случае же обнуления в точке $\mathbf{s}$ и 2-го полного дифференциала функции теоретически требуется исследование в ней d^3y и т. д. — с повышением *целочисленного* уровня либо экстремума, либо стационарной седловины (чистого или смешанного типа).

В широком смысле первое требование в (41) или (42) позволяет выявить точку $\mathbf{s}$ или точки $\mathbf{s}_k$ или подмножество $\langle \mathbf{s} \rangle$ *стационарности* функции $y(\mathbf{x})$ на области её определения. В анализе это обосновывает классическая **теорема Эйлера — Ферма** о необходимом и достаточном условии существования в точке $\mathbf{s}$ стационарности для непрерывно дифференцируемой функции $y(\mathbf{x})$. Оно выражается аналитическим образом любым из двух указанных ниже тождественных уравнений — либо через 1-й дифференциал, либо через 1-ю тензор-производную (*градиент*) функции в искомой точке стационарности $\mathbf{s} \in \mathcal{A}^n$:

$$dy(\mathbf{x} = \mathbf{s}) = 0 \leftrightarrow \frac{dy}{dx}(\mathbf{s}) = \mathbf{0}. \quad (43)$$

Кроме этого, указанные уравнения задают необходимое условие существования экстремума именно для непрерывно дифференцируемой скалярной функции, причём экстремум для неё может быть только неособым (**лемма Эйлера — Ферма**).

Полные 1-й и 2-й дифференциалы функции в (41)–(43) вычисляются обычным образом — как полные суммы частных дифференциалов, составляющие *линейную и квадратичную дифференциальные формы*:

$$dy = \sum_{i=1}^n (\partial y)_i = \sum_{i=1}^n \frac{\partial y}{\partial x_i} dx_i, \quad (44)$$

$$d^2y = \sum_{j=1}^n \sum_{i=1}^n \left(\partial^2 y \right)_{ij} = \sum_{j=1}^n \sum_{i=1}^n \frac{\partial^2 y}{\partial x_j \partial x_i} dx_j dx_i. \quad (45)$$

Причём для частных *смешанных* производных порядка $p \geq 2$ от скалярной функции $y = y(\mathbf{x})$ последовательность дифференцирования по частным *скалярным* переменным, как известно, значения не имеет.

Вышеуказанные скалярные дифференциальные выражения порядка 1 и 2 (а также подобные выражения ещё более высоких порядков)

целесообразно представлять в краткой форме записи, используя для этого *тензор-производные* от функции $y(\mathbf{x})$ по $\mathbf{x}$ соответствующего порядка и $n \times 1$ -вектор дифференциала аргумента $d\mathbf{x}$ с элементами $dx_1, dx_2, \dots, dx_n$ и его транспонированную $1 \times n$ форму:

$$dy = \frac{dy}{d\mathbf{x}} d\mathbf{x}, \quad d^2y = d\mathbf{x}' \frac{d^2y}{d\mathbf{x}' d\mathbf{x}} d\mathbf{x}. \quad (46, 47)$$

Тензор-производная от скалярной функции $y(\mathbf{x})$ по $\mathbf{x}$ порядка p при $p = 1$ есть её *градиент*; при $p = 2$ есть её симметричная *матрица Гессе* и вообще при $p > 1$ есть её функциональная симметричная p -мерная матрица всех частных производных порядка p . В дифференциальных выражениях (46), (47)

$$d\mathbf{x} = \begin{pmatrix} dx_1 \\ dx_2 \\ \dots \\ dx_n \end{pmatrix}$$

есть $n \times 1$ -вектор дифференциала $\mathbf{x}$, при этом, строго говоря, $\mathbf{x}$ в $\mathcal{A}^n$ в данной аффинной системе координат определяется именно как радиус-вектор;

$$\frac{dy}{d\mathbf{x}} = \left(\frac{\partial y}{\partial x_1} \quad \frac{\partial y}{\partial x_2} \quad \dots \quad \frac{\partial y}{\partial x_n} \right) = \left(\frac{\partial y}{\partial x_i} \right), \quad (i = \overline{1, n})$$

есть $1 \times n$ -вектор градиента, или 1-я тензор-производная функции $y(\mathbf{x})$;

$$\frac{d^2y}{d\mathbf{x}' d\mathbf{x}} = \begin{pmatrix} \frac{\partial^2 y}{\partial x_1^2} & \frac{\partial^2 y}{\partial x_1 \partial x_2} & \dots & \frac{\partial^2 y}{\partial x_1 \partial x_n} \\ \frac{\partial^2 y}{\partial x_2 \partial x_1} & \frac{\partial^2 y}{\partial x_2^2} & \dots & \frac{\partial^2 y}{\partial x_2 \partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial^2 y}{\partial x_n \partial x_1} & \frac{\partial^2 y}{\partial x_n \partial x_2} & \dots & \frac{\partial^2 y}{\partial x_n^2} \end{pmatrix} = \left(\frac{\partial^2 y}{\partial x_j \partial x_i} \right), \quad (i, j = \overline{1, n})$$

есть симметричная $n \times n$ -матрица Гессе, или 2-я тензор-производная функции $y(\mathbf{x})$. Её детерминант именуется как *гессиан*.

Нетрудно видеть, что при формальном перемножении по известным *правилам линейной алгебры* этих тензор-производных на дифференциал аргумента — один раз в (46) и два раза в (47) получаются линейная (44) и квадратичная (45) дифференциальные формы.

С учётом этих обозначений, теорема Эйлера — Ферма (43) и правила Эйлера (41), (42) приводятся к компактному виду, сходному по форме с ранее изложенными правилами (1)—(3):

$$dy = \frac{dy}{dx}(s)dx = 0 \leftrightarrow \frac{dy}{dx}(s) = 0; \quad (48)$$

$$dy = \frac{dy}{dx}(s^+)dx = 0, \quad d^2y = dx' \frac{d^2y}{dx'dx}(s^+)dx < 0; \quad (49)$$

$$dy = \frac{dy}{dx}(s^-)dx = 0, \quad d^2y = dx' \frac{d^2y}{dx'dx}(s^-)dx > 0. \quad (50)$$

Геометрический смысл обеих формул (48) легко демонстрируется именно в *метрическом евклидовом координатном пространстве* $\mathcal{E}^n$. Как было принято ещё в начале данного параграфа, $\mathbf{x}$ есть *независимая* векторная переменная в координатном пространстве (радиус-вектор). Поэтому она и её 1-й дифференциал $d\mathbf{x}$ принимают *любые* направления в некоторой допустимой области $\mathcal{T}^n$. Следовательно, для обнуления 1-го дифференциала функции dy её градиент должен быть ортогонален $d\mathbf{x}$, принимающему всевозможные направления. Но такое требование реализуется геометрически тогда и только тогда, когда градиент $dy/d\mathbf{x}$ в точке стационарности нулевой.

В случае *нестрогой* стационарности (экстремума) целевой функции $y(\mathbf{x})$ 2-го уровня тут имеется некоторое непрерывное множество $\langle s \rangle$ решений уравнения (48). Соответственно в формулах (49) и (50) для дифференциала d^2y применяются знаки неравенств $\leq$ и $\geq$.

Отметим то особое обстоятельство, что для $p \geq 2$ непрерывно дифференцируемой $y(\mathbf{x})$ или, гораздо менее общо, — для аналитической $y(\mathbf{x})$, в отличие от таковой целевой функции от скалярного аргумента $y(x)$, стационарный перегиб также может реализовываться при *любом* целочисленном уровне $p \geq 2$ (в некоторой точке $s^\pm$). Стационарный перегиб в многомерном варианте общепринято называется *седловиной*. Геометрически наглядно седловина проявляется для функции $y(\mathbf{x})$, заданной на $\mathcal{A}^2$. В случае стационарной седловины 2-го уровня для $y(\mathbf{x})$ в точке $s^\pm$ её 2-й дифференциал $d^2y(\mathbf{x}=s^\pm)$, отсчитываемый от этой точки, принимает и положительное, и отрицательное, и, возможно, нулевое значение в зависимости от направления вектора $d\mathbf{x}$:

$$\frac{dy}{dx}(s^\pm)dx = 0, \quad dx' \frac{d^2y}{dx'dx}(s^\pm)dx \geq \wedge \leq 0. \quad (51)$$

Естественно, возникает вопрос об оценке каким-либо достаточно простым способом знака или знаков d^2y в (49)–(51), так как вычислять все эти знаки по всем мыслимым направлениям dx в $\mathcal{T}^n$ совершенно нереально.

Из теории квадратичных форм (см., например, [25 и 31]) известно следующее. Для того чтобы данная $n \times n$ -матрица была отрицательно определённая как в (49), необходимо и достаточно, чтобы все её n собственных значений (спектр) были отрицательные. Для того чтобы данная $n \times n$ -матрица была положительно определённая как в (50), необходимо и достаточно, чтобы все её n собственных значений (спектр) были положительные. Разумеется, все эти правила относятся именно к *симметричным* матрицам, каковой, в частности, является матрица Гессе. Но, однако, в случае *нестроого* экстремума матрица Гессе *полуопределённая* — либо отрицательно (для нестроого максимума), либо положительно (для нестроого минимума); при этом она имеет и нулевые собственные значения. В случае седловины как в (51) матрица Гессе *знаконеопределённая*; но при этом она имеет и положительные, и отрицательные, и, возможно, нулевые собственные значения.

Итак, один из способов оценки знака d^2y для реализации правил Эйлера сводится к вычислению собственных значений матрицы Гессе. Хотя, по сути, требуются только их знаки. Есть гораздо более простые возможности для оценки этих знаков и её знакоопределённости.

Пожалуй, самый простой и известный способ идентификации характера знакоопределённости для *симметричной* $n \times n$ -матрицы даёт **критерий Сильвестра**. Согласно данному критерию, о характере судят по знакам n её последовательных угловых главных, или диагональных миноров (т. е. детерминантов) размеров 1×1 , 2×2 , ..., $k \times k$, ..., $n \times n$, формируемых вдоль главной диагонали слева направо. Если все они положительные, то и матрица положительно определённая, и обратно. Если же все они последовательно знакопереваются в порядке знаков « $-$ », « $+$ », « $-$ », ... , то матрица отрицательно определённая, и обратно. Отсюда в первом случае её гессиан обязательно положительный, а во втором случае её гессиан имеет знак коэффициента $(-1)^n$.

Но симметричная $n \times n$ -матрица — *положительно полуопределённая* имеет положительные и нулевые диагональные миноры (в том числе все нулевые только размера более $g \times g$); — *отрицательно полуопределённая* имеет с переменным знаком коэффициента $(-1)^k$ и нулевые диагональные миноры (в том числе все нулевые только размера более $g \times g$).

Отметим, что теоретической основой этого критерия является **закон инерции квадратичных форм Сильвестра** (относительно базиса).

Иной способ оценки знакоопределённости матрицы Гессе в точке стационарности осуществляется по знакам коэффициентов её *векового уравнения* с использованием классического **правила знаков Декарта** и того обстоятельства, что для вещественной *симметричной* матрицы её собственные значения (корни уравнения) обязательно *вещественные*. Напомним, что коэффициенты векового уравнения в *знакопередающей* форме (см. гл. 4) для любой квадратной матрицы (или же её *скалярные характеристические коэффициенты*) суть полные суммы всевозможных диагональных миноров этой матрицы одного и того же с ними порядка. Если все скалярные коэффициенты положительные, то матрица Гессе положительно определённая (при этом $d^2y > 0$). Если же коэффициенты знакопередаются (как выше миноры), то матрица Гессе отрицательно определённая (при этом $d^2y < 0$). Но если они далее обнуляются при превышении некоторого порядка r (ранга симметричной матрицы), то матрица Гессе знакополуопределённая. В иных случаях матрица Гессе *законеопределённая*. Учитывая прямую взаимосвязь коэффициентов и *характеристических следов* матрицы (§ 4.4), аналогичные признаки выражаются через *следы* степеней матрицы Гессе от 1 до n .

Покажем инвариантность вышеизложенных правил относительно *линейных* модальных преобразований V ($\det V \neq 0$), а, следовательно, и их общую применимость для оптимизации скалярной функции $y(\mathbf{x})$ на *аффинном* координатном пространстве $\mathcal{A}^n$. Пусть новый базис выражается через исходный как $\tilde{\mathbf{E}}_Z = V \cdot \tilde{\mathbf{E}}_X$. Тогда имеем:

$$\mathbf{x} = V\mathbf{z} \rightarrow \frac{dy}{d\mathbf{x}} = \frac{dy}{d(V\mathbf{z})} = \frac{dy}{d\mathbf{z}} \cdot V^{-1} = \mathbf{0} = \frac{dy}{d\mathbf{z}},$$

$$\frac{d^2y}{d\mathbf{x}'d\mathbf{x}} = \frac{d^2y}{d(V\mathbf{z})'d(V\mathbf{z})} = V'^{-1} \cdot \frac{d^2y}{d\mathbf{z}'d\mathbf{z}} \cdot V^{-1}.$$

Следовательно, во-первых, обнуление обоих градиентов происходит здесь всегда в эквивалентных точках $\mathbf{s}$ и $V^{-1}\mathbf{s}$; во-вторых, в силу несингулярности модальной матрицы, матрицы Гессе в обоих базисах знакоопределяются одинаково (т. е. знаки их собственных значений).

Более того, в следующем параграфе будет отдельно показано, что инвариантность этих правил имеет место и по отношению к общим *нелинейным* модальным преобразованиям. (Но степени их гладкости и регулярности, в принципе, должны отвечать уровню экстремума.)

* * *

Пример 1. Проанализировать с точки зрения оптимизации скалярную целевую функцию 2-го порядка от $\mathbf{x}$:

$$q = q(\mathbf{x}) = 1/2 \cdot \mathbf{x}' G \mathbf{x} + \mathbf{g}'_0 \cdot \mathbf{x} + q_0 \quad (\det G \neq 0).$$

$$\left(\frac{dq}{d\mathbf{x}} \right)' = G\mathbf{x} + \mathbf{g}_0 = \mathbf{0} \rightarrow \mathbf{x} = \mathbf{s} = -G^{-1} \mathbf{g}_0.$$

В точке $\mathbf{s}$ имеем:

- максимум, если G отрицательно определённая;
- минимум, если G положительно определённая;
- стационарную седловину, если G знаконеопределённая.

Экстремальное значение целевой функции выражается как:

$$q(\mathbf{s}) = -1/2 \cdot \mathbf{g}'_0 G^{-1} \mathbf{g}_0 + q_0.$$

Пример 2. Выявить и проанализировать экстремумы целевых функций y_1 и y_2 соответственно для разности и для отношения среднего арифметического и среднего геометрического от двух вещественных чисел $x_1 > 0$ и $x_2 > 0$.

Экстремальные задачи такого рода, как хорошо известно, при количестве переменных $n = 2$ решаются тривиальным путём исходя из положительности x_1 и x_2 :

$$y_1 = \frac{x_1 + x_2}{2} - \sqrt{x_1 \cdot x_2} = \left(\frac{\sqrt{x_1} - \sqrt{x_2}}{2} \right)^2 \geq 0,$$

$$y_2 = \frac{x_1 + x_2}{2 \cdot \sqrt{x_1 \cdot x_2}} = \frac{1}{4} \cdot \left(\sqrt{\frac{x_1}{x_2}} + \sqrt{\frac{x_2}{x_1}} + 2 \right) = \frac{1}{4} (t + 1/t + 2) \geq 1.$$

Отсюда сразу же следует, что обе функции, неограниченные сверху, имеют минимумы $y_1 = 0$ и $y_2 = 1$ при $x_1 = x_2 > 0$. Однако такие тривиальные способы решения задачи при $n > 2$ не реализуются.

Аналитический способ, напротив, является универсальным, т. е. он может применяться при любом $n \geq 2$. В наиболее полном виде аналитический способ решения таких задач будет продемонстрирован в гл. 4 при доказательстве *генерального неравенства для средних величин*. Здесь же этот способ демонстрируется только на примере поставленной конкретной частной задачи при $n = 2$. Причём, в силу положительности x_1 и x_2 , в функциях y_1 и y_2 для упрощения решения и анализа применяются квадраты средних величин.

$$1) y_1 = y_1(\mathbf{x}) = y_1(x_1, x_2) = \left(\frac{x_1 + x_2}{2}\right)^2 - (\sqrt{x_1 \cdot x_2})^2 = \left(\frac{x_1 - x_2}{2}\right)^2.$$

$$\mathbf{g} = \mathbf{g}(\mathbf{x}) = \frac{dy_1}{d\mathbf{x}} = \left(\frac{\partial y_1}{\partial x_1}, \frac{\partial y_1}{\partial x_2}\right) = (1/2 \cdot (x_1 - x_2); -1/2 \cdot (x_2 - x_1)) = \mathbf{0} \rightarrow$$

$$\rightarrow \mathbf{x} = \mathbf{s}: x_1 = x_2 > 0, y_1(\mathbf{s}) = 0;$$

$$G = G(\mathbf{x}) = \frac{d^2 y_1}{d\mathbf{x}' d\mathbf{x}} = \begin{pmatrix} \frac{\partial^2 y_1}{\partial x_1^2}, & \frac{\partial^2 y_1}{\partial x_1 \partial x_2} \\ \frac{\partial^2 y_1}{\partial x_1 \partial x_2}, & \frac{\partial^2 y_1}{\partial x_2^2} \end{pmatrix} = \begin{pmatrix} 1/2, & -1/2 \\ -1/2, & 1/2 \end{pmatrix} = \text{Const},$$

$$\text{tr}G = 1/2 + 1/2 = 1 > 0, \quad \det G = 1/4 - 1/4 = 0.$$

Итак, целевая функция $y_1(x_1, x_2)$ имеет тут нестрогий минимум на биссектрисе 1-го квадранта, т. е. при $x_1^- = x_2^- > 0$ (рис. 6). Её матрица Гессе здесь вырождена именно вдоль этой биссектрисы. Поверхность $y_1(x_1, x_2)$ по своей геометрии в целом полувогнутая.

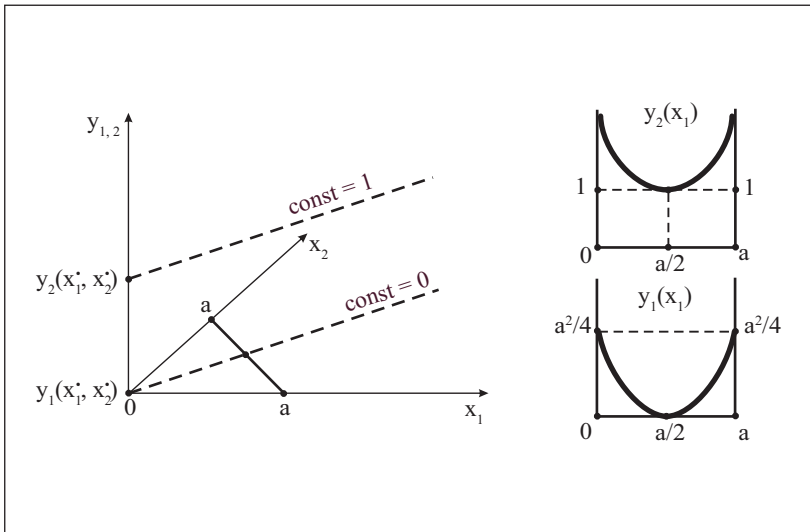


Рис. 6. К рассмотрению функций разности (y_1) и отношения (y_2) среднего арифметического и среднего геометрического в аффинном 3-х координатном и 2-х координатном (профильном) базисах при двух числовых переменных x_1 и $x_2 > 0$.

$$2) y_2 = y_2(\mathbf{x}) = y_2(x_1, x_2) = \left(\frac{x_1 + x_2}{2} / \sqrt{x_1 \cdot x_2} \right)^2 = 1/4 \cdot \left(\frac{x_1}{x_2} + \frac{x_2}{x_1} + 2 \right).$$

$$\mathbf{g} = \mathbf{g}(\mathbf{x}) = \frac{dy_2}{d\mathbf{x}} = \left(\frac{\partial y_2}{\partial x_1}, \frac{\partial y_2}{\partial x_2} \right) = \left(1/4 \cdot \left(x_2 - x_2 / x_1^2 \right); 1/4 \cdot \left(x_1 - x_1 / x_2^2 \right) \right) = \mathbf{0} \rightarrow$$

$$\rightarrow \mathbf{x} = \mathbf{s}: x_1 = x_2 > 0; y_2(\mathbf{s}) = 1;$$

$$G = G(\mathbf{x}) = \frac{d^2 y_2}{d\mathbf{x}^T d\mathbf{x}} = \begin{pmatrix} \frac{\partial^2 y_2}{\partial x_1^2}, & \frac{\partial^2 y_2}{\partial x_1 \partial x_2} \\ \frac{\partial^2 y_2}{\partial x_1 \partial x_2}, & \frac{\partial^2 y_2}{\partial x_2^2} \end{pmatrix} =$$

$$= \begin{pmatrix} x_2 / 2x_1^3, & -1/4 \cdot (1/x_1^2 + 1/x_2^2) \\ -1/4 \cdot (1/x_1^2 + 1/x_2^2), & x_1 / 2x_2^3 \end{pmatrix},$$

$$\text{tr}G(\mathbf{x}) = x_2 / 2x_1^3 + x_1 / 2x_2^3 = (x_1^4 + x_2^4) / 2x_1^3 x_2^3 > 0,$$

$$\det G(\mathbf{x}) = 1/16 \cdot \left[4/x_1^2 x_2^2 - (1/x_1^2 + 1/x_2^2)^2 \right] = -1/16 \cdot (1/x_1^2 - 1/x_2^2)^2 \leq 0.$$

И здесь $y_2(x_1, x_2)$ имеет нестрогий минимум на биссектрисе 1-го квадранта, т. е. при $x_1^- = x_2^- = u > 0$ (рис. 6), где матрица Гессе вырождена, а поверхность $y_2(x_1, x_2)$ локально полувогнутая, так как

$$G(\mathbf{s}) = 1/2u \cdot \begin{pmatrix} +1, & -1 \\ -1, & +1 \end{pmatrix}, \quad \text{tr}G(\mathbf{s}) = 1/u^2 > 0, \quad \det G(\mathbf{s}) = 0.$$

Но вне этой биссектрисы (!) поверхность $y_2(x_1, x_2)$ в пределах 1-го квадранта имеет повсюду седловинную форму. Исключением является только начало системы координат, где эта поверхность асимптотически приближается к оси ординат ($y_2 \geq 1$). Особо отметим, что поверхность $y_2(x_1, x_2)$ имеет семейство образующих линий в виде параллелей на координатной плоскости $\langle x_1, x_2 \rangle$ с возрастанием их высоты (ординаты) при удалении от биссектрисы 1-го квадранта. Все они исходят от оси координат — каждая со своей высотой: $y_2 = 1/4 \cdot (k + 1/k + 2) = \text{const}$, где $k = x_2/x_1$. В частности, для биссектрисы 1-го квадранта имеем $k = 1$, $y_2 = 1$; но при $k \rightarrow 0$ или при $k \rightarrow \infty$ имеем $y_2 = \text{const} \rightarrow \infty$. Вековое уравнение для матрицы Гессе $\mu^2 - \text{tr}G \cdot \mu + \det G = 0$ имеет 2 решения:

$$\mu_{1,2} = \frac{x_1^4 + x_2^4}{x_1^3 \cdot x_2^3} \pm \sqrt{\left(\frac{x_1^4 + x_2^4}{x_1^3 \cdot x_2^3} \right)^2 + \left(\frac{1}{x_1^2} - \frac{1}{x_2^2} \right)^2}.$$

В точках вне биссектрисы 1-го квадранта собственные значения μ_1 и μ_2 суть разнознаковые, а на биссектрисе $\mu_1 > 0$, $\mu_2 = 0$. Отсюда вытекает весьма важный вывод. *Решение данной задачи наглядно иллюстрирует тот факт, что для идентификации нестрогой стационарности целевой функции $y(\mathbf{x})$ нужно знать знакоопределённость её матрицы Гессе именно на области стационарности, а не вообще в её окрестности!* Каков же конкретно этот нестрогий экстремум — глобальный или локальный устанавливается затем дополнительным исследованием. Так, в этой конкретной задаче имеем глобальный нестрогий минимум целевой функции $y_2(x_1, x_2)$ на биссектрисе 1-го квадрата, поскольку здесь любой другой точке вне биссектрисы отвечает проходящая через неё образующая поверхности с постоянным значением на ней целевой функции, большим 1 (см. выше), и на которой её градиент ненулевой.

Интересно также проследить изменение целевой функции $y_2(x_1, x_2)$ в направлении $\mathbf{x}$, перпендикулярном к биссектрисе 1-го квадранта, например, в зависимости от x_1 (см. рис. 6):

$$y_2(x_1) = 1/4 \cdot [x_1 / (a - x_1) + (a - x_1) / x_1 + 2] = a^2 / 4x_1(a - x_1);$$

$$\frac{dy_2}{dx_1^2} = 1/4 \cdot [a / (a - x_1)^2 - a / x_1^2] = 0 \rightarrow x_1^- = a/2,$$

$$\frac{d^2 y_2}{dx_1^2} = 1/2 \cdot [(a - x_1)^3 + x_1^3] / (a - x_1)^3 x_1^3 > 0; \quad \min y_2 = 1.$$

Рассмотренная выше весьма замечательная поверхность $y_2(x_1, x_2)$ дополняется своим симметричным отражением в 3-м квадранте, где обе величины x_1 и x_2 и их те же самые средние формально отрицательные. Чисто геометрически поверхность получается в результате совместного вращения вокруг оси y_2 и поступательного движения вверх по её направлению вверх и вниз прямолинейной образующей, исходящей из начала координат и параллельной плоскости $\langle x_1, x_2 \rangle$. При равномерной скорости её вращения в направлении по или против часовой стрелки скорость движения этой образующей по вертикали возрастает по мере приближения её проекции к оси x_1 или x_2 . Среди всех положений этой образующей единственное исходное положение как биссектрисы 1-го и 3-го квадрантов на уровне $y_2 = 1$ является областью стационарности функции $y_2(x_1, x_2)$, так как только на данной области поверхности градиент и, следовательно, производные по направлениям нулевые!

Глава 2. Аналитическая условная оптимизация

§ 2.1. Условные экстремумы уровня $p = 2$

для функций от зависимой переменной типа

$$\tilde{\mathbf{x}} \in \mathcal{A}^n: \tilde{\mathbf{x}} = \mathbf{x}(\mathbf{u}), \mathbf{u} \in \mathcal{A}^q, q \leq n$$

В данном параграфе $n \times 1$ -вектор $\tilde{\mathbf{x}}$ рассматривается как *зависимая* переменная от $q \times 1$ -вектора $\mathbf{u}$ — некоей независимой переменной, или аргумента, где $q \leq n$. Суть задачи заключается в аналитическом поиске *условного* экстремума эволюционной целевой функции y на гладком и регулярном *вложенном многообразии* $\langle \tilde{\mathbf{x}} \rangle \subset \mathcal{T}^n \subset \mathcal{A}^n$, заданном *параметрическим* способом. Причём $\mathcal{T}^n \subset \mathcal{A}^n$ есть некая n -мерная закрытая область. Для заданной целевой двухступенчатой сложной функции $y = y[\mathbf{x}(\mathbf{u})] = f(\mathbf{u})$ аргумент $\mathbf{u} \in \mathcal{A}^q$ (параметр) может принимать любые значения из области определения функции f в $\mathcal{A}^q$.

Пусть $y = y(\mathbf{x})$ и $\tilde{\mathbf{x}} = \mathbf{x}(\mathbf{u})$ — скалярная и векторная функции, однозначно определённые и, по крайней мере, дважды непрерывно дифференцируемые на $\mathcal{T}^n \subset \mathcal{A}^n$ и соответственно на $\mathcal{A}^q$. Причём (см. далее) имеет место непрерывное и полное отображение переменной $\mathbf{u}$ из $\mathcal{A}^q$ в некоторую геометрическую поверхность $\langle \tilde{\mathbf{x}} \rangle$ в $\mathcal{A}^n$, где $q \leq n$. Однако особенно здесь подчеркнём то, что при дифференцировании по переменной $\tilde{\mathbf{x}}$ она понимается обычным образом — как *свободная* переменная $\mathbf{x}$, изменяющаяся от своего точечного значения, но на $\langle \tilde{\mathbf{x}} \rangle$, во всевозможных направлениях в координатном пространстве $\mathcal{A}^n$! По правилу дифференцирования двухступенчатой сложной функции находим её 1-й полный дифференциал и 1-ю тензор-производную по $\mathbf{u}$.

$$dy = \frac{dy}{d\tilde{\mathbf{x}}} d\tilde{\mathbf{x}} = \frac{dy}{d\tilde{\mathbf{x}}} \cdot \frac{d\tilde{\mathbf{x}}}{d\mathbf{u}} d\mathbf{u}, \left(d\tilde{\mathbf{x}} = \frac{d\tilde{\mathbf{x}}}{d\mathbf{u}} d\mathbf{u} \right), \quad (52)$$

где $d\tilde{\mathbf{x}} = \begin{pmatrix} \frac{\partial \tilde{x}_1}{\partial \tilde{x}_2} \\ \dots \\ \frac{\partial \tilde{x}_n}{\partial \tilde{x}_2} \end{pmatrix}$ — условный 1-й дифференциал $\mathbf{x}$, отсчитываемый, как

и свободный дифференциал $d\mathbf{x}$, от какой-либо точки $\tilde{\mathbf{c}} \in \langle \tilde{\mathbf{x}} \rangle$. Различие между $d\mathbf{x}$ и $d\tilde{\mathbf{x}}$ состоит в том, что первый из них принимает, как указано выше, любые направления в $\mathcal{A}^n$ в точке $\tilde{\mathbf{c}} \in \langle \tilde{\mathbf{x}} \rangle$, а второй — только те направления здесь же в $\mathcal{A}^n$, которые производятся дифференциалом $d\mathbf{u}$ при его отображении в $\mathcal{A}^n$. В иной форме имеем:

$$\frac{d\mathbf{y}}{d\mathbf{u}} = \frac{d\mathbf{y}}{d\tilde{\mathbf{x}}} \cdot \frac{d\tilde{\mathbf{x}}}{d\mathbf{u}}. \quad (53)$$

$1 \times q \quad 1 \times n \quad n \times q$

В (52) и (53) фигурируют дифференциальные характеристики:

$\frac{d\mathbf{y}}{d\mathbf{x}}$ — градиент $y(\tilde{\mathbf{x}})$, или 1-я тензор-производная функции $y(\tilde{\mathbf{x}})$;

$\frac{d\tilde{\mathbf{x}}}{d\mathbf{u}}$ — $n \times q$ -матрица Якоби, или 1-я тензор-производная функции $\mathbf{x}(\mathbf{u})$.

Соотношения (52), (53) мнемонически и по смыслу аналогичны первым соотношениям из систем (22), (23) для скалярных переменных $\tilde{x}$ и u . Структура $n \times q$ -матрицы Якоби, данная ниже, получается в результате последовательного частного дифференцирования $n \times 1$ -вектора $\tilde{\mathbf{x}}$ в $\mathcal{A}^n$ по компонентам $q \times 1$ -вектора $\mathbf{u}$ в $\mathcal{A}^q$. Поэтому матрица принадлежит аффинному пространству $\mathcal{A}^{n \times q}$. В матричной форме она имеет вид:

$$\frac{d\tilde{\mathbf{x}}}{d\mathbf{u}} = \begin{pmatrix} \frac{\partial \tilde{x}_1}{\partial u_1} & \frac{\partial \tilde{x}_1}{\partial u_2} & \dots & \frac{\partial \tilde{x}_1}{\partial u_q} \\ \frac{\partial \tilde{x}_2}{\partial u_1} & \frac{\partial \tilde{x}_2}{\partial u_2} & \dots & \frac{\partial \tilde{x}_2}{\partial u_q} \\ \dots & \dots & \dots & \dots \\ \frac{\partial \tilde{x}_n}{\partial u_1} & \frac{\partial \tilde{x}_n}{\partial u_2} & \dots & \frac{\partial \tilde{x}_n}{\partial u_q} \end{pmatrix} = \left(\frac{\partial \tilde{x}_i}{\partial u_j} \right), \quad (i = \overline{1, n}; \quad j = \overline{1, q}).$$

В случае $q = n$ матрица Якоби принимает квадратную форму, а её детерминант тогда называется *якобианом*. Отдельный и весьма важный частный случай отвечает здесь *постоянной* $n \times q$ -матрице Якоби, т. е. прямоугольной числовой матрице A размера $n \times q$ и ранга q . Например, этот случай реализуется на вложенной q -плоскости $\langle \tilde{\mathbf{x}} \rangle$: $\tilde{\mathbf{x}} = A\mathbf{u} + \mathbf{b}$. Тогда матрица Якоби выражается как $d\tilde{\mathbf{x}} / d\mathbf{u} = A = \text{Const}$.

Более общий случай отвечает $p \times q$ -матрице Якоби *постоянного* ранга $q < n$ и *конечной* по величине, что обеспечивает тут непрерывное и изоморфное ранга q *векторное отображение* $\tilde{\mathbf{x}} = \mathbf{x}(\mathbf{u})$ параметра $\mathbf{u}$ из $\mathcal{A}^q$ в $\mathcal{A}^n$. Её дифференцируемость k раз обеспечивает *гладкость* порядка k , конечность тензор-производных до порядка k обеспечивает *регулярность* порядка k . Вектор-столбцы $p \times q$ -матрицы Якоби ранга q суть линейно независимые. Пространство $\mathcal{A}^q$ изоморфно отображается тогда в гладкую и регулярную q -поверхность $\langle \tilde{\mathbf{x}} \rangle$ в $\mathcal{A}^n$. Отметим здесь, что альтернативным способом та же q -поверхность $\langle \tilde{\mathbf{x}} \rangle$ в $\mathcal{A}^n$ задаётся через *сужающее отображение* $p \times 1$ -векторной переменной $\mathbf{u}$ из $\mathcal{A}^n$ в $\mathcal{A}^q$; $p \times p$ -матрица Якоби сингулярная и имеет постоянный ранг $q < p$.

Если $d\tilde{\mathbf{x}} / d\mathbf{u}$ повсюду существует и имеет постоянный ранг q , то она *полнозначимая*, а отображение $\mathbf{x}(\mathbf{u})$ и q -поверхность $\langle \tilde{\mathbf{x}} \rangle$ суть *гладкие* и *регулярные* порядка, по крайней мере, 1 и ранга q .

Пусть $y = y[\mathbf{x}(\mathbf{u})] = f(\mathbf{u})$ стационарная в точке $\mathbf{s} \in \mathcal{A}^q$. Согласно правилу Эйлера — Ферма (48) и соотношению (52), для *условной стационарности* $y(\mathbf{x})$ необходимо и достаточно, чтобы выполнялись нижеуказанные уравнения, а градиент $dy/d\mathbf{x}$ в точке $\tilde{\mathbf{s}}$ был ненулевым:

$$df(\mathbf{u} = \mathbf{s}) = \frac{dy}{d\mathbf{u}}(\mathbf{s}) d\mathbf{u} = \frac{dy}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) d\tilde{\mathbf{x}} = \frac{dy}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \cdot \frac{d\tilde{\mathbf{x}}}{d\mathbf{u}}(\mathbf{s}) d\mathbf{u} = \mathbf{0} \leftrightarrow \quad (54)$$

$1 \times q \quad q \times 1 \quad 1 \times n \quad n \times 1 \quad 1 \times n \quad n \times q \quad q \times 1$

$$\leftrightarrow \frac{dy}{d\mathbf{u}}(\mathbf{s}) = \frac{dy}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \cdot \frac{d\tilde{\mathbf{x}}}{d\mathbf{u}}(\mathbf{s}) = \mathbf{0}. \quad (55)$$

В уравнениях стационарности (54), (55) значения тензор-производных и дифференциалов даны в *эквивалентных* точках $\mathbf{s} \in \mathcal{A}^q$ и $\tilde{\mathbf{s}} \in \mathcal{A}^n$.

Идентификация условного экстремума или седловины 2-го уровня для данной целевой двухступенчатой сложной функции в самом общем случае выполняется через анализ либо её 2-го полного дифференциала, либо её 2-й тензор-производной. Последние выводятся аналогично тем же характеристикам в скалярных формулах (26) и (27) из § 1.5, но, с учётом правил дифференцирования по векторным переменным:

$$d^2 f(\mathbf{u} = \mathbf{s}) = d\mathbf{u}' \frac{d^2 y}{d\mathbf{u}' d\mathbf{u}}(\mathbf{s}) d\mathbf{u} = d^2 y(\tilde{\mathbf{x}} = \tilde{\mathbf{s}}) = d\tilde{\mathbf{x}}' \frac{d^2 y}{d\tilde{\mathbf{x}}' d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) d\tilde{\mathbf{x}} + \frac{dy}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \cdot d^2 \tilde{\mathbf{x}} \leftrightarrow$$

$1 \times q \quad q \times q \quad q \times 1 \quad 1 \times n \quad n \times n \quad n \times 1 \quad 1 \times n \quad n \times 1$

$$\leftrightarrow \frac{d^2 y}{d\mathbf{u}' d\mathbf{u}}(\mathbf{s}) = \left(\frac{d\tilde{\mathbf{x}}}{d\mathbf{u}}(\mathbf{s}) \right)' \cdot \frac{d^2 y}{d\tilde{\mathbf{x}}' d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \cdot \frac{d\tilde{\mathbf{x}}}{d\mathbf{u}}(\mathbf{s}) + \frac{dy}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \cdot \frac{d^2 \tilde{\mathbf{x}}}{d\mathbf{u}' d\mathbf{u}}(\mathbf{s}); \quad (56, 57)$$

$q \times q \quad q \times n \quad n \times n \quad n \times q \quad 1 \times n \quad n \times q \times q$

где $d^2\tilde{\mathbf{x}} = \begin{vmatrix} \partial^2\tilde{x}_1 \\ \partial^2\tilde{x}_2 \\ \dots \\ \partial^2\tilde{x}_n \end{vmatrix}$ — условный 2-й дифференциал $\tilde{\mathbf{x}}$, $\frac{d^2\tilde{\mathbf{x}}}{d\mathbf{u}'d\mathbf{u}}$ — 2-я тензор-

производная функции $\tilde{\mathbf{x}} = \mathbf{x}(\mathbf{u})$.

Последняя производная представляет собой исходно трёхмерную $n \times q \times q$ -матрицу, состоящую из 2-ых частных производных:

$$\frac{\partial^2 x_i}{\partial u_k \partial u_j} \quad (i = \overline{1, n}; \quad j, k = \overline{1, q}).$$

Как трёхмерная матрица, она имеет три типа плоских сечений:

- 1) фронтальные $n \times q$ -сечения $\left(\frac{\partial^2 x_i}{\partial u_k \partial u_j} \right)$, $(i = \overline{1, n}; \quad j = \overline{1, q}; \quad k = \text{const})$;
- 2) боковые $q \times n$ -сечения $\left(\frac{\partial^2 x_i}{\partial u_k \partial u_j} \right)$, $(i = \overline{1, n}; \quad k = \overline{1, q}; \quad j = \text{const})$;
- 3) горизонтальные $q \times q$ -сечения $\left(\frac{\partial^2 x_i}{\partial u_k \partial u_j} \right)$, $(j, k = \overline{1, q}; \quad i = \text{const})$.

(Последние из них обязательно симметричные.)

По сути, они отображают возможность существования трёх форм трёхмерной матрицы — одной исходной и двух транспонированных. Однако из-за перестановочности частных *скалярных* дифференциалов ∂u_j и ∂u_k первые две формы идентичны друг другу с точностью до двумерного транспонирования, в том числе и по своим сечениям при $j = k$. Следовательно, для вычисления, например, второго слагаемого в (57) $1 \times n$ -вектор градиента поэтапно q раз умножается слева на q фронтальных сечений (т. е. при k от 1 до q), производя в результате обязательно симметричную $q \times q$ -матрицу из векторов-строк произведения $(d\mathbf{y} / d\tilde{\mathbf{x}}) \cdot (\partial^2 \tilde{\mathbf{x}} / \partial u_k \partial \mathbf{u})$; $k = \overline{1, q}$. В свою очередь, первое слагаемое в (57) есть тоже симметричная $q \times q$ -матрица, вычисляемая по известным правилам линейной алгебры. Таким образом, после суммирования обоих слагаемых в (57) в итоге вычисляется опять-таки симметричная $q \times q$ -матрица Гессе $d^2 \mathbf{y} / d\mathbf{u}' d\mathbf{u} (\mathbf{s})$.

Когда все дифференциальные характеристики в уравнениях (54)–(57) существуют, тогда, в принципе, возможны три *непредельных* варианта стационарности целевой функции, а, следовательно, и её экстремума.

1) Либо $dy / d\tilde{\mathbf{x}}(\tilde{\mathbf{s}}) = \mathbf{0}$ и $d\tilde{\mathbf{x}} / d\mathbf{u}$ существует. Отображение $\tilde{\mathbf{x}} = \mathbf{x}(\mathbf{u})$ регулярное, причём, по крайней мере, в окрестности точки $\mathbf{s}$. Это соответствует стационарности целевой функции в $\tilde{\mathbf{s}}$ 1-й степени, или её безусловной стационарности.

2) Либо $dy / d\tilde{\mathbf{x}}(\tilde{\mathbf{s}}) \neq \mathbf{0}$ и существует (значимая), но $d\tilde{\mathbf{x}} / d\mathbf{u}(\mathbf{s}) = \mathbf{Z}$. Отображение $\tilde{\mathbf{x}} = \mathbf{x}(\mathbf{u})$ в самой точке $\mathbf{s}$ негладкое, но регулярное. Это соответствует стационарности целевой функции в $\tilde{\mathbf{s}}$ 2-й степени. Она реализуется в *особых* точках негладкой поверхности $\tilde{\mathbf{x}} = \mathbf{x}(\mathbf{u})$, в которых матрица Якоби отображения обнуляется. В этих точках отображение осуществляется деформацией бесконечного сжатия по всем направлениям дифференциала $d\mathbf{u}$ в $\mathcal{A}^q$. Деформация растяжения-сжатия в конкретном направлении $d\tilde{\mathbf{x}}$ в $\mathcal{A}^n$ и $d\mathbf{u}$ в $\mathcal{A}^q$ оценивается через отношение Релея в следующей форме:

$$\mathcal{R}e(d\mathbf{u}) = \frac{d\tilde{\mathbf{x}}' d\tilde{\mathbf{x}}}{d\mathbf{u}' d\mathbf{u}} = \frac{d\mathbf{u}' \left[\left(\frac{d\tilde{\mathbf{x}}}{d\mathbf{u}} \right)' \cdot \frac{d\tilde{\mathbf{x}}}{d\mathbf{u}} \right] d\mathbf{u}}{d\mathbf{u}' d\mathbf{u}}. \quad (58)$$

В рассматриваемом тут крайнем варианте (нулевой матрицы Якоби) отношение Релея в точке $\mathbf{s}$ нулевое во всевозможных направлениях дифференциала аргумента $\mathbf{u}$.

Выше пока имелась аналогия с первым и вторым вариантом для случая скалярных переменных (см. § 1.5). Далее аналогии нет.

3) Либо, как в варианте 2, градиент в точке $\mathbf{s}$ *значимый*, причём $\frac{dy}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \in \langle \ker \frac{d\tilde{\mathbf{x}}}{d\mathbf{u}}(\mathbf{s}) \rangle_L$, при этом матрица $d\tilde{\mathbf{x}} / d\mathbf{u}$ также *значимая*,

но $1 \leq \text{rang} \{ d\tilde{\mathbf{x}} / d\mathbf{u} \} \leq q$, т. е. отображение $\tilde{\mathbf{x}} = \mathbf{x}(\mathbf{u})$ полугладкое или гладкое и регулярное (по крайней мере, в окрестности точки $\mathbf{s}$). Это соответствует тут *полуусловной* стационарности целевой функции в $\tilde{\mathbf{s}}$. В этом варианте матрица Якоби функции отображения $\mathbf{x}(\mathbf{u})$ обязательно сингулярная *слева* (на что указывает нижний индекс L у её ядра). Последнее имеет место либо по причине того, что $q < n$, либо по причине того, что лишь только $< q$ её столбцов линейно независимые, либо по обоим этим причинам.

Пусть конкретнее матрица Якоби *полнозначимая*, т. е. она конечная и $1 \leq \text{rang} \{ d\tilde{\mathbf{x}} / d\mathbf{u} \} = q = \text{const}$ в точке $\mathbf{s}$ и, по крайней мере, на некоторой её окрестности в $\mathcal{A}^q$ (По сути, её ранг равен размерности $\mathbf{u}$)

Отсюда $q \times q$ -матрица *внутренней гомомультпликации* матрицы Якоби в (58) на окрестности $\mathbf{s}$ обязательно несингулярная и существует. (Обе матрицы здесь *полнозначимые*). Тогда отношение Релея (58) есть положительная величина для любого направления $\mathbf{du}$ в $\mathcal{A}^q$ и, кроме того, имеет глобальные максимум и минимум — см. об этом в § 4.6. Следовательно, при изоморфном преобразовании дифференциала $\mathbf{du}$ в (52) осуществляется деформация конкретного сжатия-растяжения. В этом весьма важном субварианте варианта 3 реализуется *истинная условная стационарность* $y = y(\tilde{\mathbf{x}})$ на заданной именно параметрическим способом области $\langle \tilde{\mathbf{x}} \rangle \subset \mathcal{A}^n$. Геометрически она трактуется в $\mathcal{A}^n$ так. В точке $\tilde{\mathbf{s}}$ вдоль допустимых направлений *условного* дифференциала $d\tilde{\mathbf{x}} \neq \mathbf{0}$ линейная часть приращения целевой функции всегда нулевая, но по другим направлениям *свободного* дифференциала $d\mathbf{x} \neq \mathbf{0}$ она ненулевая. (Разумеется, *безусловная стационарность* целевой функции $y = y(\tilde{\mathbf{x}})$ реализуется только по 1-му варианту — см. выше.)

Теперь пусть $1 \leq \text{rang}\{d\tilde{\mathbf{x}} / d\mathbf{u}\} < q$ в точке $\mathbf{s}$ и, по крайней мере, на некоторой её окрестности в $\mathcal{A}^q$. Тогда, в силу сингулярности $q \times q$ -матрицы внутренней гомомультпликации в (58), отношение Релея для $\mathbf{du}$ в направлении её сингулярных собственных векторов в $\mathcal{A}^q$ нулевое. Соответственно изоморфизма дифференциалов $d\tilde{\mathbf{x}}$ в $\mathcal{A}^n$ и $\mathbf{du}$ в $\mathcal{A}^q$ при этом нет. Причём по некоторым направлениям вектора $\mathbf{du}$ в $\mathcal{A}^q$ в результате преобразования дифференциала аргумента, согласно (52), происходит деформация бесконечного сжатия. (Отношение Релея есть тут положительно полуопределённая, или неотрицательная величина.) Тогда в этом субварианте варианта 3 проявляется *смешанная условная стационарность* целевой функции. Она реализуется частично от того, что линейное приращение такой целевой функции вдоль допускаемых направлений $d\tilde{\mathbf{x}} \neq \mathbf{0}$ нулевое, и частично от того, что дифференциал $\mathbf{du}$ при его отображении подвергается деформации бесконечного сжатия.

Пусть в *первом варианте* стационарности $d^2\tilde{\mathbf{x}} / d\mathbf{u}^2 d\mathbf{u}(\mathbf{s})$ также существует (как и 1-я тензор-производная). Тогда, согласно (57), имеем:

$$\frac{d^2 y}{d\mathbf{u}^2 d\mathbf{u}}(\mathbf{s}) = \left(\frac{d\tilde{\mathbf{x}}}{d\mathbf{u}}(\mathbf{s}) \right)' \cdot \frac{d^2 y}{d\tilde{\mathbf{x}}' d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \cdot \frac{d\tilde{\mathbf{x}}}{d\mathbf{u}}(\mathbf{s}).$$

Если данная 2-я тензор-производная ненулевая, то по характеру её знакоопределённости здесь идентифицируется либо тот или иной *безусловный экстремум*, либо *безусловная седловина* (1-й степени, 2-го уровня) с применением правил Эйлера (49)—(51) и критерия Сильвестра.

Обе фигурирующие тут матрицы Гессе имеют один и тот же характер знакоопределённости, если $\text{rang}\{\mathbf{d}\tilde{\mathbf{x}}/\mathbf{d}\mathbf{u}\} = q$, так как конгруэнтное преобразование никак не затрагивает знакоопределённость матрицы. При $\text{rang}\{\mathbf{d}\tilde{\mathbf{x}}/\mathbf{d}\mathbf{u}\} < q$ присутствуют нулевые собственные значения у матрицы Гессе, вследствие чего стационарность является нестрогой.

Представляет особый интерес такой субвариант варианта 1, когда $\text{rang}\{\mathbf{d}\tilde{\mathbf{x}}/\mathbf{d}\mathbf{u}\} = q = n$. Он, по сути, отвечает *нелинейному модальному преобразованию* базиса в задаче на *безусловный* экстремум (см. об этом в конце § 1.9); причём $\det\{\mathbf{d}\tilde{\mathbf{x}}/\mathbf{d}\mathbf{u}\} \neq 0$, что обязательно для модальной матрицы. В случае нелинейного модального преобразования имеется изоморфное отображение одной и той же *безусловной стационарности* целевой функции из одного $\mathcal{A}^n$ в другое $\mathcal{A}^n$, которые связаны между собой нелинейно, как и переменные $\mathbf{x}$ и $\mathbf{z}$:

$$\frac{dy}{d\mathbf{x}} = \frac{dy}{d[\mathbf{x}(\mathbf{z})]} = \frac{dy}{dz} \cdot \left(\frac{d\mathbf{x}}{dz}\right)^{-1} = \mathbf{0} = \frac{dy}{dz},$$

$$\frac{d^2y}{d\mathbf{x}'d\mathbf{x}} = \frac{d^2y}{d[\mathbf{x}(\mathbf{z})]'d[\mathbf{x}(\mathbf{z})]} = \left(\frac{d\mathbf{x}}{dz}\right)^{-1} \cdot \frac{d^2y}{dz'dz} \cdot \left(\frac{d\mathbf{x}}{dz}\right)^{-1}.$$

Следовательно, во-первых, обращение в нуль градиентов $dy/d\mathbf{x}$ и dy/dz происходит всегда в эквивалентных точках $\mathbf{x}$ и $\mathbf{z}(\mathbf{x})$; во-вторых, в силу несингулярности матрицы Якоби $d\mathbf{x}/dz$, матрицы Гессе $d^2y/d\mathbf{x}'d\mathbf{x}$ и $d^2y/dz'dz$ знакоопределены всегда одинаково. Обратим внимание здесь на тот весьма интересный факт, что модальное преобразование для *тензор-производных функции* — как линейное, так и нелинейное, осуществляется по одним и тем же правилам. Попросту говоря, при линейном преобразовании базиса $\tilde{\mathbf{E}}_Z = \mathbf{V} \cdot \tilde{\mathbf{E}}_X$ его матрица $\mathbf{V} = \mathbf{A}$ играет роль той же матрицы Якоби для отображения $\mathbf{x}(\mathbf{z}) = \mathbf{A}\mathbf{z}$!

Пусть во *втором варианте* стационарности $d^2y/d\tilde{\mathbf{x}}'d\tilde{\mathbf{x}}(\tilde{\mathbf{s}})$ также существует, как и 1-я условная тензор-производная. Тогда, согласно (57), имеем:

$$\frac{d^2y}{du'du}(\mathbf{s}) = \frac{dy}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \cdot \frac{d^2\tilde{\mathbf{x}}}{du'du}(\mathbf{s}).$$

Если данная 2-я тензор-производная ненулевая, то по характеру её знакоопределённости идентифицируется либо тот или иной особый экстремум, либо особая седловина (2-й степени и 2-го уровня) с применением правил Эйлера (49)—(51) и критерия Сильвестра.

Наконец, в *третьем варианте* стационарности итоговая матрица Гессе выражается самой общей формулой (57). Если данная 2-я тензор-производная ненулевая, то по характеру её знакоопределённости идентифицируется либо *условный экстремум*, либо *условная седловина* (2-го уровня) с применением правил Эйлера (49)—(51) и критерия Сильвестра. Здесь представляет отдельный интерес также субвариант варианта 3, когда $d\tilde{\mathbf{x}} / d\mathbf{u} = \mathbf{A} = \text{Const}$ размера $n \times q$ и ранга q . Тогда, во-первых, в точке *истинной* условной стационарности $\tilde{\mathbf{s}}$ функции на q -плоскости $\langle \tilde{\mathbf{x}} \rangle$ имеем $\frac{dy}{d\mathbf{u}}(\mathbf{s}) = \frac{dy}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \cdot \mathbf{A} = \mathbf{0} \leftrightarrow \left(\frac{dy}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \right)' \in \langle \ker \mathbf{A}' \rangle$

и, во-вторых, в силу того что матрица $d^2\tilde{\mathbf{x}} / d\mathbf{u}'d\mathbf{u}$ нулевая, имеем $\frac{d^2y}{d\mathbf{u}'d\mathbf{u}}(\mathbf{s}) = \mathbf{A}' \cdot \frac{d^2y}{d\tilde{\mathbf{x}}'d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \cdot \mathbf{A}$.

* * *

Полная адекватность природы стационарности 2-го порядка функций $y = y(\tilde{\mathbf{x}})$ на q -поверхности $\tilde{\mathbf{x}} = \mathbf{x}(\mathbf{u})$ в $\mathcal{A}^n$ и $y = f(\mathbf{u})$ в $\mathcal{A}^q$ имеет место тогда, когда тензор-производные $d\tilde{\mathbf{x}} / d\mathbf{u}$ и $d^2\tilde{\mathbf{x}} / d\mathbf{u}'d\mathbf{u}$ повсюду ранга q и существуют (*полнозначимые*), т. е. при гладком и регулярном ранга q *изоморфном отображении* из $\mathcal{A}^q$ в $\mathcal{A}^n$, по крайней мере, порядка 2. Этот случай наиболее интересен в теоретическом и практическом плане. Если здесь $q < n$, то тогда имеем условную стационарность $y(\tilde{\mathbf{x}})$ на $\langle \tilde{\mathbf{x}} \rangle \subset \mathcal{A}^n$ и стационарность $f(\mathbf{u})$ в $\mathcal{A}^q$ в эквивалентных точках $\tilde{\mathbf{s}}$ и $\mathbf{s}$. При этом дифференциалы $d\tilde{\mathbf{x}}$ и $d\mathbf{u}$ в (59) суть *изоморфизмы*:

$$d\tilde{\mathbf{x}} = \frac{d\tilde{\mathbf{x}}}{d\mathbf{u}} d\mathbf{u} \leftrightarrow d\mathbf{u} = \frac{d\mathbf{u}}{d\tilde{\mathbf{x}}} d\tilde{\mathbf{x}}, \quad (59)$$

$$\frac{d\mathbf{u}}{d\tilde{\mathbf{x}}} = \left(\frac{d\tilde{\mathbf{x}}}{d\mathbf{u}} \right)_L^{-1} = \left(\frac{d\tilde{\mathbf{x}}}{d\mathbf{u}} \right)^+ = \left[\left(\frac{d\tilde{\mathbf{x}}}{d\mathbf{u}} \right)' \cdot \frac{d\tilde{\mathbf{x}}}{d\mathbf{u}} \right]^{-1} \cdot \left(\frac{d\tilde{\mathbf{x}}}{d\mathbf{u}} \right)'. \quad (60)$$

Именно в этом случае ($1 \leq \text{rang}(d\tilde{\mathbf{x}} / d\mathbf{u}) = q$) однозначно определяется и вычисляется функциональная матрица Якоби для *квазиобратного отображения* $d\mathbf{u} / d\tilde{\mathbf{x}}$ как единственная левая обратная матрица $(d\tilde{\mathbf{x}} / d\mathbf{u})_L^{-1}$, или также как тождественная ей *квазиобратная матрица* Мура — Пенроуза $(d\tilde{\mathbf{x}} / d\mathbf{u})^+$, единственная тут *правая квазиобратная матрица* вообще [27, с. 52]. Причём прямая и обратная матрицы Якоби связаны здесь совершенно естественным образом и через единичную $q \times q$ -матрицу, и через симметричный $n \times n$ -проектор.

Имеем соответствующие формулы перемножения матриц:

$$\frac{d\mathbf{u}}{d\tilde{\mathbf{x}}} \cdot \frac{d\tilde{\mathbf{x}}}{d\mathbf{u}} = I_{q \times q}, \quad \frac{d\tilde{\mathbf{x}}}{d\mathbf{u}} \cdot \frac{d\mathbf{u}}{d\tilde{\mathbf{x}}} = \left[\frac{d\tilde{\mathbf{x}}}{d\mathbf{u}} \cdot \left(\frac{d\tilde{\mathbf{x}}}{d\mathbf{u}} \right)' \right]^{-1}. \quad (61, 62)$$

В формуле (62) фигурирует симметричный характеристический проектор [27, с. 50], осуществляющий аффинное проецирование в $\mathcal{A}^n$ на $\langle \text{im } d\tilde{\mathbf{x}} / d\mathbf{u} \rangle$ параллельно $\langle \ker d\tilde{\mathbf{x}} / d\mathbf{u} \rangle_L$, т. е. на образ матрицы Якоби параллельно её ядру (слева). Этот симметричный проектор вычисляется теоретически однозначно через сингулярную $n \times n$ -матрицу внешней гомомультимпликации матрицы Якоби в формуле (62). С его использованием геометрически наглядным образом определяются все условные дифференциальные характеристики 1-го порядка, а также формулируется иллюстративное проективное уравнение для условной стационарности в задачах на условный экстремум данного типа.

Линеаризация размерности q для $q \times 1$ -вектор-функции отображения $\tilde{\mathbf{x}} = \mathbf{x}(\mathbf{u})$ в точке $\tilde{\mathbf{c}} \in \langle \tilde{\mathbf{x}} \rangle$ определяется здесь её линейной частью, а геометрически — касательной q -плоскостью $\langle l \rangle$ к q -поверхности $\langle \tilde{\mathbf{x}} \rangle$:

$$l(\mathbf{u}) = \tilde{\mathbf{c}} + d\tilde{\mathbf{x}}(\mathbf{u} = \mathbf{c}) = \tilde{\mathbf{c}} + \frac{d\tilde{\mathbf{x}}}{d\mathbf{u}}(\mathbf{c}) (\mathbf{u} - \mathbf{c}) \rightarrow \langle l \rangle \equiv \tilde{\mathbf{c}} + \left\langle \text{im } \frac{d\tilde{\mathbf{x}}}{d\mathbf{u}}(\mathbf{c}) \right\rangle. \quad (63)$$

Лемма 1 (о линеаризации). Линеаризация размерности q для $q \times 1$ -вектор-функции $\tilde{\mathbf{x}} = \mathbf{x}(\mathbf{u})$ имеет место в точке $\tilde{\mathbf{c}} \in \langle \tilde{\mathbf{x}} \rangle$ тогда и только тогда, когда $n \times q$ -матрица Якоби $d\tilde{\mathbf{x}} / d\mathbf{u}$ в эквивалентной ей точке $\mathbf{c}$ существует и на некоторой её окрестности имеет постоянный ранг $q \geq 1$, т. е. состоит из q линейно независимых векторов-столбцов, или $\langle \ker d\tilde{\mathbf{x}} / d\mathbf{u} \rangle = \mathbf{0}$. (Доказательство этой леммы вполне тривиально.)

Условный 1-й дифференциал переменной $d\tilde{\mathbf{x}} \equiv d\mathbf{l}$ в соответствии с (52) и (63) определяется как:

$$d\tilde{\mathbf{x}} = \frac{d\tilde{\mathbf{x}}}{d\mathbf{u}} d\mathbf{u} = \left[\frac{d\tilde{\mathbf{x}}}{d\mathbf{u}} \cdot \left(\frac{d\tilde{\mathbf{x}}}{d\mathbf{u}} \right)' \right]^{-1} d\mathbf{x}. \quad (64)$$

Условный 1-й дифференциал целевой функции $y(\tilde{\mathbf{x}})$ определяется и обозначается как:

$$[dy] = \frac{dy}{d\tilde{\mathbf{x}}} \cdot \frac{d\tilde{\mathbf{x}}}{d\mathbf{x}} d\mathbf{u} = \frac{dy}{d\tilde{\mathbf{x}}} d\tilde{\mathbf{x}} = \frac{dy}{d\tilde{\mathbf{x}}} \cdot \left[\frac{d\tilde{\mathbf{x}}}{d\mathbf{u}} \cdot \left(\frac{d\tilde{\mathbf{x}}}{d\mathbf{u}} \right)' \right]^{-1} d\mathbf{x}. \quad (65)$$

Условный градиент целевой функции $y(\tilde{\mathbf{x}})$ определяется здесь и обозначается как:

$$\left[\frac{dy}{dx} \right] = \frac{dy}{d\tilde{\mathbf{x}}} \cdot \frac{d\tilde{\mathbf{x}}}{dx} = \frac{dy}{d\tilde{\mathbf{x}}} \cdot \left[\frac{d\tilde{\mathbf{x}}}{du} \cdot \left(\frac{d\tilde{\mathbf{x}}}{du} \right)' \right]^{\leftarrow}. \quad (66)$$

Условный градиент геометрически трактуется как аффинная в $\mathcal{A}^n$ проекция градиента на $\langle \text{im } d\tilde{\mathbf{x}} / du \rangle$ (т. е. касательную q -плоскость $\langle \mathcal{I} \rangle$ к q -поверхности $\langle \tilde{\mathbf{x}} \rangle$) параллельно $\langle \ker d\tilde{\mathbf{x}} / du \rangle_L$.

Отсюда вытекает формулируемая ниже *проективная теорема об условной стационарности* для задачи на условный экстремум данного типа, или условный аналог теоремы Эйлера — Ферма (43).

Теорема 1 (проективная). Для того чтобы в точке $\tilde{\mathbf{s}} \in \langle \tilde{\mathbf{x}} \rangle$ имела место условная стационарность целевой функции $y(\tilde{\mathbf{x}})$, необходимо и достаточно, чтобы её условный градиент в данной точке равнялся нулю, а сам градиент в ней был ненулевым:

$$\left[dy(\tilde{\mathbf{x}} = \tilde{\mathbf{s}}) \right] = 0 \leftrightarrow \left[\frac{dy}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \right] = \mathbf{0} \leftrightarrow \left(\frac{dy}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \right)' \in \left\langle \ker \frac{d\tilde{\mathbf{x}}}{du}(\tilde{\mathbf{s}}) \right\rangle_L; \left(\frac{dy}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \neq \mathbf{0} \right).$$

Срединное выражение выше есть проективный вариант уравнения условной стационарности. Правое выражение даёт геометрическую интерпретацию формулы (55) для точки условной стационарности.

А именно, в точке условной стационарности обнуление 1-го дифференциала целевой функции происходит в силу того, что ненулевой градиент в ней коллинеарен некоему левому сингулярному собственному вектору матрицы Якоби $d\tilde{\mathbf{x}} / du$, т. е. условный градиент находится в линейном подпространстве ядра матрицы Якоби (слева). Поэтому и вышеуказанная аффинная проекция градиента на её образ параллельно её ядру (слева) по величине нулевая. (В евклидовом метрическом пространстве $\mathcal{E}^n$ с декартовыми координатами этот факт геометрическим образом интерпретируется как *ортogonalность* градиента касательной q -плоскости в точке $\tilde{\mathbf{s}}$, или всем векторам-столбцам матрицы Якоби.) Понятно, что в полной мере условный аналог теоремы Эйлера — Ферма (43) здесь не имеет места, так как нужно ещё убедиться в том, что найденная точка $\tilde{\mathbf{s}}$ (где условный градиент целевой функции нулевой) принадлежит именно q -поверхности $\langle \tilde{\mathbf{x}} \rangle$. Это требование выполняется, если существует *обратное* отображение $\tilde{\mathbf{s}} \rightarrow \mathbf{s}$, т. е. отображение точки стационарности из $\mathcal{A}^n$ в $\mathcal{A}^q$.

Далее обратимся к условным дифференциальным характеристикам 2-го порядка. Если $1 \leq \text{rang}\{\mathbf{d}\tilde{\mathbf{x}} / \mathbf{d}\mathbf{u}\} = q$, то по-прежнему $\mathbf{d}\mathbf{u}$ и $\mathbf{d}\tilde{\mathbf{x}}$ изоморфны. Более того, каждому $\mathbf{d}\mathbf{u}$ однозначно отвечают $\mathbf{d}\tilde{\mathbf{x}}$ и $\mathbf{d}^2\tilde{\mathbf{x}}$ от функции отображения $\tilde{\mathbf{x}} = \mathbf{x}(\mathbf{u})$, заданной на $\mathcal{A}^q$, а каждому $\mathbf{d}\tilde{\mathbf{x}}$ однозначно отвечают $\mathbf{d}\mathbf{u}$ и $\mathbf{d}^2\mathbf{u}$ от обратной функции отображения $\mathbf{u} = \mathbf{u}(\tilde{\mathbf{x}})$, заданной на $\langle \tilde{\mathbf{x}} \rangle$. Покажем, что при таком подходе $\mathbf{d}^2\mathbf{u}$ и $\mathbf{d}^2\tilde{\mathbf{x}}$ также изоморфны (т. е. как это было ранее для скалярных переменных u и x в § 1.6). Имеем:

$$\mathbf{d}^2\tilde{\mathbf{x}} = \mathbf{d}\mathbf{u}' \left(\frac{\mathbf{d}^2\tilde{\mathbf{x}}}{\mathbf{d}\mathbf{u}'\mathbf{d}\mathbf{u}} \right)^{(1 \leftrightarrow 2)} \quad \mathbf{d}\mathbf{u} \leftrightarrow \mathbf{d}^2\mathbf{u} = \mathbf{d}\tilde{\mathbf{x}}' \left(\frac{\mathbf{d}^2\mathbf{u}}{\mathbf{d}\tilde{\mathbf{x}}'\mathbf{d}\tilde{\mathbf{x}}} \right)^{(1 \leftrightarrow 2)} \mathbf{d}\tilde{\mathbf{x}}, \quad (67)$$

$$\frac{\mathbf{d}^2\mathbf{u}}{\mathbf{d}\tilde{\mathbf{x}}'\mathbf{d}\tilde{\mathbf{x}}} = -\frac{\mathbf{d}\mathbf{u}}{\mathbf{d}\tilde{\mathbf{x}}} \cdot \left[\left(\frac{\mathbf{d}\mathbf{u}}{\mathbf{d}\tilde{\mathbf{x}}} \right)' \cdot \left(\frac{\mathbf{d}^2\tilde{\mathbf{x}}}{\mathbf{d}\mathbf{u}'\mathbf{d}\mathbf{u}} \right)^{(1 \leftrightarrow 2)} \cdot \frac{\mathbf{d}\mathbf{u}}{\mathbf{d}\tilde{\mathbf{x}}} \right], \quad (68)$$

$$\mathbf{d}^2\tilde{\mathbf{x}} = -\frac{\mathbf{d}\tilde{\mathbf{x}}}{\mathbf{d}\mathbf{u}} \mathbf{d}^2\mathbf{u} \neq \mathbf{0} \leftrightarrow \mathbf{d}^2\mathbf{u} = -\frac{\mathbf{d}\mathbf{u}}{\mathbf{d}\tilde{\mathbf{x}}} \cdot \mathbf{d}^2\tilde{\mathbf{x}} \neq \mathbf{0}. \quad (69)$$

Здесь и дальше знак сверху « $'(1 \leftrightarrow 2)$ » обозначает транспонирование трёхмерной матрицы по типу $(1 \leftrightarrow 2)$, т. е. с перестановкой элементов с 1-ми и 2-ми одинаковыми индексами; причём $\mathbf{d}\mathbf{u} / \mathbf{d}\tilde{\mathbf{x}}$ вычисляется по формуле (60). Тут имеется мнемоническая аналогия со скалярной формулой (31). При этом заметим, что, если бы связи между 1-ми дифференциалами переменных $\mathbf{u}$ и $\tilde{\mathbf{x}}$ не было, то тогда $\mathbf{d}^2\mathbf{u}$ был бы нулевым! В свою очередь, условный дифференциал целевой функции 2-го порядка в точке $\tilde{\mathbf{s}}$ определяется с учётом (56) как:

$$\begin{aligned} \left[\mathbf{d}^2 y(\tilde{\mathbf{x}} = \tilde{\mathbf{s}}) \right] &= \mathbf{d}\tilde{\mathbf{x}}' \frac{\mathbf{d}^2 y}{\mathbf{d}\tilde{\mathbf{x}}'\mathbf{d}\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \mathbf{d}\tilde{\mathbf{x}} + \frac{\mathbf{d}y}{\mathbf{d}\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \mathbf{d}^2\tilde{\mathbf{x}} = \\ &= \mathbf{d}\tilde{\mathbf{x}}' \left\{ \frac{\mathbf{d}^2 y}{\mathbf{d}\tilde{\mathbf{x}}'\mathbf{d}\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) + \frac{\mathbf{d}y}{\mathbf{d}\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \cdot \left[\left(\frac{\mathbf{d}\mathbf{u}}{\mathbf{d}\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \right)' \cdot \left(\frac{\mathbf{d}^2\tilde{\mathbf{x}}}{\mathbf{d}\mathbf{u}'\mathbf{d}\mathbf{u}}(\mathbf{s}) \right)^{(1 \leftrightarrow 2)} \cdot \frac{\mathbf{d}\mathbf{u}}{\mathbf{d}\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \right] \right\} \mathbf{d}\tilde{\mathbf{x}}. \quad (70) \end{aligned}$$

$1 \times n$
 $n \times n$
 $1 \times n$
 $n \times q$
 $q \times n \times q$
 $q \times n$
 $n \times 1$

Вспомогательные размеры объектов, выделенные жирным шрифтом, здесь и дальше указывают направление перемножения тензор-объекта с трёхмерной матрицей, т. е. по каким индексам происходит свёртка тензорных производных при их перемножении. Свёртка происходит ввиду того, что n элементов вектора градиента поэтапно умножаются на n симметричных $q \times q$ -сечений трёхмерной матрицы, а затем всё это суммируется в двумерную симметричную $q \times q$ -матрицу в скобках.

Разумеется, изложенный *проективный* подход к анализу задачи на условный экстремум этого типа имеет чисто иллюстративное значение. Однако именно *проективный* подход позволяет связать воедино все три рассматриваемых в монографии аналитических метода постановки и решения задачи на условный экстремум.

Полное аналитическое решение задачи этого типа осуществимо на практике изложенным двухступенчатым дифференциальным методом. В общем виде решение задачи заключается в нахождении условного экстремума целевой функции на некотором нелинейном или линейном образе $\langle \tilde{\mathbf{x}} \rangle$ преобразования $\mathbf{x} = \mathbf{x}(\mathbf{u})$. В задачах на экстремум данного типа сложность их анализа и решения увеличивается именно с ростом размерности q векторного параметра $\mathbf{u}$. Наиболее просто они решаются при $q = 1$, т. е. как задачи на условный экстремум целевой функции на некоторой *траектории* или *прямой*, вложенной в $\mathcal{A}^n$. При $q = 1$ тензор-производные $d\tilde{\mathbf{x}}/d\mathbf{u}$ и $d^2\tilde{\mathbf{x}}/d\mathbf{u}'d\mathbf{u}$ попросту вырождаются в вектор-производные $d\tilde{\mathbf{x}}/du$ и $d^2\tilde{\mathbf{x}}/du^2$. Причём, согласно (55), в $\mathcal{E}^n$ векторы $(dy/d\tilde{\mathbf{x}})'$ и $d\tilde{\mathbf{x}}/du$ в точке стационарности ортогональны!

Покажем *инвариантность вышеизложенных процедур решения задач на условный экстремум данного типа по отношению к линейным модальным преобразованиям базисов* в $\mathcal{A}^n$ и в $\mathcal{A}^q$.

Для двухступенчатого метода установим инвариантность формул (55) и (57) к преобразованиям $\tilde{\mathbf{E}}_Z = V_1 \cdot \tilde{\mathbf{E}}_X$ и $\tilde{\mathbf{E}}_t = V_2 \cdot \tilde{\mathbf{E}}_u$:

$$1) \mathbf{x} = V_1 \cdot \mathbf{z} \rightarrow \frac{dy}{d(V_1\tilde{\mathbf{z}})} \cdot \frac{d(V_1\tilde{\mathbf{z}})}{d\mathbf{u}} = \frac{dy}{d\tilde{\mathbf{z}}} V_1^{-1} \cdot V_1 \frac{d\tilde{\mathbf{z}}}{d\mathbf{u}} = \frac{dy}{d\tilde{\mathbf{z}}} \cdot \frac{d\tilde{\mathbf{z}}}{d\mathbf{u}} = \mathbf{0},$$

$$\begin{aligned} \frac{d^2y}{d\mathbf{u}'d\mathbf{u}} &= \left[\frac{d(V_1\tilde{\mathbf{z}})}{d\mathbf{u}} \right]' \cdot \frac{d^2y}{d(V_1\tilde{\mathbf{z}})'d(V_1\tilde{\mathbf{z}})} \cdot \frac{d(V_1\tilde{\mathbf{z}})}{d\mathbf{u}} + \frac{dy}{d(V_1\tilde{\mathbf{z}})} \cdot \frac{d^2(V_1\tilde{\mathbf{z}})}{d\mathbf{u}'d\mathbf{u}} = \\ &= \left(\frac{d\tilde{\mathbf{z}}}{d\mathbf{u}} \right)' V_1' \cdot V_1'^{-1} \cdot \frac{d^2y}{d\tilde{\mathbf{z}}'d\tilde{\mathbf{z}}} V_1^{-1} \cdot V_1 + \frac{dy}{d\tilde{\mathbf{z}}} V_1^{-1} \cdot V_1 \frac{d^2\tilde{\mathbf{z}}}{d\mathbf{u}'d\mathbf{u}} = \\ &= \left(\frac{d\tilde{\mathbf{z}}}{d\mathbf{u}} \right)' \cdot \frac{d^2y}{d\tilde{\mathbf{z}}'d\tilde{\mathbf{z}}} \cdot \frac{d\tilde{\mathbf{z}}}{d\mathbf{u}} + \frac{dy}{d\tilde{\mathbf{z}}} \cdot \frac{d^2\tilde{\mathbf{z}}}{d\mathbf{u}'d\mathbf{u}}; \end{aligned}$$

$$2) \mathbf{u} = V_2 \cdot \mathbf{t} \rightarrow \frac{dy}{d\tilde{\mathbf{z}}} \cdot \frac{d\mathbf{z}}{d(V_2\mathbf{t})} = \frac{dy}{d\tilde{\mathbf{z}}} \cdot \frac{d\mathbf{z}}{d\mathbf{t}} V_2^{-1} = \mathbf{0} = \frac{dy}{d\tilde{\mathbf{z}}} \cdot \frac{d\tilde{\mathbf{z}}}{d\mathbf{t}},$$

$$\frac{d^2y}{d\mathbf{u}'d\mathbf{u}} = \frac{d^2y}{d(V_2\mathbf{t})'d(V_2\mathbf{t})} = V_2' \cdot \frac{d^2y}{d\mathbf{t}'d\mathbf{t}} V_2^{-1}.$$

Итак, в пространстве $\mathcal{A}^n$ вообще нет никакого влияния от линейного преобразования переменной (так как итоговые производные берутся по переменной $\mathbf{u}$). В $\mathcal{A}^q$ влияние имеется только на матрицу Гессе. Но её конгруэнтное преобразование при этом никак не изменяет характер знакоопределённости матрицы, а, следовательно, и стационарности.

Для доказательства аналогичной инвариантности в проективной версии анализа задачи проектор применяется в форме (62) с учётом (61):

$$\left[\frac{d\tilde{\mathbf{x}}}{d\mathbf{u}} \cdot \left(\frac{d\tilde{\mathbf{x}}}{d\mathbf{u}} \right)' \right] = \frac{d\tilde{\mathbf{x}}}{d\mathbf{u}} \cdot \frac{d\mathbf{u}}{d\tilde{\mathbf{x}}} = \frac{d\tilde{\mathbf{x}}}{d\mathbf{u}} \cdot \left(\frac{d\tilde{\mathbf{x}}}{d\mathbf{u}} \right)_L^{-1}.$$

Далее из (64), (65) и (71) имеем:

$$\begin{aligned} d\tilde{\mathbf{x}} &= \left[\frac{d\tilde{\mathbf{x}}}{d\mathbf{u}} \cdot \left(\frac{d\tilde{\mathbf{x}}}{d\mathbf{u}} \right)' \right] d\mathbf{x} = \frac{d\tilde{\mathbf{x}}}{d\mathbf{u}} \cdot \left(\frac{d\tilde{\mathbf{x}}}{d\mathbf{u}} \right)_L^{-1} d\mathbf{x} = V_1 \frac{d\tilde{\mathbf{z}}}{dt} V_2^{-1} \cdot V_2 \left(\frac{d\tilde{\mathbf{z}}}{dt} \right)_L^{-1} \cdot V_1^{-1} \cdot V_1 d\mathbf{z} = \\ &= V_1 \left[\frac{d\tilde{\mathbf{z}}}{dt} \cdot \left(\frac{d\tilde{\mathbf{z}}}{dt} \right)' \right] d\mathbf{z} = V_1 d\tilde{\mathbf{z}} \leftrightarrow d\tilde{\mathbf{z}} = \left[\frac{d\tilde{\mathbf{z}}}{dt} \cdot \left(\frac{d\tilde{\mathbf{z}}}{dt} \right)' \right] d\mathbf{z}; \\ [d\mathbf{y}] &= \frac{d\mathbf{y}}{d\tilde{\mathbf{z}}} V_1^{-1} \cdot V_1 \left[\frac{d\tilde{\mathbf{z}}}{dt} \cdot \left(\frac{d\tilde{\mathbf{z}}}{dt} \right)' \right] d\mathbf{z} = \frac{d\mathbf{y}}{d\tilde{\mathbf{z}}} \cdot \left[\frac{d\tilde{\mathbf{z}}}{dt} \cdot \left(\frac{d\tilde{\mathbf{z}}}{dt} \right)' \right] d\mathbf{z} = \frac{d\mathbf{y}}{d\tilde{\mathbf{z}}} d\tilde{\mathbf{z}}; \\ \left[\frac{d\mathbf{y}}{d\mathbf{z}} \right] &= \frac{d\mathbf{y}}{d\tilde{\mathbf{z}}} \left[\frac{d\tilde{\mathbf{z}}}{dt} \cdot \left(\frac{d\tilde{\mathbf{z}}}{dt} \right)' \right], \quad \left[\frac{d\mathbf{y}}{d\mathbf{x}} \right] = \mathbf{0} \leftrightarrow \left[\frac{d\mathbf{y}}{d\mathbf{z}} \right] = \mathbf{0}; \\ \left[d^2\mathbf{y} \right] &= d\tilde{\mathbf{z}}' V_1' \left\{ V_1'^{-1} \frac{d^2\mathbf{y}}{d\tilde{\mathbf{z}}' d\tilde{\mathbf{z}}} V_1^{-1} + \frac{d\mathbf{y}}{d\tilde{\mathbf{z}}} V_1^{-1} \left[V_1'^{-1} \left(\frac{d\mathbf{t}}{d\tilde{\mathbf{z}}} \right)' V_1 \left(\frac{d^2\mathbf{z}}{dt' dt} \right)^{(1 \leftrightarrow 2)} \frac{d\mathbf{t}}{d\tilde{\mathbf{z}}} V_1^{-1} \right] \right\} \cdot \\ &\quad \cdot V_1 d\tilde{\mathbf{z}} = d\tilde{\mathbf{z}}' \left\{ \frac{d^2\mathbf{y}}{d\tilde{\mathbf{z}}' d\tilde{\mathbf{z}}} + \frac{d\mathbf{y}}{d\tilde{\mathbf{z}}} \cdot \left[\left(\frac{d\mathbf{t}}{d\tilde{\mathbf{z}}} \right)' \cdot \left(\frac{d^2\tilde{\mathbf{z}}}{dt' dt} \right)^{(1 \leftrightarrow 2)} \cdot \frac{d\mathbf{t}}{d\tilde{\mathbf{z}}} \right] \right\} d\tilde{\mathbf{z}}. \end{aligned}$$

Матрица V_1 аннигилирует тут, как показано выше (только в сложных случаях) жирными индексами снизу и сверху, а матрица V_2 аннигилирует тут совершенно естественным образом (поэтому здесь не показывается).

Матрица Гессе вычисляется как в (71):

$$\left[\frac{d^2 y}{dz' dz} \right]_{n \times n} = \left[\frac{d\tilde{\mathbf{z}}}{dt} \cdot \left(\frac{d\tilde{\mathbf{z}}}{dt} \right)' \right]_{n \times n}^{\leftarrow} \cdot \left\{ \frac{d^2 y}{d\tilde{\mathbf{z}}' d\tilde{\mathbf{z}}} + \frac{dy}{d\tilde{\mathbf{z}}} \cdot \left[\left(\frac{dt}{d\tilde{\mathbf{z}}} \right)' \cdot \left(\frac{d^2 \tilde{\mathbf{z}}}{dt' dt} \right)^{(1 \leftrightarrow 2)} \cdot \frac{dt}{d\tilde{\mathbf{z}}} \right]_{n \times q} \cdot \frac{dt}{d\tilde{\mathbf{z}}} \right\}_{q \times n \times q} \cdot \left[\frac{d\tilde{\mathbf{z}}}{dt} \cdot \left(\frac{d\tilde{\mathbf{z}}}{dt} \right)' \right]_{n \times n}^{\leftarrow}.$$

Доказательство инвариантности этих процедур к наиболее общим — *нелинейным* модальным преобразованиям $\tilde{\mathbf{x}}$ и $\mathbf{u}$ осуществляется также, но формально приводит к анализу стационарности *трёхступенчатой* сложной функции. Степени гладкости и регулярности нелинейных модальных преобразований должны отвечать тут уровню условной стационарности или экстремума (выше это $p = 2$).

* * *

Пример. Проанализируем с точки зрения условной оптимизации скалярную функцию 2-го порядка от линейно зависимой переменной

$$q = q(\tilde{\mathbf{x}}) = 1/2 \cdot \tilde{\mathbf{x}}' \cdot G \tilde{\mathbf{x}} + \mathbf{g}_0' \cdot \tilde{\mathbf{x}} + q_0 \quad (\det G \neq 0),$$

$$\tilde{\mathbf{x}} = \mathbf{x}(\mathbf{u}) = \mathbf{A}\mathbf{u} = \tilde{\mathbf{c}} + \mathbf{A}(\mathbf{u} - \mathbf{c}), \quad \text{где } \tilde{\mathbf{c}} = \mathbf{A}\mathbf{c}, \quad \text{rang } \mathbf{A} = q < n.$$

Без тривиальной подстановки $\mathbf{x} \rightarrow \mathbf{A}\mathbf{u}$ последовательно находим для $q(\tilde{\mathbf{x}})$ градиент по $\mathbf{u}$ и условный градиент по $\tilde{\mathbf{x}}$, матрицу Гессе по $\mathbf{u}$ и условную матрицу Гессе по $\tilde{\mathbf{x}}$. (Здесь очевидно $\mathbf{A}'\mathbf{g}_0' \in \langle \text{im } \mathbf{A}'\mathbf{G}\mathbf{A} \rangle$.)

$$\frac{dq}{d\mathbf{u}} = \frac{dq}{d\tilde{\mathbf{x}}} \cdot \frac{d\tilde{\mathbf{x}}}{d\mathbf{u}} \rightarrow$$

$$\left(\frac{dq}{d\mathbf{u}} \right)' = \mathbf{A}' \cdot (\mathbf{G}\tilde{\mathbf{x}} + \mathbf{g}_0') = \mathbf{A}' \cdot \mathbf{G}\tilde{\mathbf{x}} + \mathbf{A}'\mathbf{g}_0' = \mathbf{A}' \cdot \mathbf{G} \cdot \mathbf{A}\mathbf{u} + \mathbf{A}'\mathbf{g}_0',$$

$$\left(\frac{dq}{d\mathbf{u}}(\mathbf{s}) \right)' = \mathbf{A}' \cdot \mathbf{G} \cdot \mathbf{A}\mathbf{s} + \mathbf{A}'\mathbf{g}_0' = \mathbf{0} \rightarrow$$

$$\mathbf{s} = -(\mathbf{A}' \cdot \mathbf{G} \cdot \mathbf{A})^{-1} \cdot \mathbf{A}'\mathbf{g}_0' = -\mathbf{A}^{+} \cdot \mathbf{G}^{-1} \cdot \mathbf{A}'^{+} \cdot \mathbf{A}'\mathbf{g}_0' = -\mathbf{A}^{+} \cdot \mathbf{G}^{-1} \cdot \overset{\leftarrow}{\mathbf{A}}\mathbf{A}'\mathbf{g}_0',$$

$$\tilde{\mathbf{s}} = \mathbf{A}\mathbf{s} = -\left(\overset{\leftarrow}{\mathbf{A}}\mathbf{A}' \cdot \mathbf{G}^{-1} \cdot \overset{\leftarrow}{\mathbf{A}}\mathbf{A}' \right) \mathbf{g}_0',$$

$$\frac{d^2 q}{d\mathbf{u}' d\mathbf{u}} = \mathbf{A}'\mathbf{G}\mathbf{A} = \text{Const},$$

$$\left[\frac{d^2 q}{d\mathbf{x}' d\mathbf{x}} \right] = \left(\frac{d\mathbf{u}}{d\tilde{\mathbf{x}}} \right)' \cdot \frac{d^2 q}{d\mathbf{u}' d\mathbf{u}} \cdot \frac{d\mathbf{u}}{d\tilde{\mathbf{x}}} = \mathbf{A}'^{+} \cdot \mathbf{A}' \cdot \mathbf{G} \cdot \mathbf{A} \cdot \mathbf{A}^{+} = \overset{\leftarrow}{\mathbf{A}}\mathbf{A}' \cdot \mathbf{G} \cdot \overset{\leftarrow}{\mathbf{A}}\mathbf{A}' = \text{Const}.$$

Поскольку ($\det G \neq 0$), то в точке $\tilde{\mathbf{s}}$ имеем (уровня $p = 2$):

- условный минимум, если G или $[G]$ положительно определённая;
- условный максимум, если G или $[G]$ отрицательно определённая;
- условную стационарную седловину, если $[G]$ знаконеопределённая.

$$q(\tilde{\mathbf{s}}) = q[\mathbf{x}(\mathbf{s})] = -1/2 \cdot \mathbf{g}_0' \cdot \overleftarrow{\mathbf{A}} \mathbf{A}' \cdot G^{-1} \cdot \overleftarrow{\mathbf{A}} \mathbf{A}' \cdot \mathbf{g}_0' + q_0.$$

(Сравните результат с таковым в примере 1 из § 1.9.)

В следующем параграфе рассматривается постановка и решение задачи на условную стационарность целевой функции от ограниченной переменной, т. е. именно по Лагранжу, но с проективным подходом.

§ 2.2. Условные экстремумы уровня $p = 2$ для функций от ограниченной переменной типа $\tilde{\mathbf{x}} \in \mathcal{A}^n$: $\mathbf{h}(\tilde{\mathbf{x}}) = \mathbf{0}$, $\mathbf{h}(\mathbf{x}) \in \mathcal{A}^m$, $m \leq n$

В данном параграфе $n \times 1$ -вектор $\tilde{\mathbf{x}}$ рассматривается иначе — как *ограниченная* переменная, подчиняющаяся уравнению связи $\mathbf{h}(\tilde{\mathbf{x}}) = \mathbf{0}$, $\mathbf{h}(\mathbf{x} \notin \langle \tilde{\mathbf{x}} \rangle) \neq \mathbf{0}$. Здесь $\mathbf{h} = \mathbf{h}(\mathbf{x}) \in \mathcal{A}^m$ есть $m \times 1$ -вектор, $m \leq n$. Вектор-функция $\mathbf{h}(\mathbf{x})$ принимает в $\mathcal{A}^m$ нулевое значение на некоем множестве $\langle \tilde{\mathbf{x}} \rangle \subset \mathcal{A}^n$. Суть задачи состоит в поиске экстремума эволюционной целевой функции на гладком и регулярном *вложенном многообразии* $\langle \tilde{\mathbf{x}} \rangle \subset \mathcal{A}^n$, заданном *ограничительным* способом.

Пусть $y = y(\mathbf{x})$ и $\mathbf{h} = \mathbf{h}(\mathbf{x})$ — скалярная и векторная функции, однозначно определённые и, по крайней мере, дважды непрерывно дифференцируемые на некотором компактном подмножестве в $\mathcal{A}^n$, например, n -мерной закрытой области $\mathcal{T}^n \subset \mathcal{A}^n$. (Причём $\langle \tilde{\mathbf{x}} \rangle \subset \mathcal{T}^n$.) Особенно подчеркнём здесь то, что при дифференцировании по $\tilde{\mathbf{x}}$ последняя понимается обычным образом — как *свободная* переменная $\mathbf{x}$, изменяющаяся от своего точечного значения, но на $\langle \tilde{\mathbf{x}} \rangle$, во всевозможных направлениях в координатном пространстве $\mathcal{A}^n$! Соответственно $d\mathbf{x}$ и $d\tilde{\mathbf{x}}$ — *свободный* и *условный* дифференциалы векторной переменной. Дополнительно примем, что на закрытой области $\mathcal{T}^n$ выполняется требование: $\text{rang}\{\mathbf{dh} / d\tilde{\mathbf{x}}\} = m < n$, где $\mathbf{dh} / d\tilde{\mathbf{x}}$ есть $m \times n$ -матрица Якоби, или 1-я тензор-производная функции $\mathbf{h}(\mathbf{x})$ по $\tilde{\mathbf{x}}$ в окрестности q -поверхности $\langle \tilde{\mathbf{x}} \rangle$, где $q = n - m$.

Такое, по сути, *обратное* функциональное отображение ранга m обеспечивает задание вложенной в $\mathcal{A}^n$ некоторой геометрической поверхности размерности $q = n - m$ ограничительным способом.

$$\frac{d\mathbf{h}}{d\tilde{\mathbf{x}}} = \begin{pmatrix} \frac{\partial h_1}{\partial \tilde{x}_1} & \frac{\partial h_1}{\partial \tilde{x}_2} & \dots & \frac{\partial h_1}{\partial \tilde{x}_n} \\ \frac{\partial h_2}{\partial \tilde{x}_1} & \frac{\partial h_2}{\partial \tilde{x}_2} & \dots & \frac{\partial h_2}{\partial \tilde{x}_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial h_m}{\partial \tilde{x}_1} & \frac{\partial h_m}{\partial \tilde{x}_2} & \dots & \frac{\partial h_m}{\partial \tilde{x}_n} \end{pmatrix} = \left(\frac{\partial h_i}{\partial \tilde{x}_j} \right), \quad (i = \overline{1, m}; j = \overline{1, n}).$$

Дифференцируемость $\mathbf{h}(\tilde{\mathbf{x}})$ k раз обеспечивает функции обратного отображения и поверхности гладкость порядка k ; конечность её тензор-производных до порядка k обеспечивает им же регулярность порядка k .

При указанных допущениях уравнение связи $\mathbf{h}(\tilde{\mathbf{x}}) = \mathbf{0}$ задаёт в $\mathcal{A}^n$ на $\mathcal{T}^n \subset \mathcal{A}^n$ некоторую гладкую регулярную порядка не менее двух q -мерную поверхность $\langle \tilde{\mathbf{x}} \rangle$, где $q = n - m$. Взяв только линейную часть разложения вектор-функции $\mathbf{h}(\mathbf{x})$ по формуле Тейлора в окрестности точки $\tilde{\mathbf{c}} \in \langle \tilde{\mathbf{x}} \rangle$, получаем уравнение для касательной q -плоскости:

$$\mathbf{h}(l) = \frac{d\mathbf{h}}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{c}}) d\tilde{\mathbf{x}} = \frac{d\mathbf{h}}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{c}}) \cdot (l - \tilde{\mathbf{c}}) = \mathbf{0} \rightarrow \langle l \rangle \equiv \tilde{\mathbf{c}} + \left\langle \ker \frac{d\mathbf{h}}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{c}}) \right\rangle. \quad (73)$$

Линеаризация размерности $q = n - m$ для уравнения $\mathbf{h}(\tilde{\mathbf{x}}) = \mathbf{0}$ в точке $\tilde{\mathbf{c}} \in \langle \tilde{\mathbf{x}} \rangle$ определяется здесь линейной частью $\langle \tilde{\mathbf{x}} \rangle$, а геометрически — *касательной q -плоскостью $\langle l \rangle$ к q -поверхности $\langle \tilde{\mathbf{x}} \rangle$* .

Лемма 2 (о линеаризации). Линеаризация размерности $q = n - m$ для уравнения $\mathbf{h}(\tilde{\mathbf{x}}) = \mathbf{0}$, где $\mathbf{h}(\mathbf{x})$ — $m \times 1$ -вектор-функция, имеет место в точке $\tilde{\mathbf{c}} \in \langle \tilde{\mathbf{x}} \rangle$ тогда и только тогда, когда $n \times m$ -матрица Якоби $d\mathbf{h}/d\tilde{\mathbf{x}}$ в ней существует и на некоторой её окрестности имеет постоянный ранг $m < n$, т. е. состоит из m линейно независимых векторов-строк, или $\langle \ker d\mathbf{h}/d\tilde{\mathbf{x}} \rangle_L = \mathbf{0}$. (Доказательство этой леммы вполне тривиально.)

Далее из (73) на касательной q -плоскости $\langle l \rangle$ к q -поверхности $\langle \tilde{\mathbf{x}} \rangle$ в какой-либо точке $\tilde{\mathbf{c}} \in \langle \tilde{\mathbf{x}} \rangle$ находится 1-й условный дифференциал $d\tilde{\mathbf{x}} \equiv dl$ как аффинная проекция свободного дифференциала $d\mathbf{x}$ в точке $\tilde{\mathbf{c}}$ на ядро матрицы Якоби (т. е. на $\langle l \rangle$) *параллельно* её образу (справа):

$$\frac{d\mathbf{h}}{d\tilde{\mathbf{x}}} d\tilde{\mathbf{x}} = \mathbf{0} \leftrightarrow d\tilde{\mathbf{x}} = \left[\left(\frac{d\mathbf{h}}{d\tilde{\mathbf{x}}} \right)^{\rightarrow} \cdot \frac{d\mathbf{h}}{d\tilde{\mathbf{x}}} \right] d\mathbf{x}. \quad (74)$$

Обратим внимание на то, что в формуле фигурирует симметричный характеристический проектор иного типа [27, с. 50], нежели проектор из § 2.1. *Фактически они оба геометрически тождественны!* Данный проектор осуществляет аффинное проецирование в $\mathcal{A}^n$ на $\langle \ker \mathbf{dh} / \mathbf{d}\tilde{\mathbf{x}} \rangle$ параллельно $\langle \text{im } \mathbf{dh} / \mathbf{d}\tilde{\mathbf{x}} \rangle_{\mathbf{R}}$. Этот симметричный проектор вычисляется теоретически однозначно через сингулярную $n \times n$ -матрицу *внешней гомомultiпликации* матрицы Якоби в (74). С его использованием на касательной q -плоскости $\langle \mathcal{L} \rangle$ в точке $\tilde{\mathbf{c}}$ также геометрически наглядным образом (проективно) определяются все условные дифференциальные характеристики 1-го порядка и формулируется уравнение для условной стационарности в задачах данного типа.

Аналогично (65) и (66) определяются условный 1-й дифференциал и *условный градиент* целевой функции $y(\tilde{\mathbf{x}})$ — её дифференциальные характеристики 1-го порядка:

$$\left[dy \right] = \frac{dy}{d\tilde{\mathbf{x}}} d\tilde{\mathbf{x}} = \frac{dy}{d\tilde{\mathbf{x}}} \cdot \left[\left(\frac{\mathbf{dh}}{d\tilde{\mathbf{x}}} \right)' \cdot \frac{\mathbf{dh}}{d\tilde{\mathbf{x}}} \right] d\mathbf{x}, \quad (75)$$

$$\left[\frac{dy}{d\mathbf{x}} \right] = \frac{dy}{d\tilde{\mathbf{x}}} \cdot \frac{d\tilde{\mathbf{x}}}{d\mathbf{x}} = \frac{dy}{d\tilde{\mathbf{x}}} \cdot \left[\left(\frac{\mathbf{dh}}{d\tilde{\mathbf{x}}} \right)' \cdot \frac{\mathbf{dh}}{d\tilde{\mathbf{x}}} \right]. \quad (76)$$

Условный градиент геометрически трактуется и вычисляется здесь как аффинная в $\mathcal{A}^n$ проекция градиента на $\langle \ker \mathbf{dh} / \mathbf{d}\tilde{\mathbf{x}} \rangle$ (т. е. на касательную q -плоскость $\langle \mathcal{L} \rangle$ к q -поверхности $\langle \tilde{\mathbf{x}} \rangle$) параллельно $\langle \text{im } \mathbf{dh} / \mathbf{d}\tilde{\mathbf{x}} \rangle_{\mathbf{R}}$. Он тождествен условному градиенту (66) из § 2.1.

Отсюда вытекает формулируемая ниже *проективная теорема об условной стационарности* для задачи на условный экстремум данного типа, или условный аналог теоремы Эйлера—Ферма (43), как и в § 2.1.

Теорема 2 (проективная). Для того чтобы в точке $\tilde{\mathbf{s}} \in \langle \tilde{\mathbf{x}} \rangle$ имела место условная стационарность целевой функции $y(\tilde{\mathbf{x}})$, необходимо и достаточно, чтобы её условный градиент в данной точке равнялся нулю, а сам градиент в ней был ненулевым:

$$\left[\frac{dy}{d\mathbf{x}}(\tilde{\mathbf{x}}) \right] = \frac{dy}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{x}}) \cdot \left[\left(\frac{\mathbf{dh}}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{x}}) \right)' \cdot \frac{\mathbf{dh}}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{x}}) \right] = \mathbf{0}, \quad \leftrightarrow \quad \tilde{\mathbf{x}} = \tilde{\mathbf{s}} \quad \left(\frac{dy}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \neq \mathbf{0} \right). \quad (77)$$

$\mathbf{h}(\tilde{\mathbf{x}}) = \mathbf{0}$;

Они же суть необходимые требования для реализации в $\tilde{\mathbf{s}} \in \langle \tilde{\mathbf{x}} \rangle$ *условного экстремума* целевой функции $y(\mathbf{x})$. Это есть условный аналог леммы Эйлера — Ферма (43).

Отдельный весьма важный частный случай отвечает постоянной $m \times n$ -матрице Якоби, а именно прямоугольной матрице A размера $m \times n$ и ранга $m < n$. Например, этот случай реализуется на вложенной q -плоскости $\langle \tilde{\mathbf{x}} \rangle : A\tilde{\mathbf{x}} - \mathbf{a} = \mathbf{0}$, где $\mathbf{a} \in \langle \text{im } A \rangle$. Тогда $d\mathbf{h} / d\tilde{\mathbf{x}} = A = \text{Const}$ есть матрица размера $m \times n$ и ранга $m < n$. Причём имеем соответствующую упрощённую формулу для условного градиента, а также первое в (77) требование для условной стационарности:

$$\left[\frac{dy}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \right] = \frac{dy}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \cdot \overrightarrow{A' A} = \mathbf{0} \leftrightarrow \left(\frac{dy}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \right)' \in \langle \text{im } A' \rangle \leftrightarrow \frac{dy}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \in \langle \text{im } A \rangle_{\mathbb{R}}.$$

При *условной* стационарности функции $(dy / d\tilde{\mathbf{x}}(\tilde{\mathbf{s}}) \neq \mathbf{0})$ внешняя гомомультпликация матрицы Якоби (используемая в проекторе) есть обязательно сингулярная $n \times n$ -матрица — либо по причине того, что $m < n$; либо по причине того, что m строк матрицы Якоби линейно зависимы между собой; либо по обоим причинам, вместе взятым. При этом именно в точке условной стационарности $\tilde{\mathbf{s}} \in \langle \tilde{\mathbf{x}} \rangle$ обнуление 1-го дифференциала целевой функции происходит от того, что ненулевой градиент в ней всегда коллинеарен некоему правому несингулярному собственному вектору матрицы Якоби $d\mathbf{h}/d\tilde{\mathbf{x}}$. Поэтому и вышеуказанная аффинная проекция градиента на её ядро *параллельно* её образу (справа) по своей величине здесь обязательно нулевая. (В евклидовом метрическом пространстве $\mathcal{E}^n$ с декартовыми координатами этот же факт интерпретируется геометрическим образом как ортогональность градиента касательной q -плоскости, т. е. аналогично тому, что имело место в предыдущем параграфе.)

Для обоих типов характеристических симметричных проекторов, производимых матрицей Якоби, имеют место формулы взаимосвязи:

$$\left. \begin{aligned} \left[\left(\frac{d\mathbf{h}}{d\tilde{\mathbf{x}}} \right)' \cdot \frac{d\mathbf{h}}{d\tilde{\mathbf{x}}} \right] + \left[\left(\frac{d\mathbf{h}}{d\tilde{\mathbf{x}}} \right)' \cdot \frac{d\mathbf{h}}{d\tilde{\mathbf{x}}} \right] &= I_{n \times n}, \\ \left[\left(\frac{d\mathbf{h}}{d\tilde{\mathbf{x}}} \right)' \cdot \frac{d\mathbf{h}}{d\tilde{\mathbf{x}}} \right] \cdot \left[\left(\frac{d\mathbf{h}}{d\tilde{\mathbf{x}}} \right)' \cdot \frac{d\mathbf{h}}{d\tilde{\mathbf{x}}} \right] &= \left[\left(\frac{d\mathbf{h}}{d\tilde{\mathbf{x}}} \right)' \cdot \frac{d\mathbf{h}}{d\tilde{\mathbf{x}}} \right] \cdot \left[\left(\frac{d\mathbf{h}}{d\tilde{\mathbf{x}}} \right)' \cdot \frac{d\mathbf{h}}{d\tilde{\mathbf{x}}} \right] = Z_{n \times n}. \end{aligned} \right\} \quad (78)$$

Здесь первый проектор проецирует на ядро параллельно образу (справа), а второй проектор проецирует на образ (справа) параллельно ядру. Причём образ (справа) и ядро матрицы Якоби всегда составляют прямую аффинную сумму в $\mathcal{A}^n$ (ортогональную в $\mathcal{E}^n$). В силу этого данные две аффинные проекции для неё всегда существуют! Заметим, что в предыдущем параграфе для проекторов обоих типов выполнялись аналогичные (78) соотношения в силу того, что образ и ядро (слева) матрицы Якоби для функции прямого отображения $\mathbf{x}(\mathbf{u})$ также всегда образуют прямую аффинную сумму в $\mathcal{A}^n$ (ортогональную в $\mathcal{E}^n$).

В рассматриваемом случае ($1 \leq \text{rang} \{d\mathbf{h} / d\tilde{\mathbf{x}}\} = m < n$) проектор на образ матрицы Якоби (справа) параллельно её ядру выражается весьма просто через единственную *правую обратную матрицу* $(d\mathbf{h} / d\tilde{\mathbf{x}})_R^{-1}$, или тождественную *квазиобратную матрицу* Мура — Пенроуза $(d\mathbf{h} / d\tilde{\mathbf{x}})^+$, единственную тут *левую квазиобратную матрицу* вообще [27, с. 53]:

$$\left(\frac{d\mathbf{h}}{d\tilde{\mathbf{x}}}\right)_R^{-1} = \left(\frac{d\mathbf{h}}{d\tilde{\mathbf{x}}}\right)^+ = \left(\frac{d\mathbf{h}}{d\tilde{\mathbf{x}}}\right)' \cdot \left[\frac{d\mathbf{h}}{d\tilde{\mathbf{x}}} \cdot \left(\frac{d\mathbf{h}}{d\tilde{\mathbf{x}}}\right)'\right]^{-1}, \quad (79)$$

$$\frac{d\mathbf{h}}{d\tilde{\mathbf{x}}} \cdot \left(\frac{d\mathbf{h}}{d\tilde{\mathbf{x}}}\right)_R^{-1} = I_{m \times m}, \quad \left(\frac{d\mathbf{h}}{d\tilde{\mathbf{x}}}\right)_R^{-1} \cdot \frac{d\mathbf{h}}{d\tilde{\mathbf{x}}} = \left[\left(\frac{d\mathbf{h}}{d\tilde{\mathbf{x}}}\right)' \cdot \frac{d\mathbf{h}}{d\tilde{\mathbf{x}}}\right]^{-1}, \quad (80, 81)$$

С учётом формул (78)—(81), уравнение для условного градиента в системе (77) преобразуется следующим образом — *именно для точки $\tilde{\mathbf{s}}$* :

$$\begin{aligned} \left[\frac{dy}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}})\right] &= \frac{dy}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) - \frac{dy}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \cdot \left[\left(\frac{d\mathbf{h}}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}})\right)' \cdot \frac{d\mathbf{h}}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}})\right]^{-1} \\ &= \frac{dy}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) - \left[\frac{dy}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \cdot \left(\frac{d\mathbf{h}}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}})\right)_R^{-1}\right] \cdot \frac{d\mathbf{h}}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) = \\ &= \frac{dy}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) - \lambda \cdot \frac{d\mathbf{h}}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) = \frac{d(y - \lambda \cdot \mathbf{h})}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) = \mathbf{0}, \end{aligned} \quad (82)$$

где

$$\lambda = \frac{dy}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \cdot \left(\frac{d\mathbf{h}}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}})\right)_R^{-1} \quad (83)$$

есть *явный* $1 \times m$ -вектор множителей Лагранжа, т. е. $\lambda = (\lambda_1, \dots, \lambda_m)$.

В свою очередь, этот *вектор-строка* (как и градиент!) является при $\text{rang } d\mathbf{h} / d\tilde{\mathbf{x}} = m$ *единственным* решением линейного уравнения:

$$\lambda \cdot \frac{d\mathbf{h}}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) = \frac{dy}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \neq \mathbf{0}.$$

Обратим внимание здесь на то, что левая квазиобратная матрица и она также правая обратная матрица для матрицы Якоби, в силу своей единственности, *не требует ввода метрики* в $\mathcal{A}^n$. Все производимые новые аффинные характеристики при этом определяются однозначно. Оставаясь в аффинном пространстве, приходим к фундаментальному следствию и к классическому методу множителей Лагранжа [60].

А именно: преобразование (82) сводит решение задачи на *условную стационарность* функции $y(\tilde{\mathbf{x}})$, согласно системе (77), к решению задачи на *безусловную стационарность* функции Лагранжа (84) от 2-х независимых друг от друга векторных переменных (**метод Лагранжа**)

$$\mathcal{L} = \mathcal{L}(\tilde{\mathbf{x}}, \lambda) = y(\tilde{\mathbf{x}}) - \lambda \cdot \mathbf{h}(\tilde{\mathbf{x}}), \quad (84)$$

причём с увеличением размерности координатного пространства и задачи с n до $(n + m)$. Из условия стационарности функции $\mathcal{L}(\tilde{\mathbf{x}}, \lambda)$ на суммарном аффинном координатном пространстве $\mathcal{A}^{n+m}$ получаем в итоге классическую **теорему Лагранжа** о необходимых и достаточных требованиях для *условной стационарности* $y(\tilde{\mathbf{x}})$ в точке $\tilde{\mathbf{s}}$:

$$\left. \begin{aligned} \frac{\partial \mathcal{L}(\tilde{\mathbf{x}}, \lambda)}{\partial \tilde{\mathbf{x}}} &= \frac{dy}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{x}}) - \lambda \cdot \frac{d\mathbf{h}}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{x}}) = \mathbf{0}, \\ \frac{\partial \mathcal{L}(\tilde{\mathbf{x}}, \lambda)}{\partial \lambda} &= -\mathbf{h}(\tilde{\mathbf{x}}) = \mathbf{0}; \end{aligned} \right\} \leftrightarrow \tilde{\mathbf{x}} = \tilde{\mathbf{s}} \left(\frac{dy}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \neq \mathbf{0} \right). \quad (85)$$

В точке условной стационарности целевой функции *с точностью до 1-го полного дифференциала* функции Лагранжа тогда имеем: $\mathcal{L} = \mathcal{L}(\tilde{\mathbf{x}}, \lambda) = y(\tilde{\mathbf{x}}) - \lambda \cdot \mathbf{h}(\tilde{\mathbf{x}}) = \text{const}$, или $d\mathcal{L}(\tilde{\mathbf{x}}, \lambda) = 0$.

В части вычисления условной стационарности в задачах *данного типа* метод Лагранжа, как видим, полностью тождествен по результату вышеизложенному *проективному методу*, основанному на вполне наглядных и естественных геометрических соображениях. В свою очередь, проективный метод через преобразование (82) здесь же вполне естественным путём приводит к классическому методу Лагранжа.

Более того, докажем полную тождественность обоих методов при идентификации характера условной стационарности (в том числе и условного экстремума) целевой функции в точке $\tilde{\mathbf{s}}$ — решении как системы (77), так и системы (85).

С целью вычисления условных дифференциальных характеристик 2-го порядка (как квадратичных проекций на касательную q -поверхность 2-го порядка $\langle q \rangle$ к q -поверхности $\langle \tilde{\mathbf{x}} \rangle$ в точке $\tilde{\mathbf{s}}$) разложим функции $y(\tilde{\mathbf{x}})$ и $\mathbf{h}(\tilde{\mathbf{x}})$ совместно в окрестности точки $\tilde{\mathbf{s}}$ по формулам Тейлора, учтя только слагаемые 2-го порядка:

$$\left. \begin{aligned} \left[d^2 y(\tilde{\mathbf{x}} = \tilde{\mathbf{s}}) \right] &= d\tilde{\mathbf{x}}' \frac{d^2 y}{d\tilde{\mathbf{x}}' d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) d\tilde{\mathbf{x}} + \frac{dy}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) d^2 \tilde{\mathbf{x}}, \\ &\quad \begin{matrix} 1 \times 1 & 1 \times n & n \times n & n \times 1 & 1 \times n & n \times 1 \end{matrix} \\ \left[d^2 \mathbf{h}(\tilde{\mathbf{x}} = \tilde{\mathbf{s}}) \right] &= d\tilde{\mathbf{x}}' \left(\frac{d^2 \mathbf{h}}{d\tilde{\mathbf{x}}' d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \right)^{(1 \leftrightarrow 2)} d\tilde{\mathbf{x}} + \frac{d\mathbf{h}}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) d^2 \tilde{\mathbf{x}} = \mathbf{0}, \\ &\quad \begin{matrix} 1 \times m & 1 \times n & n \times m \times n & n \times 1 & m \times n & n \times 1 \end{matrix} \end{aligned} \right\} \quad (86)$$

где $d^2 \tilde{\mathbf{x}}$ — условный 2-й дифференциал $\mathbf{x}$, $d^2 \mathbf{h} / d\tilde{\mathbf{x}}' d\tilde{\mathbf{x}}$ — 2-я тензор-производная функции $\mathbf{h}(\mathbf{x})$.

Последняя в своей *исходной* форме представляет собой трёхмерную $m \times n \times n$ -матрицу, состоящую именно из вторых частных производных $\frac{\partial^2 h_i}{\partial \tilde{x}_k \partial \tilde{x}_j}$ ($i = \overline{1, m}; j, k = \overline{1, n}$). Как трёхмерная матрица она имеет три типа плоских сечений:

- 1) фронтальные $m \times n$ -сечения $\left(\frac{\partial^2 h_i}{\partial \tilde{x}_k \partial \tilde{x}_j} \right)$, ($i = \overline{1, m}; j = \overline{1, n}; k = \text{const}$);
- 2) боковые $n \times m$ -сечения $\left(\frac{\partial^2 h_i}{\partial \tilde{x}_k \partial \tilde{x}_j} \right)$, ($i = \overline{1, m}; k = \overline{1, n}; j = \text{const}$);
- 3) горизонтальные $n \times n$ -сечения $\left(\frac{\partial^2 h_i}{\partial \tilde{x}_k \partial \tilde{x}_j} \right)$, ($j, k = \overline{1, n}; i = \text{const}$).

(Последние из них обязательно симметричные.)

По сути, это означает возможность существования трёх форм трёхмерной матрицы: одной исходной и двух транспонированных. Из-за перестановочности частных *скалярных* дифференциалов $\partial \tilde{x}_j$ и $\partial \tilde{x}_k$ первые две формы матрицы идентичны друг другу с точностью до двумерного транспонирования, в том числе по своим сечениям при $j = k$.

Следовательно, для вычисления, например, первого слагаемого во 2-м уравнении в системе (86) дифференциалы аргумента слева и справа умножаются именно на горизонтальные $n \times n$ -сечения при i от 1 до m , производя при этом вектор $d^2\mathbf{h}$, и т. д.

Далее из системы (86) исключаем общее неизвестное $d^2\tilde{\mathbf{x}}$, выражая его из 2-го уравнения и подставляя в 1-е уравнение. Здесь следует учесть, что, согласно (77) и (78), аффинная проекция градиента целевой функции в точке $\tilde{\mathbf{s}}$ на образ матрицы Якоби (справа) параллельно её ядру тождественна самому этому градиенту. Тогда дополнительно, с учётом формулы (80), имеем:

$$\frac{dy}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) = \frac{dy}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \cdot \left[\left(\frac{d\mathbf{h}}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \right)' \cdot \frac{d\mathbf{h}}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \right]^{-1} = \frac{dy}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \cdot \left(\frac{d\mathbf{h}}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \right)_R^{-1} \cdot \frac{d\mathbf{h}}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}). \quad (87)$$

Подставив это выражение вместо градиента в 1-е уравнение в (86) и сделав сюда же подстановку второго слагаемого из 2-го уравнения в (86), предварительно умноженного слева на правую обратную матрицу, в итоге получаем нижеуказанную формулу (88) для условного 2-го дифференциала целевой функции, отсчитываемого от её значения в точке $\tilde{\mathbf{s}}$:

$$\left[d^2y(\tilde{\mathbf{x}} = \tilde{\mathbf{s}}) \right] = \underbrace{d\tilde{\mathbf{x}}'}_{1 \times n} \left\{ \underbrace{\frac{d^2y}{d\tilde{\mathbf{x}}' d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}})}_{n \times n} - \underbrace{\left[\frac{dy}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \cdot \left(\frac{d\mathbf{h}}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \right)_R^{-1} \right]}_{1 \times m} \cdot \underbrace{\left(\frac{d^2\mathbf{h}}{d\tilde{\mathbf{x}}' d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \right)'_{(1 \leftrightarrow 2)}}_{n \times m \times n} \right\} \underbrace{d\tilde{\mathbf{x}}}_{n \times 1}.$$

Здесь m элементов специального $1 \times m$ -вектора в квадратных скобках поэтапно умножаются на m симметричных $n \times n$ -сечений трёхмерной матрицы, а затем всё это суммируется в двумерную симметричную $n \times n$ -матрицу в фигурных скобках. Причём этот вектор имеет размерность функции $\mathbf{h}$ и, по сути, образован скалярными множителями Лагранжа. В итоге выражение (88) в целом представляет тут собой вырожденную квадратичную форму, ввиду хотя бы вырождённости дифференциала $d\tilde{\mathbf{x}}$. Порядок перемножения сомножителей имеет вполне естественный вид. Далее отсюда, с учётом формулы (74), выражаем искомую сингулярную *условную матрицу Гессе* в точке $\tilde{\mathbf{s}}$:

$$\begin{aligned}
 \left[\frac{d^2 y}{d\mathbf{x}' d\mathbf{x}} (\tilde{\mathbf{s}}) \right]_{n \times n} &= \left[\left(\frac{d\mathbf{h}}{d\tilde{\mathbf{x}}} (\tilde{\mathbf{s}}) \right)' \cdot \frac{d\mathbf{h}}{d\tilde{\mathbf{x}}} (\tilde{\mathbf{s}}) \right]_{n \times n} \cdot \left\{ \frac{d^2 y}{d\tilde{\mathbf{x}}' d\tilde{\mathbf{x}}} (\tilde{\mathbf{s}}) - \right. \\
 &\quad \left. - \left[\frac{dy}{d\tilde{\mathbf{x}}} (\tilde{\mathbf{s}}) \cdot \left(\frac{d\mathbf{h}}{d\tilde{\mathbf{x}}} (\tilde{\mathbf{s}}) \right)_R^{-1} \right]_{1 \times m} \cdot \left(\frac{d^2 \mathbf{h}}{d\tilde{\mathbf{x}}' d\tilde{\mathbf{x}}} \right)'^{(1 \leftrightarrow 2)}_{n \times m \times n} \right\} \cdot \left[\left(\frac{d\mathbf{h}}{d\tilde{\mathbf{x}}} (\tilde{\mathbf{s}}) \right)' \cdot \frac{d\mathbf{h}}{d\tilde{\mathbf{x}}} (\tilde{\mathbf{s}}) \right]_{n \times n}. \quad (89)
 \end{aligned}$$

Обратим внимание на то, что в мнемонически близких формулах (71), (72) и (89) в 2-х различных по постановке и решению типах задач на условный экстремум реализуется эта общая схема перемножения с трёхмерной матрицей. Такая мнемоническая аналогия отображает некую функциональную взаимосвязь 2-х проективных методов анализа и решения данных задач. *Причём размерности специальных векторов в (72) и (89) в сумме составляют n .* Кроме того, формулы (88) и (89) мнемонически аналогичны скалярным прототипам (29) и (30) из § 1.6.

В простейшем частном случае — на q -плоскости матрица Якоби размера $m \times n$ постоянная и имеет ранг $m < n$: $d\mathbf{h} / d\mathbf{x} = A = \text{Const}$. При этом условная матрица Гессе выражается весьма простой формулой:

$$\left[\frac{d^2 y}{d\mathbf{x}' d\mathbf{x}} \right] = A' A \cdot \frac{d^2 y}{d\tilde{\mathbf{x}}' d\tilde{\mathbf{x}}} \cdot A' A.$$

Причём, если в таком случае безусловная матрица Гессе $d^2 y / d\tilde{\mathbf{x}}' d\tilde{\mathbf{x}}$ знакоопределённая (положительно или отрицательно), то она сама по себе определяет характер условного экстремума в точке $\tilde{\mathbf{s}}$. Это объясняется тем, что выпуклая или вогнутая на пространстве $\mathcal{A}^n$ целевая функция $y(\mathbf{x})$ остаётся в том же качестве и на q -плоскости $\langle \tilde{\mathbf{x}} \rangle$.

Выраженные (88) и (89) в $\tilde{\mathbf{s}}$ условные на касательной q -поверхности в ней 2-го порядка $\langle \mathbf{g} \rangle$ 2-й дифференциал и матрица Гессе для исходной функции $y(\tilde{\mathbf{x}})$ тождественны условным на касательной q -плоскости в ней $\langle \mathcal{L} \rangle$ частным 2-му дифференциалу и матрице Гессе для $\mathcal{L}(\tilde{\mathbf{x}}, \lambda)$! Действительно, выражение в квадратных скобках в (88) и (89) в точке $\tilde{\mathbf{s}}$ в точности, согласно (83), идентично множителю Лагранжа — некоей векторной константе при заданных первоначальных предположениях. Как константу множитель Лагранжа вводим под знак дифференциала в качестве векторного множителя для трёхмерной матрицы, что делает их произведение в скобках двумерной матрицей. В итоге имеем идентичность и 2-ых дифференциалов, и 2-ых тензор-производных:

$$\begin{aligned} \left[d^2 y(\tilde{\mathbf{x}} = \tilde{\mathbf{s}}) \right] &= d\tilde{\mathbf{x}}' \left\{ \frac{\partial^2 y(\tilde{\mathbf{x}})}{\partial \tilde{\mathbf{x}}' \partial \tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) - \left[\frac{\partial^2 (\boldsymbol{\lambda} \cdot \mathbf{h}(\tilde{\mathbf{x}}))}{\partial \tilde{\mathbf{x}}' \partial \tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \right] \right\} d\tilde{\mathbf{x}} = \\ &= d\tilde{\mathbf{x}}' \frac{\partial^2 [y - \boldsymbol{\lambda} \cdot \mathbf{h}]}{\partial \tilde{\mathbf{x}}' \partial \tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) d\tilde{\mathbf{x}} = d\tilde{\mathbf{x}}' \frac{\partial^2 \mathcal{L}}{\partial \tilde{\mathbf{x}}' \partial \tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) d\tilde{\mathbf{x}} = \left[\partial_{\tilde{\mathbf{x}}\tilde{\mathbf{x}}}^2 \mathcal{L}(\tilde{\mathbf{x}} = \tilde{\mathbf{s}}) \right] \langle \mathcal{I} \rangle, \end{aligned} \quad (90)$$

$$\begin{aligned} \left[\frac{d^2 y}{d\mathbf{x}' d\mathbf{x}}(\tilde{\mathbf{s}}) \right] &= \left[\left(\frac{d\mathbf{h}}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \right)' \cdot \frac{d\mathbf{h}}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \right] \cdot \frac{\partial^2 \mathcal{L}}{\partial \tilde{\mathbf{x}}' \partial \tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \cdot \left[\left(\frac{d\mathbf{h}}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \right)' \cdot \frac{d\mathbf{h}}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \right] = \\ &= \left[\frac{\partial^2 \mathcal{L}}{\partial \mathbf{x}' \partial \mathbf{x}}(\tilde{\mathbf{s}}) \right] \langle \mathcal{I} \rangle. \end{aligned} \quad (91)$$

Формулы отражают естественную взаимосвязь метода множителей Лагранжа и метода условных тензор-производных (т. е. проективной версии) в решении задач данного типа на 2-м этапе. Итак, *во-первых*, условная стационарность $y(\tilde{\mathbf{x}})$ и стационарность $\mathcal{L}(\tilde{\mathbf{x}}, \boldsymbol{\lambda})$ в части $\tilde{\mathbf{x}}$ реализуются всегда в одной и той же точке $\tilde{\mathbf{s}}$ (или множестве точек $\langle \tilde{\mathbf{s}} \rangle$); *во-вторых*, матрица Гессе функций $y(\tilde{\mathbf{x}})$ и $\mathcal{L}(\tilde{\mathbf{x}}, \boldsymbol{\lambda})$ по $\tilde{\mathbf{x}}$, хотя вообще и различны, но зато совпадают именно в точке условной стационарности $\tilde{\mathbf{s}}$! Это, по сути, является теоретической основой классического метода множителей Лагранжа для решения задач на условный экстремум данного типа, по крайней мере, с его уровнем $p = 2$.

Условная матрица Гессе в точке $\tilde{\mathbf{s}}$ по обоим вариантам задачи (71) и (89), (91) имеет не менее, чем m нулевых собственных значений. При этом условная стационарность обязательно строгая, если их количество в точности равно параметру m . Тогда все эти m нулевых собственных значений попросту не принимаются здесь во внимание. Остаточное количество $q = n - m$ условных собственных значений матрицы Гессе в точке $\tilde{\mathbf{s}}$ теоретически обычным образом определяют (как и в § 1.9) характер условной стационарности целевой функции $y(\tilde{\mathbf{x}})$.

Если симметричная матрица Гессе для функции Лагранжа в (91) *знакоопределённая*, то она полностью задаёт характер именно *строгого* условного экстремума. Объясняется этот факт тем, что чисто выпуклая или чисто вогнутая скалярная функция $y(\mathbf{x})$, заданная на $\mathcal{A}^n$, остаётся таковой и в пределах q -плоскости $\langle \mathcal{I} \rangle \subset \mathcal{A}^n$. Ввиду особой важности этой внутренней части условной матрицы Гессе в целом, далее она будет называться *полуусловной* матрицей Гессе. В явном виде она выражается *тождественным образом* в фигурных скобках в (71), (72) и в (89)!

Разумеется, и тут также (как ранее в §§ 1.9, 2.1) возникает проблема упрощения процедуры оценки знаков условных собственных значений. В важном случае симметричная несингулярная полуусловная матрица Гессе S в точке $\tilde{s}$ *знаконеопределённая* ($\det S \neq 0$, $\text{rang } [G] = q = n - m$), что тоже имеет место при *строгой* условной стационарности функции. Необходимо выяснить характер её знакоопределённости. Для этого условную матрицу Гессе приводим неким ортогональным модальным преобразованием R ($\tilde{E}_Z = R \cdot \tilde{E}_X$) к *нуль-клеточной форме*:

$$R' \cdot \begin{pmatrix} \overrightarrow{A'A} \cdot S \cdot \overrightarrow{A'A} \\ \overleftarrow{A'A} \cdot S \cdot \overleftarrow{A'A} \end{pmatrix} \cdot R = \begin{array}{|c|c|} \hline S_{q \times q} & Z_{q \times m} \\ \hline Z_{m \times q} & Z_{m \times m} \\ \hline \end{array},$$

где $\det S \neq 0$, $\det S_{q \times q} \neq 0$; Z — нулевые матрицы соответствующих размеров. Например, такое модальное преобразование R вычисляется по способу, изложенному в монографии [27, с. 49]. При этом используется следующее свойство характеристических симметричных проекторов:

$$\begin{bmatrix} \overrightarrow{A'A} \cdot S \cdot \overrightarrow{A'A} \\ \overleftarrow{A'A} \cdot S \cdot \overleftarrow{A'A} \end{bmatrix} = \overleftarrow{A'A}, \quad \begin{bmatrix} \overrightarrow{A'A} \cdot S \cdot \overrightarrow{A'A} \\ \overleftarrow{A'A} \cdot S \cdot \overleftarrow{A'A} \end{bmatrix} = \overrightarrow{A'A}, \quad \text{где } \det S \neq 0.$$

А именно, проекторы $\overleftarrow{A'A}$ и $\overrightarrow{A'A}$ рангов m и q обязательно содержат базисные диагональные $m \times m$ - и $q \times q$ -миноры. Последние задают две базисные $m \times n$ - и $q \times n$ -субматрицы базисных столбцов. Объединяя их сначала в $n \times n$ аффинную модальную матрицу V , а затем ортогонализуя её, например, по Граму — Шмидту, получаем искомую модальную ортогональную матрицу R . Далее к полученной по вышеуказанной схеме несингулярной угловой симметричной матрице $S_{q \times q}$ применяют критерий Сильвестра, как это было описано ранее в § 1.9.

Аналитическое решение задач данного типа в самом общем случае осуществляется по вышеизложенному методу множителей Лагранжа, а при линейном характере ограничения и более простым способом — методом условных тензор-производных. В общем виде решение задачи заключается в нахождении условного экстремума целевой функции на некотором нелинейном или линейном ядре $\langle \tilde{x} \rangle$ преобразования $\mathbf{h} = \mathbf{h}(\mathbf{x})$. В задачах на экстремум данного типа сложность их анализа и решения увеличивается именно с ростом размерности m векторного параметра $\mathbf{h}$.

Наиболее просто они решаются при $m = 1$, т. е. как задачи на условный экстремум целевой функции на некоторой *гиперповерхности* или *гиперплоскости*, вложенной в $\mathcal{A}^n$. (Сравните всё это с аналогичными выводами в § 2.1.)

Следовательно, вышеизложенные в §§ 2.1 и 2.2 два принципиально различных подхода к постановке, анализу и решению задачи на условный экстремум на аффинном координатном пространстве $\mathcal{A}^n$ совершенно естественным образом дополняют друг друга, особенно при крайних значениях размерности $\langle \tilde{\mathbf{x}} \rangle$! Они отвечают двум основным классическим способам аналитического задания неких q -поверхностей $\langle \tilde{\mathbf{x}} \rangle$, вложенных в пространство $\mathcal{A}^n$ — либо через векторный параметр размерности $q < n$, либо через систему уравнений связи размерности $m = n - q < n$ [17, 23].

Покажем также инвариантность обеих изложенных выше процедур решения задачи на условный экстремум данного типа по отношению к линейным модальным преобразованиям базисов в $\mathcal{A}^n$ и в $\mathcal{A}^m$:

$$\tilde{\mathbf{E}}_z = \mathbf{V}_1 \cdot \tilde{\mathbf{E}}_x \text{ и } \tilde{\mathbf{E}}_t = \mathbf{V}_2 \cdot \tilde{\mathbf{E}}_h; \quad \mathbf{x} = \mathbf{V}_1 \cdot \mathbf{z} \text{ и } \mathbf{h} = \mathbf{V}_2 \cdot \mathbf{t}.$$

При доказательстве этого утверждения характеристический проектор применяется, согласно формулам (78) и (81), в виде

$$\left[\left(\frac{d\mathbf{h}}{d\tilde{\mathbf{x}}} \right)' \cdot \frac{d\mathbf{h}}{d\tilde{\mathbf{x}}} \right] = \mathbf{I}_{n \times n} - \left(\frac{d\mathbf{h}}{d\tilde{\mathbf{x}}} \right)_R^{-1} \cdot \frac{d\mathbf{h}}{d\tilde{\mathbf{x}}}.$$

Далее последовательно имеем:

$$\begin{aligned} d\tilde{\mathbf{x}} &= \left[\left(\frac{d\mathbf{h}}{d\tilde{\mathbf{x}}} \right)' \cdot \frac{d\mathbf{h}}{d\tilde{\mathbf{x}}} \right] d\mathbf{x} = \left[\mathbf{I}_{n \times n} - \left(\frac{d\mathbf{h}}{d\tilde{\mathbf{x}}} \right)_R^{-1} \cdot \frac{d\mathbf{h}}{d\tilde{\mathbf{x}}} \right] d\mathbf{x}, \\ \mathbf{V}_1 d\tilde{\mathbf{z}} &= \left[\mathbf{V}_1 \cdot \mathbf{V}_1^{-1} - \mathbf{V}_1 \left(\frac{d\mathbf{t}}{d\tilde{\mathbf{z}}} \right)_R^{-1} \cdot \mathbf{V}_2^{-1} \cdot \mathbf{V}_2 \frac{d\mathbf{t}}{d\tilde{\mathbf{z}}} \mathbf{V}_1^{-1} \right] \mathbf{V}_1 d\mathbf{z} = \\ &= \mathbf{V}_1 \left[\mathbf{I}_{n \times n} - \left(\frac{d\mathbf{t}}{d\tilde{\mathbf{z}}} \right)_R^{-1} \cdot \frac{d\mathbf{t}}{d\tilde{\mathbf{z}}} \right] d\mathbf{z} = \mathbf{V}_1 \left[\left(\frac{d\mathbf{t}}{d\tilde{\mathbf{z}}} \right)' \cdot \frac{d\mathbf{t}}{d\tilde{\mathbf{z}}} \right] d\mathbf{z} \leftrightarrow \\ &\leftrightarrow d\tilde{\mathbf{z}} = \left[\left(\frac{d\mathbf{t}}{d\tilde{\mathbf{z}}} \right)' \cdot \frac{d\mathbf{t}}{d\tilde{\mathbf{z}}} \right] d\mathbf{z}. \end{aligned}$$

Для метода множителей Лагранжа имеем новое уравнение условной стационарности в векторной форме:

$$\frac{dy}{d(V_1 \tilde{z})} - \lambda \cdot \frac{dh}{d(V_1 \tilde{z})} = \frac{dy}{d\tilde{z}} V_1^{-1} - \lambda \cdot \frac{dh}{d\tilde{z}} V_1^{-1} = \mathbf{0} = \frac{dy}{d\tilde{z}} - \lambda \cdot \frac{dh}{d\tilde{z}}.$$

Затем выводим новый вектор-множитель Лагранжа:

$$\begin{aligned} \lambda &= \frac{dy}{d(V_1 \tilde{z})} \cdot \left[\frac{d(V_2 \mathbf{t})}{d(V_1 \tilde{z})} \right]_{\mathbf{R}}^{-1} = \frac{dy}{d\tilde{z}} V_1^{-1} \cdot V_1 \left(\frac{d\mathbf{t}}{d\tilde{z}} \right)_{\mathbf{R}}^{-1} \cdot V_2^{-1} = \\ &= \frac{dy}{d\tilde{z}} \cdot \left(\frac{d\mathbf{t}}{d\tilde{z}} \right)_{\mathbf{R}}^{-1} \cdot V_2^{-1} = \mathbf{I} \cdot V_2^{-1} \leftrightarrow \mathbf{I} = \lambda \cdot V_2. \end{aligned}$$

Для условной стационарности функции имеем инвариантную форму:

$$\frac{dy}{d\tilde{z}} - \mathbf{I} \cdot V_2^{-1} \cdot V_2 \frac{d\mathbf{t}}{d\tilde{z}} = \mathbf{0} = \frac{dy}{d\tilde{z}} - \mathbf{I} \cdot \frac{d\mathbf{t}}{d\tilde{z}}.$$

Для условной матрицы Гессе от функции Лагранжа аналогично имеем инвариантную форму:

$$\begin{aligned} V_1' \left[\frac{\partial^2 \mathcal{L}}{\partial \mathbf{x}' \partial \mathbf{x}}(\tilde{\mathbf{s}}) \right]_{\langle \mathbf{I} \rangle} \cdot V_1 &= V_1' \left[\frac{\partial^2 \mathcal{L}}{\partial (V_1 \mathbf{z})' \partial (V_1 \mathbf{z})} \right]_{\langle \mathbf{I} \rangle} \cdot V_1 = \left[\frac{\partial^2 \mathcal{L}}{\partial \mathbf{z}' \partial \mathbf{z}} \right]_{\langle \mathbf{I} \rangle} = \\ &= \left[\left(\frac{d\mathbf{t}}{d\tilde{z}} \right)' \cdot \frac{d\mathbf{t}}{d\tilde{z}} \right] \cdot V_1' \left\{ \frac{d^2 y}{d(V_1 \tilde{z})' d(V_1 \tilde{z})} - \frac{\partial^2 (\mathbf{I} \cdot V_2^{-1} \cdot V_2 \cdot \mathbf{t})}{\partial (V_1 \tilde{z})' \partial (V_1 \tilde{z})} \right\} \cdot V_1 \cdot \left[\left(\frac{d\mathbf{t}}{d\tilde{z}} \right)' \cdot \frac{d\mathbf{t}}{d\tilde{z}} \right] \leftrightarrow \\ &\leftrightarrow \left[\frac{\partial^2 \mathcal{L}}{\partial \mathbf{z}' \partial \mathbf{z}} \right] = \left[\left(\frac{d\mathbf{t}}{d\tilde{z}} \right)' \cdot \frac{d\mathbf{t}}{d\tilde{z}} \right] \cdot \left\{ \frac{d^2 y}{d\tilde{z}' d\tilde{z}} - \frac{\partial^2 (\mathbf{I} \cdot \mathbf{t})}{\partial \tilde{\mathbf{z}}' \partial \tilde{\mathbf{z}}} \right\} \cdot \left[\left(\frac{d\mathbf{t}}{d\tilde{z}} \right)' \cdot \frac{d\mathbf{t}}{d\tilde{z}} \right]. \end{aligned}$$

Доказательство той же общей инвариантности для метода условных тензор-производных (т. е. для проективной версии) отличается только тем, что множитель Лагранжа фигурирует в развёрнутой форме.

Доказательство инвариантности обоих методов к наиболее общим — *нелинейным* модальным преобразованиям обеих переменных $\tilde{\mathbf{x}}$ и $\mathbf{h}$ осуществляется аналогичным образом, но формально всё это приводит к анализу условной стационарности исходной целевой функции, с учётом *двухступенчатости* сложной вектор-функции ограничения. Следует иметь в виду, что степени гладкости и регулярности обоих модальных преобразований при анализе такой задачи должны отвечать уровню условной стационарности или экстремума (выше это $p = 2$).

* * *

Пример. Проанализируем с точки зрения условной оптимизации скалярную функцию 2-го порядка от линейно ограниченной переменной

$$q = q(\tilde{\mathbf{x}}) = 1/2 \cdot \tilde{\mathbf{x}}' \mathbf{G} \tilde{\mathbf{x}} + \mathbf{g}_0' \cdot \tilde{\mathbf{x}} + q_0 \quad (\det \mathbf{G} \neq 0),$$

$$\mathbf{A}' \mathbf{A} \tilde{\mathbf{x}} = \mathbf{A}' \mathbf{a} \quad (\text{rang } \mathbf{A} = m), \quad \frac{d\mathbf{q}}{d\mathbf{x}} \neq \mathbf{0}.$$

Поэтапно находим условные градиент и матрицу Гессе по $\tilde{\mathbf{x}}$:

$$\left[\frac{d\mathbf{q}}{d\mathbf{x}} \right]' = \mathbf{A}' \mathbf{A} \cdot (\mathbf{G} \tilde{\mathbf{x}} + \mathbf{g}_0') = \mathbf{A}' \mathbf{A} \cdot \mathbf{G} \tilde{\mathbf{x}} + \mathbf{A}' \mathbf{A} \cdot \mathbf{g}_0',$$

$$\left[\frac{d^2 q}{d\mathbf{x}' d\mathbf{x}} \right] = \mathbf{A}' \mathbf{A} \cdot \mathbf{G} \cdot \mathbf{A}' \mathbf{A} = \text{Const}.$$

Очевидно, что градиентное уравнение здесь всегда имеет решение и причём вырожденное. Из системы (77) далее следует:

$$\left| \begin{array}{c} \mathbf{A}' \mathbf{A} \cdot \mathbf{G} \\ \mathbf{A}' \mathbf{A} \end{array} \right| \cdot \tilde{\mathbf{s}} = \left| \begin{array}{c} \mathbf{A}' \mathbf{A} \cdot \mathbf{g}_0' \\ \mathbf{A}' \mathbf{a} \end{array} \right|,$$

$2n \times n \quad n \times 1 \quad 2n \times 1$

$$\tilde{\mathbf{s}} = \left| \begin{array}{c} \mathbf{A}' \mathbf{A} \cdot \mathbf{G} \\ \mathbf{A}' \mathbf{A} \end{array} \right|_{\mathbf{L}}^{-1} \cdot \left| \begin{array}{c} \mathbf{A}' \mathbf{A} \cdot \mathbf{g}_0' \\ \mathbf{A}' \mathbf{a} \end{array} \right|.$$

Поскольку $(\det \mathbf{G} \neq 0)$, то в точке $\tilde{\mathbf{s}}$ имеем (уровня $p = 2$):

- условный минимум, если $\mathbf{G}$ или $[\mathbf{G}]$ положительно определённая;
- условный максимум, если $\mathbf{G}$ или $[\mathbf{G}]$ отрицательно определённая;
- условную стационарную седловину, если $[\mathbf{G}]$ знаконеопределённая.

В точке условной стационарности целевая функция имеет значение:

$$q(\tilde{\mathbf{s}}) = -1/2 \cdot \tilde{\mathbf{s}}' \cdot \mathbf{G} \cdot \tilde{\mathbf{s}} + \mathbf{g}_0' \cdot \tilde{\mathbf{s}} + q_0.$$

Это решение значительно упрощается (для линейного ограничения), если начало координат перенести на плоскость $\langle \tilde{\mathbf{x}} \rangle \equiv \langle \mathbf{L} \rangle$. Тогда в новых координатах $\mathbf{A} \tilde{\mathbf{x}} = \mathbf{0}$, $\langle \tilde{\mathbf{x}} \rangle \equiv \langle \ker \mathbf{A} \rangle$. Далее имеем:

$$\tilde{\mathbf{s}} = -\mathbf{A}' \mathbf{A} \cdot \mathbf{G}^{-1} \cdot \mathbf{A}' \mathbf{A} \cdot \mathbf{g}_0',$$

$$q(\tilde{\mathbf{s}}) = -1/2 \cdot \mathbf{g}_0' \cdot \mathbf{A}' \mathbf{A} \cdot \mathbf{G}^{-1} \cdot \mathbf{A}' \mathbf{A} \cdot \mathbf{g}_0' + q_0.$$

(Сравните данный результат с результатом в примере из § 2.1.)

§ 2.3. Клеточный условный экстремум

Пусть $m \times 1$ -вектор-функция ограничения имеет блочную структуру:

$$\mathbf{h} = \mathbf{h}(\mathbf{x}) \equiv (\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_d) = [\mathbf{t}_1(\mathbf{u}_1), \mathbf{t}_2(\mathbf{u}_2), \dots, \mathbf{t}_d(\mathbf{u}_d)],$$

где $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_d$ — независимые вектор-аргументы размерности $p_1, p_2, \dots, p_d \geq 1$ ($p_1 + \dots + p_d = n$), составляющие в прямой сумме $\mathbf{x} \in \mathcal{A}^n$, $\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_d$ — независимые вектор-функции ограничения размерности $q_1, q_2, \dots, q_d \geq 1$ ($q_1 + \dots + q_d = m < n$), составляющие в прямой сумме $\mathbf{h} \in \mathcal{A}^m$. В силу блочности структуры, решение задачи на условный экстремум целевой функции $y = y(\mathbf{x}) \equiv y(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_d)$ с ограничением $\mathbf{h}(\mathbf{x}) = \mathbf{0}$ осуществляется тут поклеточно с понижением размерностей соответствующих векторных и матричных характеристик.

Система (77) распадается на d отдельных клеток:

$$\left[\frac{\partial y}{\partial \mathbf{u}_k}(\tilde{\mathbf{s}}) \right] = \frac{\partial y}{\partial \tilde{\mathbf{u}}_k}(\tilde{\mathbf{s}}_k) \cdot \left[\left(\frac{\partial \mathbf{t}_k}{\partial \tilde{\mathbf{u}}_k}(\tilde{\mathbf{s}}_k) \right)^{\rightarrow} \cdot \frac{\partial \mathbf{t}_k}{\partial \tilde{\mathbf{u}}_k}(\tilde{\mathbf{s}}_k) \right] = \mathbf{0}, \left\{ \left(\frac{\partial y}{\partial \tilde{\mathbf{u}}_k}(\tilde{\mathbf{s}}_k) \neq \mathbf{0} \right), k = \overline{1, d}. \right.$$

$$\mathbf{t}_k(\tilde{\mathbf{s}}_k) = \mathbf{0};$$

(Её можно представить также в клеточно-диагональной форме.)

Множитель Лагранжа распадается на d частных подмножителей, т. е. $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_d)$, где

$$\lambda_k = \frac{\partial y}{\partial \tilde{\mathbf{u}}_k}(\tilde{\mathbf{s}}_k) \cdot \left(\frac{\partial \mathbf{t}_k}{\partial \tilde{\mathbf{u}}_k}(\tilde{\mathbf{s}}_k) \right)^{-1}_R.$$

Общая функция Лагранжа имеет вид:

$$\mathcal{L} = \mathcal{L}(\tilde{\mathbf{x}}, \lambda) \equiv f(\tilde{\mathbf{u}}_1, \tilde{\mathbf{u}}_2, \dots, \tilde{\mathbf{u}}_d) - \sum_{k=1}^d \lambda_k \cdot \mathbf{t}_k(\tilde{\mathbf{u}}_k).$$

Система (85) также распадается на d отдельных клеток:

$$\left. \begin{aligned} \frac{\partial \mathcal{L}}{\partial \tilde{\mathbf{u}}_k}(\tilde{\mathbf{s}}_k) &= \frac{\partial y}{\partial \tilde{\mathbf{u}}_k}(\tilde{\mathbf{s}}_k) - \lambda_k \cdot \frac{d\mathbf{t}}{d\tilde{\mathbf{u}}_k}(\tilde{\mathbf{s}}_k), \\ \frac{\partial \mathcal{L}}{\partial \lambda_k} &= -\mathbf{t}_k(\tilde{\mathbf{s}}_k) = \mathbf{0}; \end{aligned} \right\} k = \overline{1, d}.$$

Характер условной стационарности (экстремума) целевой функции в найденной из этой системы точке $\tilde{\mathbf{s}} = (\tilde{s}_1, \tilde{s}_2, \dots, \tilde{s}_d)$ задаёт, например, полуусловная матрица Гессе от функции Лагранжа размера $n \times n$. В данном случае она представима в форме прямой (клеточной) суммы по d диагональным клеткам:

$$\left[\frac{\partial^2 \mathcal{L}}{\partial \mathbf{x}' \partial \mathbf{x}}(\tilde{\mathbf{s}}) \right] = \oplus \sum_{k=1}^d \left[\frac{\partial^2 \mathcal{L}}{\partial \mathbf{u}'_k \partial \mathbf{u}_k}(\tilde{\mathbf{s}}_k) \right],$$

$$\left[\frac{\partial^2 \mathcal{L}}{\partial \mathbf{u}'_k \partial \mathbf{u}_k}(\tilde{\mathbf{s}}_k) \right] =$$

$$= \left[\left(\frac{d\mathbf{t}}{d\tilde{\mathbf{u}}_k}(\tilde{\mathbf{s}}_k) \right)' \cdot \frac{d\mathbf{t}}{d\tilde{\mathbf{u}}_k}(\tilde{\mathbf{s}}_k) \right] \cdot \frac{\partial^2 \mathcal{L}}{\partial \mathbf{u}'_k \partial \tilde{\mathbf{u}}_k}(\tilde{\mathbf{s}}_k) \cdot \left[\left(\frac{d\mathbf{t}}{d\tilde{\mathbf{u}}_k}(\tilde{\mathbf{s}}_k) \right)' \cdot \frac{d\mathbf{t}}{d\tilde{\mathbf{u}}_k}(\tilde{\mathbf{s}}_k) \right].$$

Если все указанные частные клетки отрицательно (положительно) знакоопределённые, то в точке $\tilde{\mathbf{s}}$ функция имеет условный максимум (минимум) 2-го уровня. Если клетки по разному знакоопределены, то в точке $\tilde{\mathbf{s}}$ функция имеет условную седловину 2-го уровня. Если некие клетки законоопределены, то нужен анализ условной матрицы Гессе.

§ 2.4. Предельные методы решения задач на условный экстремум с малым или с большим параметром

Пусть теперь в задаче на условный экстремум $y(\mathbf{x})$ с *ограниченной* переменной $\tilde{\mathbf{x}}$ (§ 2.2) вектор-функция $\mathbf{h}'(\mathbf{x})$ имеет природу градиента некоторой скалярной функции от $\mathbf{x}$. Тогда $\mathbf{h}(\mathbf{x})$ имеет обязательно ту же самую размерность n , что и координатное пространство $\mathcal{A}^n$:

$$\mathbf{h}'(\mathbf{x}) = \frac{d\varphi}{d\mathbf{x}} \leftrightarrow \frac{d\mathbf{h}}{d\mathbf{x}} = \left(\frac{d\mathbf{h}}{d\mathbf{x}} \right)' = S(\mathbf{x}). \quad (92)$$

$1 \times n \quad 1 \times n \quad n \times n \quad n \times n \quad n \times n$

Обратим внимание на то, что здесь уравнение связи $\mathbf{h}(\mathbf{x}) = \mathbf{0}$ задаёт полное множество $\langle \tilde{\mathbf{x}} \rangle$ нестрогой стационарности для некоей нецелевой скалярной функции $\varphi(\mathbf{x})$. Причём при $\text{rang} \{ d\mathbf{h} / d\mathbf{x} \} = r = n - q = \text{const}$ оно задаёт его как q -поверхность $\langle \tilde{\mathbf{x}} \rangle$. Тогда, если $\mathbf{c} \in \langle \tilde{\mathbf{x}} \rangle$, то исходное уравнение связи в *векторной* форме можно заменить на тождественное ему уравнение связи в *скалярной* форме:

$$\mathbf{h}(\mathbf{x}) = \mathbf{0} \leftrightarrow \int_{\mathbf{c}}^{\mathbf{x}} \mathbf{h}'(\mathbf{x}) d\mathbf{x} \equiv \mathbf{0}. \quad (93)$$

$1 \times n \quad n \times 1$

(Такая тождественность вызывается именно тем, что $d\mathbf{x}$ есть *свободный* дифференциал аргумента $\mathbf{x}$.) Далее заменим в системе (85) векторное уравнение связи на тождественное ему это скалярное уравнение связи:

$$\left. \begin{aligned} \frac{\partial \mathcal{L}_N}{\partial \mathbf{x}} = \frac{\partial \left(y(\mathbf{x}) + N \cdot \int_{\mathbf{c}}^{\mathbf{x}} \mathbf{h}'(\mathbf{x}) d\mathbf{x} \right)}{\partial \mathbf{x}} = \mathbf{0}, \\ \int_{\mathbf{c}}^{\mathbf{x}} \mathbf{h}'(\mathbf{x}) d\mathbf{x} = 0. \end{aligned} \right\} \quad (94)$$

Поскольку новое уравнение связи задаёт скалярная функция от $\mathbf{x}$, то и новый множитель Лагранжа есть также скалярный параметр $\lambda = -N$. Причём он бесконечный, так как из формулы (83) или непосредственно из (94) для нового множителя Лагранжа N следует формула:

$$N \cdot \frac{d \left[\int_{\mathbf{c}}^{\mathbf{x}} \mathbf{h}'(\mathbf{x}) d\mathbf{x} \right]}{d\mathbf{x}}(\tilde{\mathbf{x}}) = N \cdot \frac{d\varphi}{d\mathbf{x}}(\tilde{\mathbf{x}}) = N \cdot \mathbf{0} = -\frac{dy}{d\mathbf{x}}(\tilde{\mathbf{x}}) \neq \mathbf{0} \quad (95)$$

Отсюда вытекает, что $N = \pm \infty$! (См. об этом же в книге [27, с. 61], хотя автор пришёл к излагаемым *общим* предельным методам ещё в 1981 г.)

С учётом (94), определим *композиционные* функции 2-х видов:

$$\Phi_N(\mathbf{x}) = y(\mathbf{x}) + N \cdot \varphi(\mathbf{x}), \quad N \rightarrow \pm \infty; \quad (96)$$

$$\Phi_\alpha(\mathbf{x}) = \varphi(\mathbf{x}) + \alpha \cdot y(\mathbf{x}), \quad \alpha \rightarrow \pm 0; \quad (97)$$

которые прямо пропорциональны друг другу, так как

$$\Phi_\alpha(\mathbf{x}) = \alpha \cdot \Phi_N(\mathbf{x}) \leftrightarrow \Phi_N(\mathbf{x}) = N \cdot \Phi_\alpha(\mathbf{x}).$$

На основании вышесказанного пока можно только *предположить*, что тождественные градиентные уравнения вида

$$\frac{d\Phi_N}{d\mathbf{x}} = \frac{dy}{d\mathbf{x}} + N \cdot \frac{d\varphi}{d\mathbf{x}} = \frac{dy}{d\mathbf{x}} + N \cdot \mathbf{h}'(\mathbf{x}) = \mathbf{0}, \quad (98)$$

$$\frac{d\Phi_\alpha}{d\mathbf{x}} = \frac{d\varphi}{d\mathbf{x}} + \alpha \cdot \frac{dy}{d\mathbf{x}} = \mathbf{h}'(\mathbf{x}) + \alpha \cdot \frac{dy}{d\mathbf{x}} = \mathbf{0} \quad (99)$$

дают *в пределе* то же решение, что и система (85) при специальном требовании (92), т. е. точку или область условной стационарности $y(\mathbf{x})$. Левая часть этих уравнений есть *композиционный условный градиент*.

В вычислительной математике такого вида композиционные функции порождают соответственно два *общих предельных* метода решения задачи на условный экстремум с ограниченной переменной, а также *частные предельные* методы решения специфических задач типа поиска строгого условного экстремума одной функции на области нестрогого экстремума другой функции. Дадим строгое теоретическое обоснование этих методов с идентификацией природы условной стационарности целевой функции.

Рассмотрим предельное композиционное уравнение вида:

$$\mathbf{b}'_{\alpha}(\mathbf{x}) = \mathbf{h}'(\mathbf{x}) + \alpha \cdot \frac{d\mathbf{y}}{d\mathbf{x}} = \mathbf{0}, \quad \alpha \rightarrow 0, \quad (100)$$

где $\mathbf{h}(\mathbf{x}) \in \mathcal{A}^n$, $\text{rang}\{\mathbf{dh} / d\mathbf{x}\} = r = n - q < n$. Если решение этого уравнения $\tilde{\mathbf{c}}$ в пределе при $\alpha \rightarrow 0$ существует, то оно обязательно принадлежит $\langle \tilde{\mathbf{x}} \rangle$, так как при $\alpha \rightarrow 0$ уравнение (100) вырождается с точностью до бесконечно малой α в данное уравнение связи $\mathbf{h}(\mathbf{x}) = \mathbf{0}$, в том числе $\mathbf{h}(\tilde{\mathbf{c}}) = \mathbf{0}$. Обратим внимание на то, что уравнение связи опять предполагается разрешимым в виде q -поверхности $\langle \tilde{\mathbf{x}} \rangle$. Покажем, что при выполнении некоего определённого требования к функции $\mathbf{h}(\mathbf{x})$ уравнение (100) имеет решение, тождественное решению системы (77) при $m = n$ и $\text{rang}\{\mathbf{dh} / d\mathbf{x}\} = r = n - q$, т. е. тождественное именно решению задачи на условную стационарность целевой функции $\mathbf{y}(\mathbf{x})$ с уравнением связи $\mathbf{h}(\mathbf{x}) = \mathbf{0}$.

Теорема 3 (предельная). Пусть $\tilde{\mathbf{s}} \in \mathcal{A}^n$ — точка условной стационарности функции $\mathbf{y}(\mathbf{x})$ на q -поверхности $\langle \tilde{\mathbf{x}} \rangle$, заданной уравнением $\mathbf{h}(\mathbf{x}) = \mathbf{0}$, т. е. $\tilde{\mathbf{s}}$ есть решение системы (77) при $m = n$ и $\text{rang}\{\mathbf{dh} / d\mathbf{x}\} = r = n - q$. Точка $\tilde{\mathbf{s}}$ является решением предельного уравнения (100) при $\alpha \rightarrow 0$ тогда и только тогда, когда

$$\left(\frac{d\mathbf{y}}{d\mathbf{x}}(\tilde{\mathbf{s}}) \right)' \in \left\langle \ker \frac{d\mathbf{h}}{d\mathbf{x}}(\tilde{\mathbf{s}}) \cap \ker \left(\frac{d\mathbf{h}}{d\mathbf{x}}(\tilde{\mathbf{s}}) \right)' \right\rangle. \quad (101)$$

(Разумеется, предполагается, что матрица Якоби для $\mathbf{h}(\mathbf{x})$ такова, что вышеуказанное пересечение множеств не вырождено в нуль-вектор.)

Следствие 1. Точка $\tilde{\mathbf{s}}$ является решением предельного уравнения (100) при $\alpha \rightarrow 0$, когда матрица Якоби функции $\mathbf{h}(\mathbf{x})$ *нуль-нормальная*, т. е.

$$\left\langle \text{im} \frac{d\mathbf{h}}{d\mathbf{x}} \right\rangle \equiv \left\langle \text{im} \left(\frac{d\mathbf{h}}{d\mathbf{x}} \right)' \right\rangle \leftrightarrow \left\langle \ker \frac{d\mathbf{h}}{d\mathbf{x}} \right\rangle \equiv \left\langle \ker \left(\frac{d\mathbf{h}}{d\mathbf{x}} \right)' \right\rangle. \quad (102)$$

(При этом, в частности, она нормальная, симметричная и т. д.)

Следствие 2. Теорема остаётся справедливой, если уравнение (100) заменить на тождественное ему уравнение с параметром $N = 1/\alpha$:

$$\mathbf{b}'_N(\mathbf{x}) = \frac{dy}{dx} + N \cdot \mathbf{h}'(\mathbf{x}) = \mathbf{0}, \quad N \rightarrow \infty. \quad (103)$$

Доказательство. Если решение $\mathbf{x}(\alpha)$ уравнения (100) при $\alpha \rightarrow 0$ в пределе существует, то оно с точностью до конечной величины принадлежит множеству решений уравнения связи $\mathbf{h}(\mathbf{x}) = \mathbf{0}$, т. е. $\langle \tilde{\mathbf{x}} \rangle$, но с точностью до бесконечно малой $\alpha \rightarrow 0$ отстаёт от неё. В противном случае имело бы место — либо $\mathbf{h}[\mathbf{x}(\alpha)] \neq \mathbf{0}$ и уравнение (100) при $\alpha \rightarrow 0$ не выполнялось; либо $dy/dx[\mathbf{x}(\alpha)] = \mathbf{0}$ и требование задачи именно на условный экстремум $dy/dx \neq \mathbf{0}$ не соблюдалось. Эта бесконечно малая разность для некоторой точки $\mathbf{c} \subset \langle \tilde{\mathbf{x}} \rangle$ составляет $[\mathbf{x}_c(\alpha) - \tilde{\mathbf{c}}]$. При этом значение $\mathbf{h}[\mathbf{x}_c(\alpha)]$ с точностью до $\alpha \rightarrow 0$ можно выразить, используя первый член разложения вектор-функции $\mathbf{h}(\mathbf{x})$ в ряд Тейлора:

$$\mathbf{h}[\mathbf{x}_c(\alpha)] = \frac{d\mathbf{h}}{d\mathbf{x}}(\tilde{\mathbf{c}}) \cdot [\mathbf{x}_c(\alpha) - \tilde{\mathbf{c}}].$$

Подставляя это значение вместо $\mathbf{h}(\mathbf{x})$ в (100), получаем:

$$(\mathbf{x}_c(\alpha) - \tilde{\mathbf{c}})' \cdot \left(\frac{d\mathbf{h}}{d\mathbf{x}}(\tilde{\mathbf{c}}) \right)' + \alpha \cdot \frac{dy}{dx}(\tilde{\mathbf{c}}) = \mathbf{0}.$$

Это линейное уравнение с точностью до $\alpha \rightarrow 0$ разрешимо тогда и только тогда, когда, например, с использованием соответствующего характеристического симметричного проектора имеем требование:

$$\left(\frac{dy}{dx}(\tilde{\mathbf{c}}) \right)' \in \left\langle \text{im} \frac{d\mathbf{h}}{d\mathbf{x}}(\tilde{\mathbf{c}}) \right\rangle \leftrightarrow \left[\frac{d\mathbf{h}}{d\mathbf{x}}(\tilde{\mathbf{c}}) \cdot \left(\frac{d\mathbf{h}}{d\mathbf{x}}(\tilde{\mathbf{c}}) \right)' \right] \cdot \left(\frac{dy}{dx}(\tilde{\mathbf{c}}) \right)' = \mathbf{0}. \quad (104)$$

При выполнении этого требования общее решение $\mathbf{x}_c(\alpha)$ с точностью до $\alpha \rightarrow 0$ выражается явно из линейного уравнения:

$$\mathbf{x}_c(\alpha) = \tilde{\mathbf{c}} - \alpha \cdot \left(\frac{d\mathbf{h}}{d\mathbf{x}}(\tilde{\mathbf{c}}) \right)^+ \cdot \left(\frac{dy}{dx}(\tilde{\mathbf{c}}) \right)' + \alpha \cdot \left\langle \ker \frac{d\mathbf{h}}{d\mathbf{x}}(\tilde{\mathbf{c}}) \right\rangle. \quad (105)$$

Здесь общее решение линейного уравнения выражено в канонической форме — через его частное *нормальное* решение с использованием квазиобратной матрицы Мура—Пенроуза.

Соответственно при $\alpha \rightarrow 0$: $\mathbf{x}_c(\alpha) \rightarrow \tilde{\mathbf{c}} \in \langle \tilde{\mathbf{x}} \rangle$. В свою очередь, отклонение $\mathbf{h}[\mathbf{x}_c(\alpha)]$ от нулевого значения в α -окрестности $\langle \tilde{\mathbf{x}} \rangle$ с той же точностью $\alpha \rightarrow 0$ составляет

$$\mathbf{h}[\mathbf{x}_c(\alpha)] = -\alpha \cdot \frac{d\mathbf{h}}{d\mathbf{x}}(\tilde{\mathbf{c}}) \cdot \left(\frac{d\mathbf{h}}{d\mathbf{x}}(\tilde{\mathbf{c}}) \right)^+ \cdot \left(\frac{d\mathbf{y}}{d\mathbf{x}}(\tilde{\mathbf{c}}) \right)' = -\alpha \cdot \left[\frac{d\mathbf{h}}{d\mathbf{x}}(\tilde{\mathbf{c}}) \cdot \left(\frac{d\mathbf{h}}{d\mathbf{x}}(\tilde{\mathbf{c}}) \right)' \right] \cdot \left(\frac{d\mathbf{y}}{d\mathbf{x}}(\tilde{\mathbf{c}}) \right)'.$$

(Обратим внимание на то, что используемые два характеристических проектора отличаются от двух родственных симметричных проекторов, применённых ранее в § 2.2, но в сумме они также составляют $\mathbf{I}_{n \times n}$.)

Подставляя опять в (100), но уже с точностью до $\alpha \rightarrow 0$ выражение для $\mathbf{h}(\mathbf{x})$ в α -окрестности точки $\tilde{\mathbf{c}} \in \langle \tilde{\mathbf{x}} \rangle$, окончательно получаем:

$$\mathbf{b}'_\alpha(\tilde{\mathbf{c}}) = +\alpha \cdot \frac{d\mathbf{y}}{d\mathbf{x}}(\tilde{\mathbf{c}}) \cdot \left[\frac{d\mathbf{h}}{d\mathbf{x}}(\tilde{\mathbf{c}}) \cdot \left(\frac{d\mathbf{h}}{d\mathbf{x}}(\tilde{\mathbf{c}}) \right)' \right] = \mathbf{0}. \quad (106)$$

Но решение этого уравнения на множестве $\langle \tilde{\mathbf{x}} \rangle$ с точностью до $\alpha \rightarrow 0$ тождественно решению $\tilde{\mathbf{s}}$ градиентного уравнения вида

$$\alpha \cdot \left[\frac{d\mathbf{y}}{d\mathbf{x}}(\tilde{\mathbf{s}}) \right] = \alpha \cdot \frac{d\mathbf{y}}{d\mathbf{x}}(\tilde{\mathbf{s}}) \cdot \left[\left(\frac{d\mathbf{h}}{d\mathbf{x}}(\tilde{\mathbf{s}}) \right)' \cdot \frac{d\mathbf{h}}{d\mathbf{x}}(\tilde{\mathbf{s}}) \right] = \mathbf{0} \quad (107)$$

тогда и только тогда, когда выполняется требование (101), или тогда, когда выполняются менее строгие требования (102) или (92). При этом следует учесть, что, в силу (100), $\tilde{\mathbf{c}} \wedge \tilde{\mathbf{s}} \in \langle \tilde{\mathbf{x}} \rangle$. ***

По сути, точка $\tilde{\mathbf{s}}$ здесь является совместным решением систем:

$$\left. \begin{aligned} \frac{d\mathbf{y}}{d\mathbf{x}} \cdot \left[\frac{d\mathbf{h}}{d\mathbf{x}} \cdot \left(\frac{d\mathbf{h}}{d\mathbf{x}} \right)' \right] &= \mathbf{0}, \\ \mathbf{h}(\mathbf{x}) &= \mathbf{0}; \end{aligned} \right\} \quad \text{и} \quad \left. \begin{aligned} \frac{d\mathbf{y}}{d\mathbf{x}} \cdot \left[\left(\frac{d\mathbf{h}}{d\mathbf{x}} \right)' \cdot \frac{d\mathbf{h}}{d\mathbf{x}} \right] &= \mathbf{0}, \\ \mathbf{h}(\mathbf{x}) &= \mathbf{0}. \end{aligned} \right\}$$

Вторая из них есть система (77) при $m = n$. В случае выполнения менее строгих требований (102) или (92), нежели (101), проекторы в обеих системах становятся тождественными.

Отметим также то, что условие (101) требует, чтобы $n \times n$ -матрица Якоби $\left\{ \frac{d\mathbf{h}}{d\mathbf{x}} \right\}(\tilde{\mathbf{s}})$ фактически имела хотя бы один тождественный левый и правый сингулярный собственный вектор, коллинеарный в точке $\tilde{\mathbf{s}}$ вектору градиента $d\mathbf{y}/d\mathbf{x}$!

Менее жёсткое требование к $\mathbf{h}(\mathbf{x})$ состоит в *нуль-нормальности* её матрицы Якоби $d\mathbf{h}/d\mathbf{x}$. Признаки принадлежности к собственному множеству и свойства для нуль-нормальных $n \times n$ -матриц (как и более общих *нуль-простых*) были ранее установлены и подробно рассмотрены в монографии [27, с. 36 - 50]. В частности, к их полному множеству относятся нормальные $\mathbf{N}$, симметричные $\mathbf{S}$ и кососимметричные $\mathbf{K}$ матрицы, а также положительно и отрицательно полуопределённые симметричные матрицы, выражаемые как $\pm \mathbf{A}'\mathbf{A}$ и $\pm \mathbf{A}\mathbf{A}'$. Отметим тут важное геометрическое свойство любой нуль-нормальной матрицы: на евклидовом координатном пространстве её образ и ядро составляют прямую евклидово ортогональную сумму. Например, в силу именно такого свойства, *весьма существенное в общих предельных методах компенсационное требование (104) в случае произвольных целевых функций $y(\mathbf{x})$ на α -окрестности выполняется тогда и только тогда, когда матрица Якоби нуль-нормальная*. Кроме того, и более общие нуль-простые матрицы приводятся модальным преобразованием $\mathbf{V} = \mathbf{V}'$ исходного базиса к нуль-нормальной форме [27, с. 61].

Из анализа хорошо известно, что $\mathbf{h}'(\mathbf{x})d\mathbf{x} = d\varphi(\mathbf{x})$, т. е. является полным дифференциалом при *свободном $d\mathbf{x}$ и ненулевой $\mathbf{h}(\mathbf{x})$* , как в (92), тогда и только тогда, когда матрица Якоби $d\mathbf{h}/d\mathbf{x}$ симметричная. Однако, если на некоем непрерывном подмножестве $\langle \tilde{\mathbf{x}} \rangle$ $n \times 1$ -вектор-функция $\mathbf{h}(\mathbf{x})$ повсюду нулевая, то тогда и $\mathbf{h}'(\mathbf{x})d\tilde{\mathbf{x}} = 0$ формально также является полным (нулевым) дифференциалом $d\varphi$ для функции $\varphi(\mathbf{x}) = \text{const}$ (несмотря на то, что матрица Якоби $d\mathbf{h}/d\mathbf{x}$ на $\langle \tilde{\mathbf{x}} \rangle$ может и не быть симметричной). Поэтому $n \times 1$ -вектор-функция $\mathbf{h}(\mathbf{x})$, нулевая на области $\langle \tilde{\mathbf{x}} \rangle$, формально на ней же всегда интегрируемая! Причём в весьма важных частных случаях $n \times 1$ -вектор-функция $\mathbf{h}(\mathbf{x})$ удовлетворяет требованию (92), т. е. является именно градиентной функцией на области определения. Тогда, интегрируя вектор-функцию $\mathbf{b}_\alpha(\mathbf{x})$ в (100), получаем с точностью до постоянной (97). Не нарушая общности *теоремы 3* и её следствий, умножив обе части функции на параметр $N = 1/\alpha$ — множитель Лагранжа в (94), получаем композиционную функцию (96) в форме функции Лагранжа с бесконечно большим скалярным множителем. При этом правая часть уравнения (98) тождественна (100), что подтверждает следствие 2 теоремы 3. Следовательно, градиентные уравнения (98) и (99), как выше и предполагалось, дают в пределе то же самое решение, что и системы (85), (77) в задаче на условную стационарность функции $y(\mathbf{x})$ при ограничении $\mathbf{h}'(\mathbf{x}) = d\varphi/d\mathbf{x} = \mathbf{0}$.

Теперь пусть $\mathbf{h}(\mathbf{x})$ — произвольная $m \times 1$ -вектор-функция, но по-прежнему на q -поверхности $\langle \tilde{\mathbf{x}} \rangle$ и только на ней она нулевая: $\mathbf{h}(\tilde{\mathbf{x}}) = \mathbf{0}$.

Составляем композиционную **функцию Куранта** вида:

$$K_N(\mathbf{x}) = y(\mathbf{x}) + N \cdot \varphi(\mathbf{x}) = y(\mathbf{x}) + N \cdot \frac{1}{2} \mathbf{h}'(\mathbf{x}) \cdot \mathbf{h}(\mathbf{x}). \quad (108)$$

Такие функции применяются в особом численном методе условной оптимизации, который предложил Р. Курант в середине XX века, — *методе штрафной функции* [10]. Роль как бы штрафа, налагаемого в процессе оптимизации на целевую функцию $y(\mathbf{x})$, играет значительно увеличиваемое множителем N второе слагаемое в (108), образуемое, например, квадратом евклидова модуля вектор-функции ограничения, — *штрафная функция* $N \cdot \varphi(\mathbf{x})$. В **методе штрафной функции Куранта** скалярный множитель N выбирают достаточно большим исходя из возможностей вычислительной техники и имеющейся погрешности.

Роль $\varphi(\mathbf{x})$ в (108), т. е. нецелевой скалярной функции $\varphi(\mathbf{x})$ из (96), например, выполняет специальная функция $\varphi(\mathbf{x}) = \frac{1}{2} \mathbf{h}'(\mathbf{x}) \cdot \mathbf{h}(\mathbf{x})$ — такая, чтобы уравнение $\frac{d\varphi}{d\mathbf{x}} = \mathbf{h}'(\mathbf{x}) \cdot \frac{d\mathbf{h}}{d\mathbf{x}} = \mathbf{g}(\mathbf{x}) = \mathbf{0}$ имело то же общее решение $\langle \tilde{\mathbf{x}} \rangle$, что и исходное уравнение связи $\mathbf{h}(\mathbf{x}) = \mathbf{0}$. Это имеет место, так как $\text{rang} \{ d\mathbf{h} / d\mathbf{x} \} = m$, или $\langle \ker d\mathbf{h} / d\mathbf{x} \rangle = \mathbf{0}$. Тогда композиционное дифференциальное уравнение, производимое из (108), вида

$$\frac{dK_N}{d\mathbf{x}} = \frac{dy}{d\mathbf{x}} + N \cdot \frac{d\varphi}{d\mathbf{x}} = \frac{dy}{d\mathbf{x}} + N \cdot \mathbf{g}(\mathbf{x}) = \frac{dy}{d\mathbf{x}} + N \cdot \mathbf{h}'(\mathbf{x}) \cdot \frac{d\mathbf{h}}{d\mathbf{x}} = \mathbf{0} \quad (109)$$

теоретически при $N \rightarrow \infty$ даёт на $\langle \tilde{\mathbf{x}} \rangle$ то же самое решение $\tilde{\mathbf{s}}$ задачи на условную стационарность функции $y(\mathbf{x})$ при ограничении $\mathbf{h}(\mathbf{x}) = \mathbf{0}$, что и тождественные системы уравнений (85) и (77).

Разумеется, аналогичным образом эту же скалярную функцию $\varphi(\mathbf{x})$ можно использовать в тождественном методе с малым параметром α :

$$K_\alpha(\mathbf{x}) = \varphi(\mathbf{x}) + \alpha \cdot y(\mathbf{x}) = \frac{1}{2} \mathbf{h}'(\mathbf{x}) \cdot \mathbf{h}(\mathbf{x}) + \alpha \cdot y(\mathbf{x}), \quad (110)$$

$$\frac{dK_\alpha}{d\mathbf{x}} = \frac{d\varphi}{d\mathbf{x}} + \alpha \cdot \frac{dy}{d\mathbf{x}} = \mathbf{g}(\mathbf{x}) + \alpha \cdot \frac{dy}{d\mathbf{x}} = \mathbf{h}'(\mathbf{x}) \cdot \frac{d\mathbf{h}}{d\mathbf{x}} + \alpha \cdot \frac{dy}{d\mathbf{x}} = \mathbf{0}. \quad (111)$$

Обе композиционные предельные функции, аналогичные (96) и (97), связаны друг с другом через соотношения:

$$K_\alpha(\mathbf{x}) = \frac{1}{N} \cdot K_N(\mathbf{x}), \quad K_N(\mathbf{x}) = \frac{1}{\alpha} \cdot K_\alpha(\mathbf{x}). \quad (112)$$

* * *

Если в обеих композиционных функциях (96) и (97) поменять местами целевую $y(\mathbf{x})$ и вспомогательную $\varphi(\mathbf{x})$ функции, то приходим к предельным методам решения экстремальных задач иного вида — типа поиска строгого экстремума одной функции (вспомогательной) на *нестрогом* или *плохо обусловленном* экстремуме другой функции (целевой). Составляем композиционную **функцию Тихонова** вида:

$$T_{\alpha}(\mathbf{x}) = y(\mathbf{x}) + \alpha \cdot \theta(\mathbf{x}). \quad (113)$$

Такого типа функции применяются в методе устойчивой *безусловной* оптимизации, который предложил А. Н. Тихонов в 60-х годах XX века, — **методе регуляризации Тихонова** [36, 37]. Роль вырожденного или плохо обусловленного уравнения связи играет градиентное уравнение:

$$\mathbf{g}(\mathbf{x}) = \frac{dy}{dx}(\mathbf{x}) = \mathbf{0} \rightarrow \mathbf{x} \in \langle \tilde{\mathbf{x}} \rangle. \quad (114)$$

При точных элементах в (114) стационарность имеет вырожденный характер и реализуется на гладкой и регулярной q -поверхности $\langle \tilde{\mathbf{x}} \rangle$. Однако, при неточных элементах в (114) стационарность может быть почти вырожденной (*плохо обусловленной*). Тогда даже незначительные изменения в расчётных данных из-за их исходных погрешностей и ошибок от округлений при вычислениях приводят к очень большому влиянию на конечный результат, т. е. в данном случае на решение конкретной задачи на безусловный экстремум целевой функции $y(\mathbf{x})$.

Для получения *корректного* по Тихонову устойчивого точечного решения такого типа задач безусловной оптимизации в (113) как раз и применяется вспомогательная функция $\alpha \cdot \theta(\mathbf{x})$, или *стабилизатор*. Скалярный множитель α выбирают достаточно малым по величине — так, чтобы стабилизатор был на уровне оценки погрешности для $y(\mathbf{x})$.

Типичный пример использования метода регуляризации Тихонова — вычисление *нормального решения* $\mathbf{s}^{\bullet}$ на q -поверхности $\langle \tilde{\mathbf{x}} \rangle$, задаваемой плохо обусловленным градиентным уравнением вида (114). Решение $\mathbf{s}^{\bullet}$ отвечает тут требованию минимума его евклидовой нормы на общем решении $\langle \tilde{\mathbf{x}} \rangle$ градиентного уравнения (*с переходом в $\mathcal{E}^n$ и декартов базис*), с учётом неточности исходных данных и округлений:

$$T_{\alpha}(\mathbf{x}) = y(\mathbf{x}) + \alpha \cdot \theta(\mathbf{x}) = y(\mathbf{x}) + \alpha \cdot \frac{1}{2} \mathbf{x}' \cdot \mathbf{x}, \quad (115)$$

$$\frac{dT_{\alpha}}{dx} = \frac{dy}{dx} + \alpha \cdot \frac{d\theta}{dx} = \mathbf{g}(\mathbf{x}) + \alpha \cdot \mathbf{x}' = \mathbf{0} \rightarrow \mathbf{x} = \mathbf{s}^{\bullet}. \quad (116)$$

В данном аспекте использования предельного метода уравнение связи (114) предполагается также разрешимым, но в виде q -поверхности $\langle \tilde{\mathbf{x}} \rangle$. Поэтому матрица Гессе $d^2y/d\mathbf{x}'d\mathbf{x}$ на этой области экстремума $y(\mathbf{x})$ обязательно сингулярная. Однако сам метод регуляризации Тихонова, применительно к экстремальным задачам, трактуется гораздо шире. Часто метод используется для устойчивого решения разнообразных задач безусловной оптимизации функции $y(\mathbf{x})$ с *нестрогим* или с *плохо обусловленным* экстремумом. Тогда проблемы устойчивости возникают и с анализом на знакоопределённость плохо обусловленной матрицы Гессе (§ 1.9), играющей в задаче фактически роль матрицы Якоби функции ограничения, вырожденной и симметричной по структуре:

$$\frac{d^2y}{d\mathbf{x}'d\mathbf{x}}(\tilde{\mathbf{x}}) = \frac{d\mathbf{g}}{d\mathbf{x}'}(\tilde{\mathbf{x}}) = \left[\frac{d\mathbf{g}}{d\mathbf{x}'}(\tilde{\mathbf{x}}) \right]'. \quad (117)$$

Вычисление же матрицы Гессе с применением стабилизатора $\alpha \cdot \theta(\mathbf{x})$, т. е. с последующим дифференцированием градиента в форме (116), позволяет устранить имеющуюся плохую обусловленность, так как *стабилизированная* симметричная матрица Гессе всегда несингулярная, в силу её особой *резольвентной* структуры:

$$\frac{d^2T_\alpha}{d\mathbf{x}'d\mathbf{x}}(\tilde{\mathbf{s}}) = \left\{ \frac{d^2y}{d\mathbf{x}'d\mathbf{x}}(\tilde{\mathbf{s}}) + \alpha \cdot \mathbf{I} \right\}. \quad (118)$$

Заметим, что метод регуляризации Тихонова не повышает точность решения экстремальной задачи: он лишь делает процедуру её решения устойчивой, например, к погрешностям опыта и ошибкам округления, а результат — однозначным. А это тоже немаловажно в вычислительном аспекте. Но инвариантность к масштабным преобразованиям теряется.

* * *

В изложенных выше предельных методах решение $\tilde{\mathbf{s}}$, отвечающее условной стационарности, при вычислениях может дополнительно уточняться аппроксимацией по естественным формулам:

$$\tilde{\mathbf{s}} = \frac{\tilde{\mathbf{s}}(+\alpha) + \tilde{\mathbf{s}}(-\alpha)}{2}, \quad \tilde{\mathbf{s}} = \frac{\tilde{\mathbf{s}}(+N) + \tilde{\mathbf{s}}(-N)}{2}. \quad (119, 120)$$

Кроме того, отдельно отметим, что нормальное решение задачи на вырожденный *условный* экстремум функции $y(\mathbf{x})$ предельным методом с малыми параметрами α и α^2 может рассматриваться как его *двукратное* применение. Задачи подобного типа классифицируются тут как поиск экстремума одной целевой скалярной функции на области нестрогого экстремума другой целевой скалярной функции.

Эту цепочку можно, в принципе, продолжать, составляя исходную композиционную функцию $\Phi_\alpha(\mathbf{x})$ или $\Phi_N(\mathbf{x})$ в форме многочлена либо от α , либо от N . Например, в двухзвенном варианте вычисляются:

- максимум $y_1(\mathbf{x})$ на нестрогом максимуме $y_2(\mathbf{x})$, или *максимакс*;
- максимум $y_1(\mathbf{x})$ на нестрогом минимуме $y_2(\mathbf{x})$, или *максимин*;
- минимум $y_1(\mathbf{x})$ на нестрогом максимуме $y_2(\mathbf{x})$, или *минимакс*;
- минимум $y_1(\mathbf{x})$ на нестрогом минимуме $y_2(\mathbf{x})$, или *минимин*.

Далее перейдём к анализу характера условной стационарности $y(\tilde{\mathbf{x}})$ в точке $\tilde{\mathbf{s}}$. Коль скоро в результате решения предельного уравнения типа (100) или (102), в том числе градиентного типа (98), (99) или (109), (110), вычислена точка условной стационарности $\tilde{\mathbf{s}}$, то в ней же возможно вычислить полуусловную и условную матрицу Гессе в (89). (Знакоопределённая полуусловная матрица Гессе в случае линейного уравнения связи, т. е. на плоскости в координатном пространстве, сама по себе задаёт характер строгого условного экстремума.)

Однако, возникает вопрос: а можно ли строго установить природу условной стационарности $y(\tilde{\mathbf{x}})$ тем же самым предельным методом, развивая его до логического завершения? Ответ на это даёт:

Теорема 4 (предельная). Характер условной стационарности функции $y(\mathbf{x})$ при ограничении уравнением связи $\mathbf{h}(\mathbf{x}) = \mathbf{0}$ и матрице Якоби $d\mathbf{h}/d\mathbf{x}$, удовлетворяющей требованию теоремы 3, в точке условной стационарности $\tilde{\mathbf{s}}$ задаёт симметричная по структуре композиционная условная матрица Гессе:

$$\frac{d^2\Phi_\alpha}{d\mathbf{x}'d\mathbf{x}}(\tilde{\mathbf{s}}) = \left(\frac{d\mathbf{b}_\alpha}{d\mathbf{x}}(\tilde{\mathbf{s}}) \right)',$$

вычисляемая в общем случае (для уровня стационарности $p = 2$) с точностью до $\alpha \rightarrow +0$.

Доказательство. Дифференцируя вектор-функцию $\mathbf{b}'_\alpha(\mathbf{x})$ согласно (100) или дифференцируя дважды скалярную функцию $\Phi_\alpha(\mathbf{x})$ согласно (97) в α -окрестности q -поверхности $\mathbf{h}(\mathbf{x}) = \mathbf{0}$, получаем $n \times n$ -матрицу:

$$\frac{d^2\Phi_\alpha}{d\mathbf{x}'d\mathbf{x}}(\tilde{\mathbf{x}}) = \left(\frac{d\mathbf{b}_\alpha}{d\mathbf{x}}(\tilde{\mathbf{x}}) \right)' = \frac{d\mathbf{h}}{d\mathbf{x}}(\tilde{\mathbf{x}}) + \alpha \cdot \frac{d^2y}{d\mathbf{x}'d\mathbf{x}}(\tilde{\mathbf{x}}). \quad (121)$$

Она определяется как композиционная предельная условная матрица Гессе. Вычислим её значение в точке $\mathbf{x}_s(\alpha)$ с точностью до $\alpha \rightarrow +0$, используя первые два члена разложения матрицы Якоби в ряд Тейлора, с учётом формулы (105):

$$\begin{aligned}
\frac{d^2\Phi_\alpha}{d\mathbf{x}'d\mathbf{x}}(\tilde{\mathbf{s}}) &= \left(\frac{d\mathbf{h}_\alpha}{d\mathbf{x}}(\tilde{\mathbf{s}}) \right)' = \frac{d\mathbf{h}}{d\mathbf{x}}(\tilde{\mathbf{s}}) + \frac{d^2\mathbf{h}}{d\mathbf{x}'d\mathbf{x}}(\tilde{\mathbf{s}}) \cdot [\mathbf{x}_s(\alpha) - \tilde{\mathbf{s}}] + \alpha \cdot \frac{d^2\mathbf{y}}{d\mathbf{x}'d\mathbf{x}}(\tilde{\mathbf{s}}) = \\
&= \frac{d\mathbf{h}}{d\mathbf{x}}(\tilde{\mathbf{s}}) - \alpha \cdot \frac{d^2\mathbf{h}}{d\mathbf{x}'d\mathbf{x}}(\tilde{\mathbf{s}}) \cdot \left[\left(\frac{d\mathbf{h}}{d\mathbf{x}}(\tilde{\mathbf{s}}) \right)^+ \cdot \left(\frac{d\mathbf{y}}{d\mathbf{x}}(\tilde{\mathbf{s}}) \right)' \right] + \alpha \cdot \frac{d^2\mathbf{y}}{d\mathbf{x}'d\mathbf{x}}(\tilde{\mathbf{s}}) = \\
&= \frac{d\mathbf{h}}{d\mathbf{x}}(\tilde{\mathbf{s}}) - \alpha \cdot \left[\frac{d\mathbf{y}}{d\mathbf{x}}(\tilde{\mathbf{s}}) \cdot \left(\frac{d\mathbf{h}}{d\mathbf{x}}(\tilde{\mathbf{s}}) \right)^+ \right] \cdot \frac{d^2\mathbf{h}}{d\mathbf{x}'d\mathbf{x}}(\tilde{\mathbf{s}}) + \alpha \cdot \frac{d^2\mathbf{y}}{d\mathbf{x}'d\mathbf{x}}(\tilde{\mathbf{s}}) = \\
&= \frac{d\mathbf{h}}{d\mathbf{x}}(\tilde{\mathbf{s}}) + \alpha \cdot \left\{ \frac{d^2\mathbf{y}}{d\mathbf{x}'d\mathbf{x}}(\tilde{\mathbf{s}}) - \left[\frac{d\mathbf{y}}{d\mathbf{x}}(\tilde{\mathbf{s}}) \cdot \left(\frac{d\mathbf{h}}{d\mathbf{x}}(\tilde{\mathbf{s}}) \right)^+ \right] \cdot \frac{d^2\mathbf{h}}{d\mathbf{x}'d\mathbf{x}}(\tilde{\mathbf{s}}) \right\} = \quad (122)
\end{aligned}$$

$$= \frac{d\mathbf{h}}{d\mathbf{x}}(\tilde{\mathbf{s}}) + \alpha \cdot \frac{d^2\mathcal{L}}{d\mathbf{x}'d\mathbf{x}}(\tilde{\mathbf{s}}). \quad (123)$$

Тут, во-первых, используется тот факт, что ядро матрицы Якоби переносится и на её производную: $\frac{d^2\mathbf{h}}{d\mathbf{x}'d\mathbf{x}} \cdot \left\langle \ker \frac{d\mathbf{h}}{d\mathbf{x}} \right\rangle = \mathbf{Z}_{n \times n}$. Во-вторых, во 2-й и 3-й строке цепи (122) при транспонировании 2-го слагаемого (симметричной $n \times n$ -матрицы) квазиобратная матрица Мура—Пенроуза переносится влево как есть, с учётом необходимого и достаточного требования (101) или более широких требований (102) и (92). Здесь используются те же обозначения и приёмы умножения трёхмерной матрицы частных производных, что и ранее в формуле (89). Отметим, что в (122) требуется вычисление квазиобратной матрицы Мура—Пенроуза от сингулярной $n \times n$ -матрицы (см., например, в [27, с. 51]).

Из сравнения формул (122) и (89) видно, что предельная условная матрица Гессе есть композиция из матрицы Якоби и полуусловной матрицы Гессе с множителем $+\alpha$. В общем случае, отвечающем (101), или в менее общих случаях, отвечающих (102) и (92), $n \times n$ -матрица Гессе (122) при $\alpha \rightarrow +0$ имеет $n - q = r = \text{rang}\{\mathbf{dh}/d\mathbf{x}\}$ конечных собственных значений, равных таковым же для $n \times n$ -матрицы Якоби, и $q = n - r = n - \text{rang}\{\mathbf{dh}/d\mathbf{x}\}$ бесконечно малых собственных значений, пропорциональных параметру $\alpha \rightarrow +0$.

Эта матрица при $\alpha \rightarrow +0$ проявляет себя как $n \times n$ -матрица Якоби от $\mathbf{h}(\mathbf{x})$ на подпространстве её *ненулевых* собственных значений, т. е. на $\langle \text{im } d\mathbf{h} / d\mathbf{x} \rangle_R$, и как аффинная проекция полуусловной $n \times n$ -матрицы Гессе от функции $y(\mathbf{x})$ на $\langle \ker d\mathbf{h} / d\mathbf{x} \rangle$ параллельно $\langle \text{im } d\mathbf{h} / d\mathbf{x} \rangle_R$ на подпространстве её *нулевых* собственных значений, т. е. на касательной q -плоскости $\langle L \rangle$ к q -поверхности $\langle \tilde{\mathbf{x}} \rangle$ в точке $\tilde{\mathbf{s}}$. Для того чтобы далее от её бесконечно малых собственных значений перейти к *условным* собственным значениям матрицы Гессе, нужно их поделить на α . Тогда они своими знаками и значимостью обычным образом (§ 1.9) задают характер условной стационарности функции в точке $\tilde{\mathbf{s}}$. При количестве ненулевых значений менее q условная стационарность нестрогая. ***

При линейном ограничении с матрицей Якоби $d\mathbf{h} / d\mathbf{x} = \mathbf{B} = \text{Const}$ того же размера $n \times n$ и ранга $r = n - q$ композиционная матрица (122) значительно упрощается:

$$\frac{d^2\Phi_\alpha}{d\mathbf{x}'d\mathbf{x}}(\tilde{\mathbf{s}}) = \left(\frac{d\mathbf{b}_\alpha}{d\mathbf{x}}(\tilde{\mathbf{s}}) \right)' = \frac{d\mathbf{h}}{d\mathbf{x}}(\tilde{\mathbf{s}}) + \alpha \cdot \frac{d^2y}{d\mathbf{x}'d\mathbf{x}}(\tilde{\mathbf{s}}). \quad (124)$$

Если же как, например, в (110), (111) применяется уравнение связи

$$\mathbf{g}(\mathbf{x}) = \frac{d\varphi}{d\mathbf{x}} = \mathbf{h}'(\mathbf{x}) \cdot \frac{d\mathbf{h}}{d\mathbf{x}} = \mathbf{0} \leftrightarrow \mathbf{h}(\mathbf{x}) = \mathbf{0},$$

где $\text{rang} \{d\mathbf{h} / d\mathbf{x}\} = r = n - q$, или $\langle \ker d\mathbf{h} / d\mathbf{x} \rangle_L = \mathbf{0}$, то предельная композиционная условная матрица Гессе от $y(\mathbf{x})$ с точностью $\alpha \rightarrow +0$ соответственно симметричная:

$$\begin{aligned} \frac{d^2\Phi_\alpha}{d\mathbf{x}'d\mathbf{x}}(\tilde{\mathbf{s}}) &= \frac{d\mathbf{b}_\alpha}{d\mathbf{x}}(\tilde{\mathbf{s}}) = \\ &= \frac{d\mathbf{g}'}{d\mathbf{x}}(\tilde{\mathbf{s}}) + \alpha \cdot \left\{ \frac{d^2y}{d\mathbf{x}'d\mathbf{x}}(\tilde{\mathbf{s}}) - \left[\frac{d\mathbf{y}}{d\mathbf{x}}(\tilde{\mathbf{s}}) \cdot \left(\frac{d\mathbf{g}'}{d\mathbf{x}}(\tilde{\mathbf{s}}) \right)^+ \right] \cdot \frac{d^2\mathbf{g}'}{d\mathbf{x}'d\mathbf{x}}(\tilde{\mathbf{s}}) \right\} = \end{aligned} \quad (125)$$

$$= \frac{d\mathbf{g}'}{d\mathbf{x}}(\tilde{\mathbf{s}}) + \alpha \cdot \frac{d^2\mathcal{L}}{d\mathbf{x}'d\mathbf{x}}(\tilde{\mathbf{s}}). \quad (126)$$

Здесь функция и вектор-множитель Лагранжа имеют вид:

$$\mathcal{L} = \mathcal{L}(\mathbf{x}, \lambda) = y(\mathbf{x}) - \lambda \cdot \mathbf{g}'(\mathbf{x}); \quad \lambda = \frac{d\mathbf{y}}{d\mathbf{x}}(\tilde{\mathbf{s}}) \cdot \left(\frac{d\mathbf{g}'}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \right)^+ \quad (127)$$

Но в отличие от ранее полученной формулы (122) тут не требуется вычисления квазиобратной матрицы Мура — Пенроуза от *сингулярной* $n \times n$ -матрицы Якоби (!), так как:

$$\frac{d\mathbf{g}'}{d\mathbf{x}} = \left(\frac{d\mathbf{h}}{d\mathbf{x}} \right)' \cdot \frac{d\mathbf{h}}{d\mathbf{x}} = \left(\frac{d\mathbf{g}'}{d\mathbf{x}} \right)', \quad (128)$$

$$\left(\frac{d\mathbf{g}'}{d\mathbf{x}} \right)^+ = \left[\left(\frac{d\mathbf{h}}{d\mathbf{x}} \right)' \cdot \frac{d\mathbf{h}}{d\mathbf{x}} \right]^+ = \left(\frac{d\mathbf{h}}{d\mathbf{x}} \right)_R^{-1} \cdot \left[\left(\frac{d\mathbf{h}}{d\mathbf{x}} \right)_R^{-1} \right]' \quad (129)$$

При линейном уравнении связи также имеем упрощение:

$$\frac{d^2\Phi_\alpha}{d\mathbf{x}'d\mathbf{x}}(\tilde{\mathbf{s}}) = \frac{d\mathbf{g}'}{d\mathbf{x}}(\tilde{\mathbf{s}}) + \alpha \cdot \frac{d^2y}{d\mathbf{x}'d\mathbf{x}}(\tilde{\mathbf{s}}) = \left(\frac{d\mathbf{h}}{d\mathbf{x}}(\tilde{\mathbf{s}}) \right)' \cdot \frac{d\mathbf{h}}{d\mathbf{x}}(\tilde{\mathbf{s}}) + \alpha \cdot \frac{d^2y}{d\mathbf{x}'d\mathbf{x}}(\tilde{\mathbf{s}}). \quad (130)$$

Доказательство инвариантности изложенного *общего* предельного метода решения задачи на условный экстремум $\mathbf{y}(\mathbf{x})$ к линейным модальным преобразованиям переменной $\mathbf{x}$ и вектор-функции $\mathbf{h}(\mathbf{x})$ предоставляется читателю в качестве упражнения. Причём предельное решение целесообразно рассматривать как решение методом Лагранжа с бесконечно большим скалярным множителем N .

* * *

Пример. Найти нормальное решение $\mathbf{s}^\bullet$ на области вырожденного минимума скалярной функции 2-го порядка от $\mathbf{x}$:

$$q(\mathbf{x}) = (G\mathbf{x} + \mathbf{g}_0')' \cdot (G\mathbf{x} + \mathbf{g}_0') = \mathbf{x}' \cdot G^2 \cdot \mathbf{x} + 2 \cdot \mathbf{x}' \cdot \mathbf{g}_0' + \mathbf{g}_0 \cdot \mathbf{g}_0' \quad (\det G = 0).$$

Данная задача решается аналитическим путём *точным образом* любым из 4-х вышеописанных предельных методов (общих и частных). Например, по методу регуляризации Тихонова имеем:

$$T_\alpha(\mathbf{x}) = q(\mathbf{x}) + \alpha \cdot \theta(\mathbf{x}) = q(\mathbf{x}) + \alpha \cdot \mathbf{x}' \cdot \mathbf{x},$$

$$1/2 \cdot \left(\frac{dT_\alpha}{d\mathbf{x}} \right)' = \left(G^2\mathbf{x} + \mathbf{g}_0' \right) + \alpha \cdot \mathbf{x} = \left(G^2 + \alpha \cdot I \right) \cdot \mathbf{x} + \mathbf{g}_0',$$

$$\frac{dT_\alpha}{d\mathbf{x}} = \mathbf{0} \quad (\alpha \rightarrow 0) \rightarrow \mathbf{s}^\bullet = - \lim_{\alpha \rightarrow 0} \left(G^2 + \alpha \cdot I \right)^{-1} \cdot \mathbf{g}_0'.$$

Точное вычисление этого предела будет рассмотрено в § 4.5.

Изложенные предельные методы условной оптимизации целевых функций с большим и с малым параметром реализуются на практике пока только в частных модификациях — в методе штрафной функции и в методе регуляризации. Причём они применяются именно в своих *численных формах* (см. §§ 5.11 и 5.12.3). Выше они были изучены как базовые *аналитические* методы, что позволило установить их истинную природу и дать им достаточно строгое теоретическое обоснование.

§ 2.5. Условное характеристическое (вековое) уравнение

Предельный метод решения задачи на ограничительный условный экстремум позволяет далее вывести в явном виде **характеристическое (вековое) уравнение для условных собственных значений** матрицы Гессе целевой функции в обычной неопределённой форме. Разумеется, оно, как и весь предыдущий анализ стационарности, относится к её уровню $p = 2$. Вначале введём ряд вспомогательных обозначений. Обратим внимание на то, что предельные матрицы (122, 123) и (125, 126) формально есть композиция из $n \times n$ -матрицы Якоби ранга $r = n - q$ и *полуусловной* $n \times n$ -матрицы Гессе ранга $\leq n$ с множителем $\alpha \rightarrow +0$. Обозначим эту предельную матрицу и её компоненты-слагаемые как $Q_\alpha = B + \alpha \cdot S$, где $\text{rang } B = r = n - q$, $S = S'$, $\text{rang } S \leq n$; причём B удовлетворяет, например, условию (102).

Характеристическое (вековое) уравнение матрицы Q_α относительно её собственных значений μ_i имеет вид:

$$\det(Q_\alpha - \mu I) = (-\mu)^n + q_1(\alpha) \cdot (-\mu)^{n-1} + \dots + q_{n-1}(\alpha) \cdot (-\mu) + q_n(\alpha) = 0. \quad (131)$$

Соответственно в этой форме оно имеет скалярные характеристические коэффициенты при степенях $(-\mu)$ вида:

$$\left. \begin{aligned} q_1(\alpha) &= k(Q_\alpha, 1) = \text{tr } Q_\alpha, \\ q_2(\alpha) &= k(Q_\alpha, 2), \\ &\dots\dots\dots \\ q_i(\alpha) &= k(Q_\alpha, i), \\ &\dots\dots\dots \\ q_n(\alpha) &= k(Q_\alpha, n) = \det Q_\alpha. \end{aligned} \right\}$$

Для матриц B и S аналогичные характеристические коэффициенты порядка i обозначаются далее как b_i и s_i .

Применим классическую формулу для детерминанта суммы двух матриц (в данном случае это B и $\alpha \cdot S$). При этом учтём, что в интервале основного индекса $r < i \leq n$: $b_i = 0$ (как и все миноры B порядка $> r$), но $b_r \neq 0$, так как $\text{rang } B = r = n - q$. Кроме того, вспомним, что сами характеристические коэффициенты порядка i представляют собой сумму детерминантов всевозможных диагональных (главных) миноров $n \times n$ -матрицы размера $i \times i$. В итоге получаем следующие разложения характеристических коэффициентов матрицы Q_α по степеням α :

$$\left. \begin{aligned}
 q_1(\alpha) &= b_1 + \alpha \cdot s_1, \\
 q_2(\alpha) &= b_2 + \alpha \cdot q_{2(1)} + \alpha^2 \cdot s_2, \\
 q_3(\alpha) &= b_3 + \alpha \cdot q_{3(2)} + \alpha^2 \cdot q_{3(1)} + \alpha^3 \cdot s_3, \\
 &\dots\dots\dots \\
 q_{r-1}(\alpha) &= b_{r-1} + \alpha \cdot q_{r-1(r-2)} + \dots + \alpha^{r-2} \cdot q_{r-1(1)} + \alpha^{r-1} \cdot s_{r-1}, \\
 q_r(\alpha) &= b_r + \alpha \cdot q_{r(r-1)} + \dots + \alpha^{r-1} \cdot q_{r(1)} + \alpha^r \cdot s_r, \\
 q_{r+1}(\alpha) &= \alpha \cdot q_{r+1(r)} + \dots + \alpha^r \cdot q_{r+1(1)} + \alpha^{r+1} \cdot s_{r+1}, \\
 &\dots\dots\dots \\
 q_{n-1}(\alpha) &= \alpha^{n-r-1} \cdot q_{n-1(r)} + \dots + \alpha^{n-2} \cdot q_{n-1(1)} + \alpha^{n-1} \cdot s_{n-1}, \\
 q_n(\alpha) &= \alpha^{n-r} \cdot q_{n(r)} + \dots + \alpha^{n-1} \cdot q_{n(1)} + \alpha^n \cdot s_n.
 \end{aligned} \right\} \quad (132)$$

Здесь у коэффициента $q_{i(j)}$ первый индекс i означает его порядок, а второй индекс j означает: сколько строк (или столбцов) матрицы B заменили те же строки (или столбцы) матрицы S при формировании *составных* матриц из B и S для вычисления детерминанта их суммы.

Из (132) следует, что все характеристические коэффициенты $q_i(\alpha)$ порядка $i > r$ пропорциональны параметру $\alpha \rightarrow 0$. В связи с этим преобразуем вековое уравнение (131) к следующему виду:

$$\begin{aligned}
 &(-\mu)^{n-r} \cdot \left[(-\mu)^r + q_1(\alpha) \cdot (-\mu)^{r-1} + \dots + q_{r-1}(\alpha) \cdot (-\mu) + q_r(\alpha) \right] + \\
 &+ q_{r+1}(\alpha) \cdot (-\mu)^{n-r-1} + \dots + q_{n-1}(\alpha) \cdot (-\mu) + q_n(\alpha) = 0.
 \end{aligned}$$

Отсюда видно, что при $\alpha \rightarrow 0$ вековое уравнение даёт $q = n - r$ нулевых и r ненулевых собственных значений. Последние при $\alpha \rightarrow 0$ совпадают с таковыми значениями для матрицы Якоби B . Но интересующие нас *условные* собственные значения матрицы Гессе S пока для Q_α суть бесконечно малые величины из-за пропорциональности параметру α . Поэтому сделаем дополнительное преобразование: $\mu_i = \alpha \cdot v_i$. Вековое уравнение тогда приобретает следующий вид:

$$\begin{aligned}
 &(-\alpha v)^n + q_1(\alpha) \cdot (-\alpha v)^{n-1} + \dots + q_{r-1}(\alpha) \cdot (-\alpha v)^{n-r+1} + \\
 &+ \left[q_r(\alpha) \cdot (-\alpha v)^{n-r} + q_{r+1}(\alpha) \cdot (-\alpha v)^{n-r-1} + \dots \right. \\
 &\left. \dots + q_{n-1}(\alpha) \cdot (-\alpha v) + q_n(\alpha) \right] = 0.
 \end{aligned} \quad (133)$$

С учётом указанных в (132) разложений $q_i(\alpha)$ по степеням α , видно, что первые r членов уравнения содержат α в степени $\geq n - r + 1$, а остальные $n - r + 1$ членов уравнения содержат α в степени $\leq n - r$.

Поэтому вековое уравнение в форме (133) даёт именно $г$ бесконечных собственных значений матрицы B/α и $q = n - г$ условных собственных значений v_i матрицы Гессе $y(\mathbf{x})$. Для вычисления v_i в пределе при $\alpha \rightarrow 0$ членами уравнения (133), стоящими вне его квадратных скобок, можно пренебречь. Теперь получаем предельное уравнение для v_i типа:

$$\alpha^{n-r} \cdot q_r(\alpha) \cdot (-v)^{n-r} + \alpha^{n-r-1} \cdot q_{r+1}(\alpha) \cdot (-v)^{n-r-1} + \dots \\ \dots + \alpha \cdot q_{n-1}(\alpha) \cdot (-v) + q_n(\alpha) = 0. \quad (134)$$

С целью его дальнейшего упрощения в разложениях $q(\alpha)$ из (132) пренебрегаем членами, содержащими параметр α в степени $> i - r$, т. е. так, чтобы в уравнении (134) все остающиеся члены содержали только максимальную степень α^{n-r} . Это допускается в пределе при $\alpha \rightarrow 0$. Далее постоянный множитель α^{n-r} в оставшихся членах попросту сокращается, что даёт, наконец, *непредельное* вековое уравнение для v_i :

$$q_r(r) \cdot (-v)^{n-r} + q_{r+1}(r) \cdot (-v)^{n-r-1} + \dots + q_{n-1}(r) \cdot (-v) + q_n(r) = 0. \quad (135)$$

Здесь коэффициенты $q_{i(r)}$, взятые последовательно при i от $г$ до n , уже не зависят от α (причём $q_{r(r)} = b_r$). Они представляют собой суммы детерминантов всевозможных диагональных миноров порядка i от $г$ до n от составных матриц из B и S , в которых присутствуют строки (или столбцы) матрицы B и $(i - r \geq 0)$ строк (или столбцов) матрицы S .

При $г = 0$, разумеется, (135) принимает вид классического векового уравнения. При $г > 0$ уравнение даёт $q = n - г$ условных собственных значений матрицы Гессе $y(\mathbf{x})$. (Среди них могут быть и одинаковые, и нулевые.) Отметим, что в применяемой здесь *знакопеременной форме* векового уравнения, т. е. с аргументом $(-v)$, положительным v_i отвечает условный минимум, отрицательным v_i отвечает условный максимум, разнознаковым v_i отвечает условная стационарная седловина функции $y(\mathbf{x})$ — все 2-го уровня. Однако для идентификации характера условной стационарности целевой функции $y(\mathbf{x})$ требуются всего лишь знаки v_i . Этот анализ выполняется по значениям только коэффициентов векового уравнения исходя из признаков его знакоопределённости (см. § 1.9).

Аналогично при линейном уравнении связи с $n \times n$ -матрицей Якоби $d\mathbf{h}/d\mathbf{x} = B = \text{Const}$ ранга $г$, а также при линейном уравнении связи с $n \times n$ -матрицей Якоби $d\mathbf{g}'/d\mathbf{x} = (d\mathbf{h}/d\mathbf{x})' \cdot d\mathbf{h}/d\mathbf{x} = A'A = B = \text{Const}$, где A — $m \times n$ -матрица ранга m , предельная матрица Гессе выражается в упрощённом виде как: $Q_\alpha = B + \alpha \cdot G$, где $G = d^2y/d\mathbf{x}'d\mathbf{x}(\mathbf{\xi}) = G'$.

Глава 3. Аналитическая оптимизация целевых функций от разнообразных комплексных переменных

§ 3.1. Два альтернативных варианта комплексификации

В силу природы комплексных чисел теоретически реализуются два принципиально различных подхода к математическим операциям, как с ними, так и с разнообразными объектами на их основе. Эти операции определяют сущность комплексификации.

Адекватный подход заключается в том, что комплексные числа и объекты на их основе подвергают формально тем же аналитическим операциям, включая простое транспонирование, что и вещественные аналоги. Адекватный вариант комплексификации даёт возможность использовать все результаты, полученные ранее для вещественных чисел и числовых объектов. Исключением здесь являются отношения типа неравенств, конечно, не для заведомо вещественных параметров. Отдельный важный субвариант отвечает *псевдоизации*, когда элементы комплексных чисел и объектов — только чисто вещественные и чисто мнимые. Соответственно в её рамках допускаются те математические преобразования, которые не изменяют этого свойства. Псевдоизация имеет место в бинарных комплексных евклидовых пространствах.

Симбиозный подход, помимо аналитических операций, использует для комплексных чисел и ряда числовых объектов несводимую к ним операцию комплексного сопряжения, включая содержащее её эрмитово транспонирование. В частности, как хорошо известно, эрмитов подход к комплексному векторному и матричному исчислению сопровождает каждую операцию транспонирования ещё комплексным сопряжением. Эрмитова комплексификация реализуется, например, для одно- и для двухвалентных тензорных объектов (векторов и двумерных матриц). Она даёт возможность использовать в самосопряжённой форме понятия *нормы* объектов (т. е. вещественного положительного модуля), а также сохранить для них в аналогичной форме отношения типа неравенств.

Эти альтернативные варианты определяют, по сути, два различных пути развития теорий и их приложений в комплексных пространствах.

Например, соотношение $B \cdot B' = B' \cdot B$ задаёт адекватно нормальные матрицы, а его аналог $B \cdot B^* = B^* \cdot B$ — эрмитово нормальные матрицы; соотношение $\langle \text{im } B \rangle \equiv \langle \text{im } B' \rangle$ задаёт адекватно нуль-нормальные матрицы, а его аналог $\langle \text{im } B \rangle \equiv \langle \text{im } B^* \rangle$ — эрмитово нуль-нормальные матрицы. Принципиально различно в E^n определяются комплексные адекватно и эрмитово симметричные проекторы и отвечающие им квазиобратные матрицы (ортогональные в евклидовом и в эрмитовом пространствах). С другой стороны, комплексная обратная матрица B^{-1} обязательно однозначная, так как для несингулярной матрицы B всегда имеет место тождество $\langle \text{im } B \rangle \equiv \langle \text{im } B' \rangle \equiv \langle \text{im } B^* \rangle$. Параллельность линейных объектов, как известно, есть аффинное понятие. Именно поэтому она не зависит от выбора варианта комплексификации.

К известным трём скалярным формам представления комплексного числа (алгебраическая, тригонометрическая и показательная) и к его псевдоизированной 2×1 -векторной форме представления можно ещё добавить 2×2 -матричную форму, не содержащую мнимой единицы:

$$\left. \begin{aligned} W(a), \quad a = p + i \cdot q : \\ \left(\begin{array}{cc} p, & -q \\ q, & p \end{array} \right) = \rho \cdot \left(\begin{array}{cc} \cos \varphi, & -\sin \varphi \\ \sin \varphi, & \cos \varphi \end{array} \right) = S + K, \\ \\ W'(a) = W(\bar{a}), \quad \bar{a} = p - i \cdot q : \\ \left(\begin{array}{cc} p, & q \\ -q, & p \end{array} \right) = \rho \cdot \left(\begin{array}{cc} \cos \varphi, & \sin \varphi \\ -\sin \varphi, & \cos \varphi \end{array} \right) = S - K, \end{aligned} \right\} \quad (136)$$

где $S = S'$, $K = -K'$, $S \cdot K = K \cdot S$. Эта форма представляет комплексные числа «а» и « $\bar{a}$ » геометрически в вещественном декартовом базисе евклидова пространства. Данное вещественное представление, как и три скалярные формы, коммутативно и удовлетворяет всем формулам и тождествам для комплексных чисел. Причём последние образуют вещественные транспонированные пары из этих 2×2 -матриц — аналоги комплексных сопряжённых пар чисел a и $\bar{a}$. Нормальная $n \times n$ -матрица представляет геометрически в некотором декартовом базисе $k \leq [n/2]$ комплексных чисел и $(n - 2k)$ вещественных чисел: $M = R \cdot W \cdot R'$. Простая вещественная матрица представляет те же числа в некотором аффинном базисе: $P = V \cdot W \cdot V^{-1}$. Тут W — каноническая вещественная форма простой (в том числе нормальной) матрицы, включающая в прямой сумме только 1×1 - и 2×2 -клетки. Она же, с точностью до перестановок клеток, есть простейшее вещественное решение векового уравнения этой матрицы: $c(\mu) = 0$.

Примеры адекватной комплексификации (аналогии)

Характерными примерами адекватной комплексификации являются готовые формулы решений алгебраических уравнений с комплексными коэффициентами; комплексные аналитические функции и тождества, комплексные дифференциалы, производные и интегралы. В аффинном комплексном пространстве, метризуемом по адекватному варианту, определяются комплексные меры для длины, расстояния и угла, хотя в псевдоевклидовом пространстве реализуются (кроме нулевой) лишь только чисто вещественные и чисто мнимые меры. Евклидова геометрия и тригонометрия также имеют адекватные аналоги и псевдоаналоги [27]. Рассмотрим применение адекватной комплексификации, например, в теории аналитических функций от комплексного аргумента.

Пусть $\mathbf{z} = \mathbf{u} + i \cdot \mathbf{v}$, где $\mathbf{z}$, $\mathbf{u}$ и $\mathbf{v}$ суть $n \times 1$ -вектор-аргументы в одном комплексном и двух вещественных n -мерных аффинных координатных пространствах; $y(\mathbf{z}) = y_1(\mathbf{u}, \mathbf{v}) + i \cdot y_2(\mathbf{u}, \mathbf{v})$ — *скалярная комплексная аналитическая* функция от $\mathbf{z}$. Дифференцирование и интегрирование в аффинном пространстве по $n \times 1$ -вектор-аргументу осуществляют тоже в аффинных координатах. Вещественные и комплексные адекватные аналоги имеют место для дифференциалов, производных и интегралов по $\mathbf{z}$. Отсюда естественно получаются *частные* дифференциальные и интегральные характеристики и устанавливается их взаимосвязь:

$$dy = dy_1(\mathbf{u}, \mathbf{v}) + i \cdot dy_2(\mathbf{u}, \mathbf{v}) = \mathbf{h}'(\mathbf{z})d\mathbf{z} = [\mathbf{h}'_1(\mathbf{u}, \mathbf{v}) + i \cdot \mathbf{h}'_2(\mathbf{u}, \mathbf{v})](d\mathbf{u} + i \cdot d\mathbf{v}) =$$

$$= [\mathbf{h}'_1(\mathbf{u}, \mathbf{v})d\mathbf{u} - \mathbf{h}'_2(\mathbf{u}, \mathbf{v})d\mathbf{v}] + i \cdot [\mathbf{h}'_1(\mathbf{u}, \mathbf{v})d\mathbf{v} + \mathbf{h}'_2(\mathbf{u}, \mathbf{v})d\mathbf{u}] \quad \leftrightarrow$$

$$\left. \begin{aligned} \mathbf{h}'_1(\mathbf{u}, \mathbf{v}) &= \frac{\partial y_1}{\partial \mathbf{u}} = \frac{\partial y_2}{\partial \mathbf{v}}, \\ \mathbf{h}'_2(\mathbf{u}, \mathbf{v}) &= -\frac{\partial y_1}{\partial \mathbf{v}} = \frac{\partial y_2}{\partial \mathbf{u}}. \end{aligned} \right\}$$

Это суть **уравнения Даламбера — Эйлера** в *векторной* форме (при $n > 1$) для *дифференцируемой* $y(\mathbf{z})$, но необязательно с полным дифференциалом. Если $dy(\mathbf{z}) = \mathbf{h}'(\mathbf{z})d\mathbf{z}$ — полный дифференциал, то $\mathbf{h}(\mathbf{z})$ — *градиентная* функция, т. е. $d\mathbf{h}'/d\mathbf{z} = (d\mathbf{h}'/d\mathbf{z})' = d^2y/d\mathbf{z}'d\mathbf{z}$ — комплексная матрица Гессе от y , и обратно. Применим ту же схему для *градиентной* аналитической вектор-функции $\mathbf{h}(\mathbf{z}) = \mathbf{h}_1(\mathbf{u}, \mathbf{v}) + i \cdot \mathbf{h}_2(\mathbf{u}, \mathbf{v})$:

$$\left. \begin{aligned} \frac{\partial \mathbf{h}_1}{\partial \mathbf{u}} = \frac{\partial \mathbf{h}_2}{\partial \mathbf{v}} = \frac{\partial^2 y_1}{\partial \mathbf{u}' \partial \mathbf{u}} = -\frac{\partial^2 y_1}{\partial \mathbf{v}' \partial \mathbf{v}} = \frac{\partial^2 y_2}{\partial \mathbf{u}' \partial \mathbf{v}} = \frac{\partial^2 y_2}{\partial \mathbf{v}' \partial \mathbf{u}} = \left(\frac{\partial \mathbf{h}_1}{\partial \mathbf{u}} \right)' \\ \frac{\partial \mathbf{h}_1}{\partial \mathbf{v}} = -\frac{\partial \mathbf{h}_2}{\partial \mathbf{u}} = \frac{\partial^2 y_2}{\partial \mathbf{v}' \partial \mathbf{v}} = -\frac{\partial^2 y_2}{\partial \mathbf{u}' \partial \mathbf{u}} = \frac{\partial^2 y_1}{\partial \mathbf{v}' \partial \mathbf{u}} = \frac{\partial^2 y_1}{\partial \mathbf{u}' \partial \mathbf{v}} = \left(\frac{\partial \mathbf{h}_1}{\partial \mathbf{v}} \right)' \end{aligned} \right\} \quad (138)$$

Первые два члена в цепочках этих равенств составляют **уравнения Даламбера—Эйлера** в *матричной* форме для дифференцируемой $\mathbf{h}(\mathbf{z})$. Наряду с очевидной здесь симметричностью частных матриц Якоби (ввиду симметричности матриц Гессе), они формулируют также условия полноты дифференциалов выше в квадратных скобках! В частности, из (138) для двух компонент исходной аналитической функции $y_1(\mathbf{u}, \mathbf{v})$ и $y_2(\mathbf{u}, \mathbf{v})$ от вещественных аргументов $\mathbf{u}$ и $\mathbf{v}$ одновременно следуют два **уравнения Лапласа** в *матричной* форме:

$$\frac{\partial^2 y_1}{\partial \mathbf{u}' \partial \mathbf{u}} + \frac{\partial^2 y_1}{\partial \mathbf{v}' \partial \mathbf{v}} = Z = \frac{\partial^2 y_2}{\partial \mathbf{u}' \partial \mathbf{u}} + \frac{\partial^2 y_2}{\partial \mathbf{v}' \partial \mathbf{v}}.$$

Следовательно, y_1 и y_2 , связанные тут дополнительно уравнениями Даламбера—Эйлера (137-138), — *сопряжённые гармонические функции*. Если $\mathbf{h}(\mathbf{z}) = \mathbf{h}_1(\mathbf{u}, \mathbf{v}) + i \cdot \mathbf{h}_2(\mathbf{u}, \mathbf{v})$ — дифференцируемая, но не обязательно градиентная вектор-функция, то для её двух компонент $\mathbf{h}_1(\mathbf{u}, \mathbf{v})$ и $\mathbf{h}_2(\mathbf{u}, \mathbf{v})$ от вещественных аргументов $\mathbf{u}$ и $\mathbf{v}$, как в (138), опять-таки следуют уравнения Даламбера—Эйлера в *матричной* форме и т. д.

В бинарном комплексном евклидовом координатном пространстве (изоморфном вещественному псевдоевклидову пространству-аналогу) вышеуказанные соотношения изменяются следующим образом:

$$\begin{aligned} z = \left| \frac{\mathbf{u}}{i\mathbf{v}} \right|; \quad dy = dy_1(\mathbf{u}, \mathbf{v}) + i \cdot dy_2(\mathbf{u}, \mathbf{v}) = \mathbf{h}'(\mathbf{z})d\mathbf{z} = \left| \frac{\mathbf{p}_1 + i \cdot \mathbf{p}_2}{\mathbf{q}_1 + i \cdot \mathbf{q}_2} \right| \cdot \left| \frac{d\mathbf{u}}{i d\mathbf{v}} \right| = \\ = [\mathbf{p}'_1(\mathbf{u}, \mathbf{v})d\mathbf{u} - \mathbf{q}'_2(\mathbf{u}, \mathbf{v})d\mathbf{v}] + i \cdot [\mathbf{q}'_1(\mathbf{u}, \mathbf{v})d\mathbf{v} + \mathbf{p}'_2(\mathbf{u}, \mathbf{v})d\mathbf{u}] \quad \leftrightarrow \\ \mathbf{p}'_1(\mathbf{u}, \mathbf{v}) = \frac{\partial y_1}{\partial \mathbf{u}}, \quad \mathbf{p}'_2(\mathbf{u}, \mathbf{v}) = \frac{\partial y_2}{\partial \mathbf{u}}, \quad \mathbf{q}'_1(\mathbf{u}, \mathbf{v}) = \frac{\partial y_2}{\partial \mathbf{v}}, \quad \mathbf{q}'_2(\mathbf{u}, \mathbf{v}) = -\frac{\partial y_1}{\partial \mathbf{v}}. \quad (139) \end{aligned}$$

Это суть уравнения-псевдоаналоги в *векторной* форме для y_1 и y_2 . Применяв повторно ту же схему для *градиентной* аналитической вектор-функции $\mathbf{h}(\mathbf{z})$, получаем также, в том числе, условия полноты дифференциалов выше в квадратных скобках:

$$\left. \begin{aligned} \frac{\partial \mathbf{p}_1}{\partial \mathbf{v}} &= \frac{\partial^2 y_1}{\partial \mathbf{u}' \partial \mathbf{v}} = \left(\frac{\partial^2 y_1}{\partial \mathbf{v}' \partial \mathbf{u}} \right)' = - \left(\frac{\partial \mathbf{q}_2}{\partial \mathbf{u}} \right)', \\ \frac{\partial \mathbf{p}_2}{\partial \mathbf{v}} &= \frac{\partial^2 y_2}{\partial \mathbf{u}' \partial \mathbf{v}} = \left(\frac{\partial^2 y_2}{\partial \mathbf{v}' \partial \mathbf{u}} \right)' = \left(\frac{\partial \mathbf{q}_1}{\partial \mathbf{u}} \right)', \\ \frac{\partial \mathbf{p}_1}{\partial \mathbf{u}} &= \frac{\partial^2 y_1}{\partial \mathbf{u}' \partial \mathbf{u}} = \left(\frac{\partial \mathbf{p}_1}{\partial \mathbf{u}} \right)', \quad \frac{\partial \mathbf{q}_1}{\partial \mathbf{v}} = \frac{\partial^2 y_2}{\partial \mathbf{v}' \partial \mathbf{v}} = \left(\frac{\partial \mathbf{q}_1}{\partial \mathbf{v}} \right)', \\ \frac{\partial \mathbf{p}_2}{\partial \mathbf{u}} &= \frac{\partial^2 y_2}{\partial \mathbf{u}' \partial \mathbf{u}} = \left(\frac{\partial \mathbf{p}_2}{\partial \mathbf{u}} \right)', \quad \frac{\partial \mathbf{q}_2}{\partial \mathbf{v}} = - \frac{\partial^2 y_1}{\partial \mathbf{v}' \partial \mathbf{v}} = \left(\frac{\partial \mathbf{q}_2}{\partial \mathbf{v}} \right)' \end{aligned} \right\} \quad (140)$$

Отметим, что в бинарном комплексном евклидовом пространстве уже нет свойства *гармоничности* для пары скалярных комплексных функций $y_1(\mathbf{u}, \mathbf{v})$ и $y_2(\mathbf{u}, \mathbf{v})$, составляющих аналитическую функцию $y(\mathbf{z})$.

Примеры симбиозной и эрмитовой комплексификации (аналогии)

В эрмитовых пространствах, как хорошо известно, определяются вещественные положительно определённые меры: нормы для длины, расстояния, скалярного и тензорного угла, ортогональные эрмитово симметричные проекторы и квазиобратная матрица Мура — Пенроуза. Есть эрмитовы аналоги евклидовой геометрии и тригонометрии [27]. Ряд алгебро-геометрических неравенств, имеющих прямое отношение к определению мер и норм, тоже имеют свои особые эрмитовы аналоги.

В анализе для аналитических функций от комплексной переменной хорошо известен *принцип максимума модуля* [26]. Обоснование этого принципа, в том числе для *многомерных* комплексных функций от *многомерной* комплексной переменной, будет изложено строго и весьма наглядно в этой главе с применением операций *формального* анализа.

Наиболее общий, по сравнению с эрмитовым, симбиозный подход в применении к основным операциям анализа в бинарном комплексном пространстве приводит к симбиозным аналогам. Это особые правила для *формального* дифференцирования и интегрирования функций от сопряжённых аргументов $\mathbf{x}$ и $\bar{\mathbf{x}}$; особые требования для их *формальной* дифференцируемости или аналитичности, а также особые требования к дифференциальным выражениям для их *формальной* интегрируемости (полноты дифференциала). Отсюда логично получаются отличающиеся простотой и наглядностью симбиозные аналоги классических методов решения задач на безусловный и условный экстремумы вещественных скалярных функций от ряда комплексных сопряжённых переменных.

Эти функции обязательно симметричные по отношению к сопряжённым переменным. По сути же, в последующих параграфах данной главы предпринято необходимое для большей полноты комплексного анализа развитие известной идеи формальных производных (см., например, [8]) для *неголоморфных*, в том числе, что здесь особо важно, *вещественных* функций от комплексных сопряжённых переменных. Для иллюстрации таковых в данной монографии можно указать конкретные примеры: квадрат эрмитова модуля для (псевдо)нормального решения и невязки комплексного линейного уравнения типа $A\mathbf{x} = \mathbf{a}$ — *модульные* функции; коэффициенты вещественного алгебраического, в том числе векового (характеристического) уравнения с парами комплексных сопряжённых корней — *немодульные* функции от корней уравнения.

§ 3.2. Формальное дифференцирование и интегрирование в комплексном пространстве

Решение задачи на экстремум целевой вещественной функции от комплексной переменной (как *одномерной*, так и *многомерной*) можно осуществить, в принципе, либо значительно упростить, если прибегнуть к так называемому *формальному* дифференцированию. Формальные производные в литературе известны (например, [9, 26]). Но достаточно развитый аппарат для формального анализа в литературе отсутствует. В изучаемом аспекте необходимость в формальном анализе вызвана тем, что целевая *вещественная* неголоморфная функция по природе есть симметричная функция от двух формально аналитически независимых друг от друга переменных x и $\bar{x}$, т. е. как $y = y(x, \bar{x}) = y(\bar{x}, x) = \bar{y}$. Эти переменные преобразуются друг в друга операцией комплексного сопряжения, которая не является аналитической:

$$x = u + i \cdot v \leftrightarrow \bar{x} = u - i \cdot v.$$

Объединив исходные *одномерные* аргументы x и $\bar{x}$ в *бинарный* 2×1 -вектор-аргумент $\mathbf{w} = (x, \bar{x})$, далее преобразуем вещественную функцию к универсальной форме $y = \psi(\mathbf{w})$. Такая форма в основных операциях анализа преобразуется относительно бинарного аргумента $\mathbf{w}$ обычным образом, как от некоего векторного аргумента. Она применяется только для симметричных, т. е. *вещественных* функций от x и $\bar{x}$.

Строго говоря, x и $\bar{x}$ полностью аналитически независимы друг от друга тогда и только тогда, когда их вещественная и мнимая части аналитически независимы. Последнее не выполняется в задачах на условный экстремум целевой скалярной функции $y = y(x, \bar{x}) = y(x, \bar{x})$ с *внутренне* ограниченной комплексной переменной.

В задачах такого рода нужное ограничение задаётся через уравнение связи между x и $\bar{x}$ посредством также *вещественной* симметричной функции ограничения типа $h = h(x, \bar{x}) = \bar{h}(x, \bar{x}) = 0$ — при *одномерных* комплексных сопряжённых переменных x и $\bar{x}$. В частности, подобное внутреннее ограничение и соответствующий ему условный экстремум могут иметь место для целевых функций от одномерной комплексной переменной! (Разумеется, неголоморфные функции y от x и $\bar{x}$ могут быть и комплексными, но тогда они уже несимметричны относительно этих переменных: $y = y(x, \bar{x}) \leftrightarrow \bar{y} = \bar{y}(\bar{x}, x)$.)

Пусть $\langle \mathcal{T} \rangle \equiv \langle \bar{\mathcal{T}} \rangle \subset \mathcal{E}$ — некий бинарный интервал на комплексной числовой оси $\mathcal{E}$; $(x_0, \bar{x}_0)$ — некая внутренняя бинарная точка на этом интервале. Определим, что *комплексная функция* $y = y(x, \bar{x})$ *формально дифференцируема (непрерывно) в бинарной точке* $(x_0, \bar{x}_0)$, *если в ней существуют формальные производные (непрерывные)* $\partial y / \partial x$ и $\partial y / \partial \bar{x}$. Обратим внимание здесь на то, что для *формальных дифференциалов* аргумента и функции в отличие от тех же обычных дифференциальных характеристик при их обозначениях применяется *курсив*!

Далее совершенно естественным образом определяется формальная дифференцируемость комплексной функции $y = y(x, \bar{x})$ на некоторой бинарной окрестности точки $(x_0, \bar{x}_0)$, на заданном бинарном интервале $\langle \mathcal{T} \rangle$ и на комплексной числовой оси $\mathcal{E}$ в целом (когда это возможно).

Комплексная функция $y = y(x, \bar{x})$, формально дифференцируемая на некоторой бинарной окрестности точки $(x_0, \bar{x}_0)$ и представимая на ней сходящимся двукратным степенным рядом от x и $\bar{x}$, является по определению *формально аналитической* в бинарной точке $(x_0, \bar{x}_0)$.

В частности, функция y дифференцируема по x в обычном смысле, т. е. по Даламберу — Эйлеру [9, 26], тогда и только тогда, когда $\partial y / \partial x$ существует, а $\partial y / \partial \bar{x} = 0$. Все указанные определения и свойства легко распространяются на функции многомерных переменных $y = y(\mathbf{x}, \bar{\mathbf{x}})$.

Для *вещественной* скалярной функции $y = \bar{y} = y(x, \bar{x}) \equiv \psi(\mathbf{w})$, дважды формально непрерывно дифференцируемой, имеем полные и частные дифференциалы и производные 1-го и 2-го порядков:

$$dy = \frac{dy}{d\mathbf{w}} d\mathbf{w} \equiv \frac{\partial y}{\partial x} dx + \frac{\partial y}{\partial \bar{x}} d\bar{x}, \quad (141, 142)$$

$$d^2 y = d\mathbf{w}' \frac{d^2 y}{d\mathbf{w}' d\mathbf{w}} d\mathbf{w} \equiv \frac{\partial^2 y}{\partial x^2} dx^2 + 2 \frac{\partial^2 y}{\partial \bar{x} \partial x} d\bar{x} dx + \frac{\partial^2 y}{\partial \bar{x}^2} d\bar{x}^2. \quad (143, 144)$$

Здесь $dy/d\mathbf{w}$ и $d^2y/d\mathbf{w}'d\mathbf{w}$ — градиент и матрица Гессе для функции y по бинарному аргументу $\mathbf{w} = (x, \bar{x})$. Причём 2×1 -вектор-переменная $\mathbf{w}$, как было оговорено, есть формальная прямая сумма скаляров x и $\bar{x}$. Матрица Гессе от y по аргументу $\mathbf{w}$, как ей и положено быть вообще, квадратная и симметричная: $d^2y/d\mathbf{w}'d\mathbf{w} = (d^2y/d\mathbf{w}'d\mathbf{w})'$. Но она и градиент в (141) и (143) суть комплексные характеристики!

В формуле (144) скалярная смешанная 2-я частная производная не изменяется при перестановке дифференциалов ∂x и $\partial \bar{x}$, равно как и их произведение. Поэтому здесь, как обычно, удвоен один из смешанных дифференциалов, равный другому. Тут не имеет значения: функция y вещественная или комплексная. Лишь бы она была *одномерная*!

Для вещественной функции $y = y(x, \bar{x}) = y(\bar{x}, x) = \bar{y}$, очевидно, однородные производные комплексно сопряжены по отношению друг к другу, а смешанная производная обязательно вещественная:

$$\frac{\partial y}{\partial x} = \overline{\left(\frac{\partial y}{\partial \bar{x}}\right)}, \quad \frac{\partial^2 y}{\partial x^2} = \overline{\left(\frac{\partial^2 y}{\partial \bar{x}^2}\right)}, \quad \frac{\partial^2 y}{\partial \bar{x} \partial x} = \frac{\partial^2 y}{\partial x \partial \bar{x}} = \overline{\left(\frac{\partial^2 y}{\partial \bar{x} \partial x}\right)}. \quad (145—147)$$

Отметим сразу же одну существенную особенность комплексного анализа (в отличие от вещественного прототипа). Любые независимые *одномерные* комплексные переменные — здесь это x и $\bar{x}$ или dx и $d\bar{x}$ обладают *двумя* степенями свободы, так как комплексная числовая ось для их отображения геометрически, по сути, является комплексной плоскостью с её единичным базисом $(1, i)$. Поэтому, например, нельзя говорить о конкретных дифференциалах функции y (первых, вторых и т. д.) из данной комплексной точки $(x_0, \bar{x}_0)$, т. е. не по конкретному *внутреннему* направлению аргумента. Они представляют собой всегда некие множества значений, получаемых при всевозможных внутренних направлениях $dx \leftrightarrow d\bar{x}$ на комплексной числовой оси (плоскости) $\mathcal{C}$. Но последняя в *скалярном* анализе формально принимается как бы за *одномерную* комплексную числовую ось! Так удобно в комплексном анализе ввиду получаемой схожести символики записей формулировок с аналогичными формулировками в вещественном анализе. Здесь это видно, например, по обозначениям дифференциалов и производных. Отметим, однако, что для итоговых дифференциальных характеристик в любом случае принципиально возможно выполнить *полный* анализ по всевозможным внутренним направлениям дифференциалов dx и $d\bar{x}$, причём даже с каким-либо задаваемым его внутренним или внешним ограничением, и далее получить окончательный результат.

Именно с подобным приёмом в последующих параграфах будут рассматриваться решения задач на экстремум, в том числе условный, для вещественной целевой функции от комплексных сопряжённых переменных x и $\bar{x}$, или от объединяющей их векторной переменной $\mathbf{w}$.

Как обратную операцию, определим формальное криволинейное интегрирование на $\mathcal{E}$ для пары производящих функций по dx и $d\bar{x}$:

$$\int_L p_1(x, \bar{x}) dx + p_2(x, \bar{x}) d\bar{x} = y(x, \bar{x}) + C. \quad (148)$$

Здесь $p_1(x, \bar{x}) = \frac{\partial y}{\partial x}$, $p_2(x, \bar{x}) = \frac{\partial y}{\partial \bar{x}}$. Этот криволинейный интеграл не зависит от пути L (т. е. под знаком данного *формального интеграла* стоит полный формальный дифференциал) тогда и только тогда, когда:

$$\frac{\partial p_1}{\partial \bar{x}} = \frac{\partial p_2}{\partial x}. \quad (149)$$

Это следует непосредственно из очевидного свойства смешанной 2-й (скалярной) частной производной, тождественного требованию (149):

$$\frac{\partial^2 y}{\partial \bar{x} \partial x} = \frac{\partial^2 y}{\partial x \partial \bar{x}} = b. \quad (150)$$

В частности, для *вещественной* функции $b = \bar{b} = g$. Итак, (148) при требовании (149) вычисляется как обычный криволинейный интеграл, не зависящий от пути интегрирования, выполняя этот путь только по одному из дифференциалов аргумента, например, по dx .

На аффинном n -мерном комплексном пространстве $\mathcal{E}^n$, задающем *многомерные* переменные x и $\bar{x}$, определяются такие же характеристики и операции с ними. Для *вещественной* функции $y = \bar{y} = y(x, \bar{x}) \equiv \psi(\mathbf{w})$, дважды формально непрерывно дифференцируемой, имеем полные и частные дифференциалы и производные 1-го и 2-го порядков:

$$dy = \frac{dy}{d\mathbf{w}} d\mathbf{w} \equiv \frac{\partial y}{\partial x} dx + \frac{\partial y}{\partial \bar{x}} d\bar{x}, \quad (151, 152)$$

$$\begin{aligned} d^2 y &= d\mathbf{w}' \frac{d^2 y}{d\mathbf{w}' d\mathbf{w}} d\mathbf{w} \equiv \\ &\equiv dx' \frac{\partial^2 y}{\partial x' \partial x} dx + 2dx' \frac{\partial^2 y}{\partial x' \partial \bar{x}} d\bar{x} + d\bar{x}' \frac{\partial^2 y}{\partial \bar{x}' \partial x} dx + d\bar{x}' \frac{\partial^2 y}{\partial \bar{x}' \partial \bar{x}} d\bar{x}. \end{aligned} \quad (153, 154)$$

Здесь $dy/d\mathbf{w}$ и $d^2y/d\mathbf{w}'d\mathbf{w} = (d^2y/d\mathbf{w}'d\mathbf{w})'$ — градиент и матрица Гессе для функции y по бинарному аргументу $\mathbf{w} = (\mathbf{x}, \bar{\mathbf{x}})$. Причём $2n \times 1$ -вектор-переменная $\mathbf{w}$ есть формальная прямая сумма векторов $\mathbf{x}$ и $\bar{\mathbf{x}}$! Градиент и матрица Гессе в (151), (153) — комплексные характеристики.

В формуле (154) фигурирующая в ней одна 2-я смешанная частная производная транспонируется в другую 2-ю смешанную производную при перестановке *формальных дифференциалов* $\partial\mathbf{x}$ и $\partial\bar{\mathbf{x}}$:

$$\frac{\partial^2 y}{\partial \bar{\mathbf{x}}' \partial \mathbf{x}} = \frac{\partial^2 y}{\partial \mathbf{x}^* \partial \mathbf{x}} = \mathbf{B}, \quad \frac{\partial^2 y}{\partial \mathbf{x}' \partial \bar{\mathbf{x}}} = \frac{\partial^2 y}{\partial \bar{\mathbf{x}}^* \partial \bar{\mathbf{x}}} = \mathbf{B}'. \quad (155, 156)$$

Причём $d\mathbf{x}^* \mathbf{B} d\mathbf{x} \equiv d\bar{\mathbf{x}}^* \mathbf{B}' d\bar{\mathbf{x}}$. Поэтому в формуле (154) аналогично тому, что в формуле (144), стоит только удвоенное первое слагаемое. Опять же не имеет значения: функция y вещественная или комплексная. Лишь бы она была *одномерная*!

Для *вещественной* функции $y = y(\mathbf{x}, \bar{\mathbf{x}}) = y(\bar{\mathbf{x}}, \mathbf{x}) = \bar{y}$ однородные производные комплексно сопряжены по отношению друг к другу, а смешанные производные (любая из двух) обязательно эрмитовы:

$$\frac{\partial y}{\partial \mathbf{x}} = \overline{\left(\frac{\partial y}{\partial \bar{\mathbf{x}}} \right)}, \quad \frac{\partial^2 y}{\partial \mathbf{x}' \partial \mathbf{x}} = \mathbf{S} = \mathbf{S}' = \overline{\left(\frac{\partial^2 y}{\partial \bar{\mathbf{x}}' \partial \bar{\mathbf{x}}} \right)}; \quad (157, 158)$$

$$\frac{\partial^2 y}{\partial \mathbf{x}^* \partial \mathbf{x}} = \left(\frac{\partial^2 y}{\partial \mathbf{x}^* \partial \mathbf{x}} \right)^* = \mathbf{G} = \mathbf{G}^*, \quad \frac{\partial^2 y}{\partial \bar{\mathbf{x}}^* \partial \bar{\mathbf{x}}} = \left(\frac{\partial^2 y}{\partial \bar{\mathbf{x}}^* \partial \bar{\mathbf{x}}} \right)^* = \mathbf{G}' = \bar{\mathbf{G}}. \quad (159, 160)$$

Формальный криволинейный скалярный интеграл на *векторном* комплексном пространстве $\mathcal{E}^n$ определяется как:

$$\int_L \mathbf{p}_1'(\mathbf{x}, \bar{\mathbf{x}}) d\mathbf{x} + \mathbf{p}_2'(\mathbf{x}, \bar{\mathbf{x}}) d\bar{\mathbf{x}} = y(\mathbf{x}, \bar{\mathbf{x}}) + C. \quad (161)$$

Здесь $\mathbf{p}_1'(\mathbf{x}, \bar{\mathbf{x}}) = \frac{\partial y}{\partial \mathbf{x}}$, $\mathbf{p}_2'(\mathbf{x}, \bar{\mathbf{x}}) = \frac{\partial y}{\partial \bar{\mathbf{x}}}$. Этот криволинейный интеграл не зависит от пути L (т. е. под знаком данного *формального интеграла* стоит полный формальный дифференциал) тогда и только тогда, когда:

$$\frac{\partial \mathbf{p}_1}{\partial \bar{\mathbf{x}}} = \left(\frac{\partial \mathbf{p}_2}{\partial \mathbf{x}} \right)'. \quad (162)$$

Это следует непосредственно из формул (155), (156) для смешанных 2-ых частных производных, которые тождественны данному требованию.

В частности, для вещественной функции u имеем: $B = \overline{B}' = G = G^*$. Для того чтобы под знаком интеграла в формуле (161) стоял полный вещественный дифференциал, необходимо и достаточно, чтобы тензор-производные в требовании (162) были эрмитовы матрицы G и G' . Итак, формальный интеграл (149) при выполнении (162) можно вычислять как обыкновенный криволинейный интеграл, не зависящий от пути интегрирования, выполняя его по любой траектории, например, только по одному из дифференциалов аргумента dx . Иначе тогда необходимо задавать конкретную траекторию интегрирования по взаимосвязанным дифференциалам аргументов dx и $d\bar{x}$.

Из вышеизложенного материала уже достаточно хорошо видно, что формальное дифференцирование и интегрирование вполне логично и естественным образом дополняет классический комплексный анализ. Основанием для этого послужила в широком смысле симбиозная, в том числе эрмитова, комплексификация исходного вещественного анализа (см. § 3.1). Вкратце в рассматриваемом аспекте это суть определения, основные признаки, свойства, характерные особенности и применение в комплексных координатных пространствах таких основных понятий, как формальные дифференцируемость и аналитичность одномерных и многомерных неголоморфных функций, формальные дифференциалы — полный и частные, формальные производные и интегралы.

В следующих параграфах этой главы будет подробно рассмотрено применение изложенного формального анализа (дифференцирования) с целью использования наиболее простого и логичного способа решения задач на экстремум вещественной целевой функции от комплексных сопряжённых переменных — одномерных и многомерных, в том числе с внешним и внутренним ограничением. Напомним, что традиционный путь для их решения заключается в переходе к аналогичной задаче, но с удвоенным количеством вещественных переменных.

Основные преимущества излагаемого нового подхода в сравнении с традиционным заключаются в следующем. Во-первых, используемые в таком подходе операции формального анализа и получаемые формулы решений мнемонически схожи с таковыми же для целевых функций от вещественных переменных. Во-вторых, размерность пространства при решении задачи определяется лишь исходной комплексной переменной (т. е. не удваивается как при традиционном подходе). И, в-третьих, как будет видно, критерий идентификации характера стационарности при таком подходе зависит только от модулей вторых производных — либо скалярных, либо тензорных, т. е. не зависит от углов их комплексного сдвига, что значительно упрощает процедуру анализа стационарности.

§ 3.3. Экстремумы для функций от пары независимых одномерных комплексных сопряжённых переменных

Для вещественной скалярной целевой функции $y = y(x, \bar{x})$ от пары одномерных переменных x и $\bar{x}$, согласно (142), (144), имеем её полные и частные формальные дифференциалы 1-го и 2-го порядков:

$$dy = \frac{\partial y}{\partial x} dx + \frac{\partial y}{\partial \bar{x}} d\bar{x} = \overline{dy}, \quad (163)$$

$$d^2y = \frac{\partial^2 y}{\partial x^2} dx^2 + 2 \frac{\partial^2 y}{\partial \bar{x} \partial x} d\bar{x} dx + \frac{\partial^2 y}{\partial \bar{x}^2} d\bar{x}^2 = \overline{d^2y}. \quad (164)$$

Из вещественности целевой функции y следуют дополнительные свойства частных производных (145)—(147). Так, из (147) видно, что 2-я смешанная производная в данном случае обязательно вещественная. Частные смешанные дифференциалы в (164) также вещественные.

Теорема 5. Вещественная функция $y = y(x, \bar{x})$ стационарна в бинарной точке $(s, \bar{s})$ тогда и только тогда, когда

$$\frac{\partial y}{\partial x}(s, \bar{s}) = 0 \leftrightarrow \frac{\partial y}{\partial \bar{x}}(s, \bar{s}) = 0. \quad (165)$$

(Причём вышеуказанные два требования равноценны.)

Доказательство. Действительно, из (163) и (145) непосредственно следует, что $dy = 0$ в точке $(s, \bar{s})$ в любых направлениях $dx \leftrightarrow d\bar{x}$ тогда и только тогда, когда в ней выполняется условие (165). ***

Далее найдём критерии для идентификации характера безусловной стационарности целевой функции y в бинарной точке $(s, \bar{s})$.

В тривиальном варианте, когда в (164) однородные 2-ые производные нулевые, о характере стационарности y можно судить весьма просто — по знаку ненулевой вещественной смешанной 2-й производной.

В более общем случае необходимо определить знак её 2-го полного дифференциала d^2y в точке $(s, \bar{s})$ во всевозможных из неё внутренних направлениях $dx \leftrightarrow d\bar{x}$. С этой целью вначале преобразуем формулу (164) к виду

$$d^2y = \left(\frac{\partial^2 y}{\partial \bar{x}^2} \cdot \frac{d\bar{x}}{dx} + 2 \frac{\partial^2 y}{\partial \bar{x} \partial x} + \frac{\partial^2 y}{\partial x^2} \cdot \frac{dx}{d\bar{x}} \right) d\bar{x} dx = \quad (166)$$

$$= \left(\frac{\partial^2 y}{\partial \bar{x}^2} \cdot \theta + 2 \frac{\partial^2 y}{\partial \bar{x} \partial x} + \frac{\partial^2 y}{\partial x^2} \cdot \bar{\theta} \right) d\bar{x} dx = f(\theta, \bar{\theta}) d\bar{x} dx, \quad (167)$$

где θ есть скалярный *директивный параметр* (по определению).

Поскольку произведение дифференциалов $d\bar{x}dx > 0$, то знак d^2y в точке стационарности $(s, \bar{s})$ определяется знаком в ней *вещественной* функции $f(\theta, \bar{\theta})$ в формуле (167), но при *внутреннем* ограничении уравнением связи $h(x, \bar{x}) = \bar{\theta} \cdot \theta - 1 = 0$. Отметим, что директивный параметр θ и его обращённый аналог θ^{-1} особо наглядно выражаются тригонометрическим способом:

$$\theta = \cos \varphi + i \cdot \sin \varphi \leftrightarrow \bar{\theta} = \cos \varphi - i \cdot \sin \varphi = \theta^{-1}. \quad (168)$$

С целью упрощения последующих выкладок примем обозначения для величин 2-ых производных в бинарной точке стационарности:

$$\partial^2 y / \partial x^2(s, \bar{s}) = s, \partial^2 y / \partial \bar{x}^2(s, \bar{s}) = \bar{s}, \partial^2 y / \partial \bar{x} \partial x(s, \bar{s}) = g = \bar{g}.$$

Учитывая это, из (166) и (167) далее имеем:

$$f(\theta, \bar{\theta}) = \bar{s} \cdot \theta + s \cdot \bar{\theta} + 2g = \bar{f}(\theta, \bar{\theta}) = f(\bar{\theta}, \theta). \quad (169)$$

Исходная задача на безусловный экстремум функции $y = y(x, \bar{x})$ сведена к вычислению условных экстремумов (максимума и минимума) функции $f(\theta, \bar{\theta})$ с уравнением связи $h = h(x, \bar{x}) = \theta \cdot \bar{\theta} - 1 = 0$. Решаем данную задачу, применяя *формальную модификацию* метода Лагранжа с множителем $\lambda = \bar{\lambda}$, так как в данном случае $h = \bar{h}$:

$$\mathcal{L} = \mathcal{L}(\theta, \bar{\theta}, \lambda) = f(\theta, \bar{\theta}) - \lambda \cdot (\bar{\theta} \cdot \theta - 1) = \bar{\mathcal{L}}(\theta, \bar{\theta}, \lambda); \quad (170)$$

$$\left. \begin{aligned} \frac{\partial \mathcal{L}}{\partial \bar{\theta}} = s - \lambda \cdot \theta = 0, \\ h(\theta, \bar{\theta}) = \bar{\theta} \cdot \theta - 1 = 0. \end{aligned} \right\} \quad (171)$$

Отсюда: $\lambda_{1,2} = \pm \sqrt{\bar{s} \cdot s} = \bar{\lambda}_{1,2}$; $\theta_{1,2} = \pm s / \sqrt{\bar{s} \cdot s} = \pm \sqrt{s / \bar{s}}$ при $s \neq 0$ (случай $s = 0$ обсуждён вначале, причём тогда $f = 2g$, т. е. однозначна);

$$\left. \begin{aligned} f_{\max} = 2 \cdot (g + \sqrt{\bar{s} \cdot s}), \quad \partial^2 \mathcal{L} / \partial \bar{\theta} \partial \theta = -\lambda_1 < 0; \\ f_{\min} = 2 \cdot (g - \sqrt{\bar{s} \cdot s}), \quad \partial^2 \mathcal{L} / \partial \bar{\theta} \partial \theta = -\lambda_2 > 0. \end{aligned} \right\} \quad (172)$$

Следовательно, в общем случае знак d^2u в точке стационарности целевой функции $y = y(x, \bar{x})$ задают знаки величин $f_{\max}$ и $f_{\min}$. В (167) значению параметра $\theta_1 = +s/\sqrt{\bar{s} \cdot s}$ отвечает максимум d^2u . Для θ_1 и s угол комплексного сдвига φ в (168) лежит в одном из 4-х квадрантов (если он точно не отвечает какой-либо из координатных осей). С другой стороны, значению параметра $\theta_2 = -\theta_1$ отвечает минимум d^2u . Для θ_2 и $-s$ угол комплексного сдвига в (168) равен $(\varphi \pm \pi)$. Соответственно при тех же значениях θ_1 и θ_2 в (166), (167) дифференциал dx имеет углы комплексного сдвига $\beta_1 = -\varphi/2$ и $\beta_2 = -(\varphi \pm \pi)/2$, дифференциал $d\bar{x}$ имеет углы комплексного сдвига $\beta_1 = \varphi/2$ и $\beta_2 = (\varphi \pm \pi)/2$. Определим данное явление как *эффект расширения* 2-го дифференциала скалярной вещественной функции $y = y(x, \bar{x})$. Эффект имеет место для ненулевого дифференциала d^2u при условии $\partial^2 y / \partial x^2(s, \bar{s}) \neq 0$, т. е. при ненулевых s и $\bar{s}$. Другими словами, 2-й дифференциал d^2u в точке $(s, \bar{s})$ имеет некоторый интервал, задаваемый его границами $f_{\max}$ и $f_{\min}$ при θ_1 и θ_2 .

Из вышеизложенного следует:

Теорема 6. Вещественная целевая функция $y = y(x, \bar{x})$ в бинарной точке своей стационарности $(s, \bar{s})$ имеет максимум 2-го уровня, если $f_{\max} < 0$; имеет минимум 2-го уровня, если $f_{\min} > 0$; имеет стационарный перегиб 2-го уровня, если $f_{\max} > 0$, $f_{\min} < 0$. Случаи $f_{\max} = 0$ (> 0) и $f_{\min} < 0$ ($= 0$) отвечают нестрогому максимуму (минимуму) того же 2-го уровня.

Следствие 1. Если $\partial^2 y / \partial x^2(s, \bar{s}) = 0 \leftrightarrow \partial^2 y / \partial \bar{x}^2(s, \bar{s}) = 0$, то характер стационарности функции y в точке $(s, \bar{s})$ 2-го уровня однозначно задаёт ненулевая производная $\partial^2 y / \partial \bar{x} \partial x(s, \bar{s})$.

Следствие 2. Если $\partial^2 y / \partial \bar{x} \partial x(s, \bar{s}) = 0$ и $\partial^2 y / \partial x^2(s, \bar{s}) \neq 0$, то функция y в точке $(s, \bar{s})$ имеет стационарный перегиб 2-го уровня.

Следствие 3. Если $\partial^2 y / \partial \bar{x} \partial x(s, \bar{s}) > 0$, то функция y в точке $(s, \bar{s})$ не может иметь максимума.

Следствие 4. Если $\partial^2 y / \partial \bar{x} \partial x(s, \bar{s}) < 0$, то функция y в точке $(s, \bar{s})$ не может иметь минимума.

Отметим, что для целевых функций $y = y(x, \bar{x})$ в отличие от $y = y(x)$, во-первых, экстремумы имеют бинарный характер (т. е. для обеих точек вместе $(s, \bar{s}) \leftrightarrow (\bar{s}, s)$, или вместе для точек $s = u + i \cdot v$ и $\bar{s} = u - i \cdot v$; исключение: $s = \bar{s} = u$) и, во-вторых, возможны перегибы 2-го уровня. Последний факт объясняется тем, что бинарный аргумент обладает, по сути, двумя степенями свободы — по u и по v .

Из теорем 5 и 6, как важное следствие, весьма просто вытекает классический *принцип отсутствия максимума модуля* аналитической (по Даламберу — Эйлеру) комплексной функции $z(x)$. Действительно, для функции $y = y(x, \bar{x}) = \bar{z}(\bar{x}) \cdot z(x) = \bar{y}$ в бинарной точке $(s, \bar{s})$ её возможной стационарности имеют место следующие соотношения:

$$\frac{\partial y}{\partial \bar{x}} = \frac{d\bar{z}}{d\bar{x}}(\bar{s}) \cdot z(s) = \overline{\left(\frac{dz}{dx}(s)\right)} \cdot z(s) = 0,$$

$$d^2y = \frac{d^2z}{dx^2}(s) \cdot \overline{z(s)} dx^2 + \overline{\left(\frac{d^2z}{dx^2}(s)\right)} \cdot z(s) d\bar{x}^2 + 2 \cdot \overline{\left(\frac{dz}{dx}(s)\right)} \cdot \frac{dz}{dx}(s) d\bar{x} dx.$$

Вариант 1. Если в первом уравнении $z(s) = 0$, то во второй формуле имеем $d^2y(x = s, \bar{x}_0 = \bar{s}) > 0$, что отвечает *минимуму* y и модуля $z(x)$.

Вариант 2. Если в первом уравнении $z(s) \neq 0$, то тогда $dz/dx(s) = 0$. При этом во второй формуле правый член нулевой. Это равносильно тому, что $\partial^2 y / \partial \bar{x} \partial x(s, \bar{s}) = 0$. Поэтому тут знак 2-го дифференциала $d^2y(x = s, \bar{x} = \bar{s})$ обязательно меняется, что отвечает стационарному перегибу y и соответственно модуля $z(x)$ — см. выше следствие 2.

В итоге отсюда заключаем, что модуль $z(x)$ не имеет максимума 2-го уровня. Доказательство по той же схеме распространяется на любой целочисленный уровень стационарности $y = y(x, \bar{x}) = \bar{z}(\bar{x}) \cdot z(x)$.

Для рассмотренного типа задач с одномерными аргументами x и $\bar{x}$ ещё не трудно, но и гораздо менее наглядно, получить тождественный результат их решения с использованием традиционного *вещественного* дифференциального анализа. Рассмотрим ниже кратко и последний.

Пусть в формуле (169): $s = a + i \cdot b$, $\theta = \cos \varphi + i \cdot \sin \varphi$. Далее имеем:

$$f(\theta, \bar{\theta}) \equiv r(\varphi) = 2 \cdot (a \cdot \cos \varphi + b \cdot \sin \varphi + g);$$

$$f_{\max} = 2 \cdot (g + \sqrt{a^2 + b^2}), \quad f_{\min} = 2 \cdot (g - \sqrt{a^2 + b^2}), \quad \text{где } a^2 + b^2 = s \cdot \bar{s}.$$

Эти значения параметров достигаются при $\tan \varphi = b/a$. При дальнейших извлечениях квадратного корня получаются два решения для углового аргумента функции $r(\varphi)$, отвечающие её максимуму и минимуму:

$$\cos \varphi_{1,2} = \pm a / \sqrt{a^2 + b^2}, \quad \sin \varphi_{1,2} = \pm b / \sqrt{a^2 + b^2}.$$

Геометрический смысл углов φ_1 и $\varphi_2 = -\varphi_1$ состоит также в том, что это суть углы комплексного сдвига для значений s и $\bar{s}$:

$$s = |s| \cdot (\cos \varphi_1 + i \cdot \sin \varphi_1), \quad \bar{s} = |s| \cdot (\cos \varphi_1 - i \cdot \sin \varphi_1) = |s| \cdot (\cos \varphi_2 + i \cdot \sin \varphi_2).$$

* * *

Пример. Продолжим рассмотрение 2-х задач (по примеру 2 из § 1.9) на экстремум пары целевых функций — разности (y_1) и отношения (y_2) среднего арифметического и среднего геометрического, но теперь для двух комплексных сопряжённых переменных $x = u + i \cdot v$ и $\bar{x} = u - i \cdot v$, где $u > 0$. Обе средние величины суть вещественные и положительные!

Задачи такого рода при количестве переменных $n = 2$ решаются тривиальным путём:

$$y_1 = \frac{x + \bar{x}}{2} - \sqrt{x \cdot \bar{x}} = u - \sqrt{u^2 + v^2} \leq 0,$$

$$y_2 = \frac{x + \bar{x}}{2 \cdot \sqrt{x \cdot \bar{x}}} = \frac{u}{\sqrt{u^2 + v^2}} \leq 1.$$

Отсюда следует, что обе функции формально не ограничены снизу, но имеют максимум $y_1 = 0$ и $y_2 = 1$ при $x = \bar{x} = u > 0$. Однако такие тривиальные способы решения задачи при $n > 2$ не реализуются.

Аналитический способ, напротив, и в данном случае тоже является универсальным, т. е. может применяться при любом $n \geq 2k$. Здесь же он демонстрируется на примере частной конкретной задачи при $n = 2$. Причём для упрощения, в силу положительности вещественной части сопряжённых переменных x и $\bar{x}$ ($u > 0$), в целевых функциях y_1 и y_2 применяются *квадраты* средних величин.

$$1) \quad y_1 = \psi_1(\mathbf{w}) = y_1(x, \bar{x}) = \left(\frac{x + \bar{x}}{2} \right)^2 - (\sqrt{x \cdot \bar{x}})^2 = \left(\frac{x - \bar{x}}{2} \right)^2.$$

$$\mathbf{g} = \mathbf{g}(\mathbf{w}) = \frac{dy_1}{d\mathbf{w}} = \left(\frac{\partial y_1}{\partial x}, \frac{\partial y_1}{\partial \bar{x}} \right) = (1/2(x - \bar{x}), 1/2(\bar{x} - x)) = \mathbf{0} \rightarrow$$

$$\rightarrow \mathbf{w} = \mathbf{s}^* = (x^*, \bar{x}^*), \quad x^* = \bar{x}^* = u > 0, \quad y_1(\mathbf{s}) = 0;$$

$$G = G(\mathbf{w}) = \frac{d^2 y_1}{d\mathbf{w}' d\mathbf{w}} = \begin{pmatrix} \frac{\partial^2 y_1}{\partial x^2} & \frac{\partial^2 y_1}{\partial \bar{x} \partial x} \\ \frac{\partial^2 y_1}{\partial \bar{x} \partial x} & \frac{\partial^2 y_1}{\partial \bar{x}^2} \end{pmatrix} = \begin{pmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{pmatrix} = \text{Const.}$$

Обратим внимание здесь на то, что матрица Гессе G *вещественная*! Но сразу идентифицировать характер стационарности функции $y_1(x, \bar{x})$ в точке стационарности исходя из значений вещественных параметров в ней $\text{tr } G = 1/2 + 1/2 = 1 > 0$ и $\det G = 1/4 - 1/4 = 0$ пока невозможно.

Конкретно можно только оценить, что при этих значениях данных параметров в её окрестности поверхность $y_1(x, \bar{x})$ полувыпуклая:

$$\begin{aligned} d^2 y_1 &= d\mathbf{w}' G d\mathbf{w} = \frac{\partial^2 y_1}{\partial x^2} dx^2 + 2 \frac{\partial^2 y_1}{\partial \bar{x} \partial x} d\bar{x} dx + \frac{\partial^2 y_1}{\partial \bar{x}^2} d\bar{x}^2 = \\ &= s dx^2 + 2g d\bar{x} dx + \bar{s} d\bar{x}^2 = 1/2 \cdot (dx^2 - 2d\bar{x} dx + d\bar{x}^2) = \\ &= \left(\frac{dx - d\bar{x}}{2} \right)^2 = (\mathbf{i} \cdot \mathbf{v})^2 = -v^2 \leq 0. \end{aligned}$$

Более общо имеем:

$$f_{\max} = 2 \cdot (g + \sqrt{s \cdot s}) = 2 \cdot (-1/2 + 1/2) = 0,$$

$$f_{\min} = 2 \cdot (g - \sqrt{s \cdot s}) = 2 \cdot (-1/2 - 1/2) = -2 < 0.$$

Согласно теореме 6, отсюда вытекает, что целевая функция $y_1(x, \bar{x})$ имеет на $\mathcal{O}$ нестрогий максимум на биссектрисе 1-го квадранта, т. е. при $\mathbf{x}^\bullet = \bar{\mathbf{x}}^\bullet = \mathbf{u} > 0$. Соответственно матрица Гессе G вырождена именно вдоль этой биссектрисы. Поверхность $y_1(x, \bar{x})$ в целом полувыпуклая.

$$2) y_2 = \psi_2(\mathbf{w}) = y_2(x, \bar{x}) = \left(\frac{x + \bar{x}}{2} / (\sqrt{x \cdot \bar{x}}) \right)^2 = 1/4 \cdot \left(\frac{x}{\bar{x}} + \frac{\bar{x}}{x} + 2 \right).$$

$$\mathbf{g} = \mathbf{g}(\mathbf{w}) = \frac{dy_2}{d\mathbf{w}} = \left(\frac{\partial y_2}{\partial x}, \frac{\partial y_2}{\partial \bar{x}} \right) = \left(1/4 \cdot \left(\bar{x} - \bar{x} / x^2 \right), 1/4 \cdot \left(x - x / \bar{x}^2 \right) \right) = \mathbf{0} \rightarrow$$

$$\rightarrow \mathbf{w} = \mathbf{s}^\bullet = (x^\bullet, \bar{x}^\bullet), \quad x^\bullet = \bar{x}^\bullet = \mathbf{u} > 0, \quad y_2(\mathbf{s}^\bullet) = 1;$$

$$\begin{aligned} G = G(\mathbf{w}) &= \frac{d^2 y_1}{d\mathbf{w}' d\mathbf{w}} = \begin{pmatrix} \frac{\partial^2 y_1}{\partial x^2} & \frac{\partial^2 y_1}{\partial \bar{x} \partial x} \\ \frac{\partial^2 y_1}{\partial \bar{x} \partial x} & \frac{\partial^2 y_1}{\partial \bar{x}^2} \end{pmatrix} = \\ &= \begin{pmatrix} \bar{x} / 2x^3 & -1/4 \cdot (1/x^2 + 1/\bar{x}^2) \\ -1/4 \cdot (1/x^2 + 1/\bar{x}^2) & x / 2\bar{x}^3 \end{pmatrix}. \end{aligned}$$

Обратим внимание на то, что G симметричная, а не эрмитова!

$$\begin{aligned}
 d^2 y_2 &= d\mathbf{w}' G d\mathbf{w} = \frac{\partial^2 y_2}{\partial x^2} dx^2 + 2 \frac{\partial^2 y_2}{\partial \bar{x} \partial x} d\bar{x} dx + \frac{\partial^2 y_2}{\partial \bar{x}^2} d\bar{x}^2 = \\
 &= s dx^2 + 2g d\bar{x} dx + \bar{s} d\bar{x}^2 = \\
 &= 1/2 \cdot \left[(\bar{x}/x^3) dx^2 - (1/x^2 + 1/\bar{x}^2) d\bar{x} dx + (x/\bar{x}^3) \right] d\bar{x}^2.
 \end{aligned}$$

Откуда имеем:

$$\begin{aligned}
 f_{\max} &= 2 \cdot (g + \sqrt{s \cdot s}) = -1/2 \cdot (1/x - 1/\bar{x})^2 \geq 0, \\
 f_{\min} &= 2 \cdot (g - \sqrt{s \cdot s}) = -1/2 \cdot (1/x + 1/\bar{x})^2 < 0.
 \end{aligned}$$

Согласно теореме 6, отсюда следует, что и в этом случае целевая функция $y_2(x, \bar{x})$ имеет на $\mathcal{O}$ нестрогий *максимум* на биссектрисе 1-го квадранта, т. е. при $x^* = \bar{x}^* = u > 0$, где матрица Гессе G вырождена, а поверхность $y_2(x, \bar{x})$ *локально* полувывпуклая. Но вне биссектрисы (!) та же поверхность $y_2(x, \bar{x})$ в пределах 1-го квадранта имеет повсюду седловинную форму. Исключением является только центр координат, в окрестности которого поверхность стремится к вещественной оси ординат ($y_2 \leq 1$).

Особо здесь отметим, что в отличие от ранее рассмотренного случая с вещественными аргументами x_1 и x_2 (пример 2 в § 1.9), где целевые функции y_1 и y_2 принимали на биссектрисе 1-го квадранта нестрогий *минимум*, в этом случае — при комплексных сопряжённых аргументах x и $\bar{x}$ ($u > 0$) обе целевые функции y_1 и y_2 также на биссектрисе 1-го квадранта принимают нестрогий *максимум*!!! Этот вывод, в частности, имеет применение в теории алгебраических уравнений (см. гл. 4).

Рассмотренная выше весьма замечательная вещественная двумерная поверхность $y_2(x, \bar{x})$ на оси $\mathcal{O}$ (как на *псевдоплоскости*!) дополняется симметричным отражением в 3-м комплексном квадранте, где средние величины от x и $\bar{x}$ формально отрицательные. Действительно, $f_{\max}$ и $f_{\min}$ можно трактовать как значения f , относящиеся именно к этим же двум формально возможным решениям для стационарности функции $y_2(x, \bar{x})$, т. е. $x^2 = \bar{x}^2$ относятся и к $x = \bar{x} = u$ и к $x = -\bar{x} = i \cdot v$. Эти решения отличаются только на угол своего комплексного сдвига $\pi/2$. Поэтому, как на вещественной биссектрисе 1-го квадранта, так и на мнимой биссектрисе 3-го квадранта функция y_2 имеет нестрогий *максимум*.

Итак, решение данной частной задачи рассмотрено теперь в полном объёме. Изложенный выше общий *аналитический* подход может быть использован для анализа и решения более сложных задач на экстремум вещественных целевых функций, у которых аргументы наиболее общо могут принимать как комплексные сопряжённые, так и вещественные значения. Яркими примерами таких функций являются коэффициенты вещественного алгебраического уравнения (суммы Виета) и полные суммы степеней их корней (суммы Варинга), см. далее §§ 4.1 — 4.4.

§ 3.4. Экстремумы для функций от пары независимых многомерных комплексных сопряжённых переменных

Для решения задач на безусловный экстремум целевой функции y от *многомерных* переменных $\mathbf{x}$ и $\bar{\mathbf{x}}$ вещественный анализ практически не пригоден. Напротив, формальный анализ позволяет решать такого типа задачи, сохранив подход и мнемоническую аналогию с изложенной выше процедурой решения задач с *одномерными* переменными x и $\bar{x}$.

Для вещественной скалярной функции $y = y(\mathbf{x}, \bar{\mathbf{x}})$ имеем её полные и частные формальные дифференциалы 1-го и 2-го порядков:

$$dy = \frac{\partial y}{\partial \mathbf{x}} d\mathbf{x} + \frac{\partial y}{\partial \bar{\mathbf{x}}} d\bar{\mathbf{x}} = \overline{dy}, \quad (173)$$

$$d^2 y = d\mathbf{x}' \frac{\partial^2 y}{\partial \mathbf{x}' \partial \mathbf{x}} d\mathbf{x} + 2d\mathbf{x}' \frac{\partial^2 y}{\partial \mathbf{x}' \partial \bar{\mathbf{x}}} d\bar{\mathbf{x}} + d\bar{\mathbf{x}}' \frac{\partial^2 y}{\partial \bar{\mathbf{x}}' \partial \bar{\mathbf{x}}} d\bar{\mathbf{x}} = \overline{d^2 y}. \quad (174)$$

Здесь частные *смешанные* дифференциалы y суть вещественные.

Теорема 7. Вещественная функция $y = y(\mathbf{x}, \bar{\mathbf{x}})$ стационарна в бинарной точке $(\mathbf{s}, \bar{\mathbf{s}})$ тогда и только тогда, когда

$$\frac{\partial y}{\partial \mathbf{x}}(\mathbf{s}, \bar{\mathbf{s}}) = \mathbf{0} \leftrightarrow \frac{\partial y}{\partial \bar{\mathbf{x}}}(\mathbf{s}, \bar{\mathbf{s}}) = \mathbf{0}. \quad (175)$$

(Причём вышеуказанные два требования равноценны.)

Доказательство. Действительно, из (173) и (157) непосредственно следует, что $dy = 0$ в точке $(\mathbf{s}, \bar{\mathbf{s}})$ в *любых* направлениях $d\mathbf{x} \leftrightarrow d\bar{\mathbf{x}}$ тогда и только тогда, когда в ней выполняется условие (175). ***

Далее найдём критерии для идентификации характера безусловной стационарности целевой функции y в бинарной точке $(\mathbf{s}, \bar{\mathbf{s}})$.

В самом тривиальном случае, когда в (174) однородные 2-ые тензор-производные нулевые, о характере стационарности y можно судить весьма просто — по знакоопределённости ненулевой *эрмитовой* (!) смешанной 2-й тензор-производной.

В более общем случае необходимо определить знак её 2-го полного дифференциала d^2y в точке $(\mathbf{s}, \bar{\mathbf{s}})$ во всевозможных из неё внутренних направлениях $d\mathbf{x} \leftrightarrow d\bar{\mathbf{x}}$. Вначале преобразуем формулу (174) к виду

$$d^2y = d\bar{\mathbf{x}}' \left[\frac{\partial^2 y}{\partial \bar{\mathbf{x}}' \partial \bar{\mathbf{x}}} \cdot \frac{d\bar{\mathbf{x}}}{d\mathbf{x}} + 2 \cdot \frac{\partial^2 y}{\partial \bar{\mathbf{x}}' \partial \mathbf{x}} + \left(\frac{d\mathbf{x}}{d\bar{\mathbf{x}}} \right)' \cdot \frac{\partial^2 y}{\partial \mathbf{x}' \partial \mathbf{x}} \right] d\mathbf{x} = \quad (176)$$

$$= d\mathbf{x}^* \left(\frac{\partial^2 y}{\partial \bar{\mathbf{x}}' \partial \bar{\mathbf{x}}} \cdot \theta + 2 \cdot \frac{\partial^2 y}{\partial \mathbf{x}' \partial \mathbf{x}} + \theta^* \cdot \frac{\partial^2 y}{\partial \mathbf{x}' \partial \mathbf{x}} \right) d\mathbf{x} = \quad (177)$$

$$= d\mathbf{x}^* F(\theta, \theta^*) d\mathbf{x} = d\mathbf{x}^* F^*(\theta^*, \theta) d\mathbf{x},$$

где θ есть $n \times n$ -матричный *директивный параметр* (по определению), причём $\det \theta \neq 0$. Выясним его матричную природу. Имеем:

$$d\bar{\mathbf{x}} = \theta d\mathbf{x} \rightarrow d\mathbf{x} = \bar{\theta} d\bar{\mathbf{x}} = \bar{\theta} \cdot \theta d\mathbf{x} \rightarrow d\bar{\mathbf{x}} = \theta \cdot \bar{\theta} d\bar{\mathbf{x}} \leftrightarrow d\mathbf{x} = \bar{\theta} \cdot \theta d\mathbf{x}.$$

Так как $d\mathbf{x}$ по направлению произволен, то $\bar{\theta} \cdot \theta = \theta \cdot \bar{\theta} = I$. Имеем:

$$d\mathbf{x} = \bar{\theta} d\bar{\mathbf{x}} \rightarrow d\mathbf{x}^* = d\mathbf{x}' \theta', \quad d\mathbf{x}^* d\mathbf{x} = \overline{d\mathbf{x}^* d\mathbf{x}} = d\mathbf{x}' \theta' \cdot \bar{\theta} d\bar{\mathbf{x}} = d\mathbf{x}^* \theta^* \cdot \theta d\mathbf{x}.$$

Так как $d\mathbf{x}$ по направлению произволен, то $\theta^* \cdot \theta = \theta \cdot \theta^* = I$.

Отсюда вытекает, что $\theta = \theta'$ — симметричная *унитарная* матрица. $\bar{\theta}$ и θ имеют диагональную форму в общем комплексном декартовом базисе $\tilde{E}_u = U \cdot \tilde{E}_x$, $u = U^{-1} x = U^* x$, $U^* U = U U^* = I$. Причём $U = U'$.

$$[d\bar{u}_k] = \text{diag}\{\theta\} du_k = [\theta_k du_k] \leftrightarrow [du_k] = \text{diag}\{\bar{\theta}\} d\bar{u}_k = [\bar{\theta}_k d\bar{u}_k].$$

Поскольку в (177) $d\mathbf{x}^* d\mathbf{x} > 0$, то знак d^2y в точке стационарности $(\mathbf{s}, \bar{\mathbf{s}})$ задаётся знакоопределённостью в ней же эрмитовой матричной функции $F(\theta, \theta^*)$ при имеющемся внутреннем ограничении $\theta^* \cdot \theta = I$.

Для упрощения дальнейших выкладок примем обозначения:

$$\partial^2 y / \partial \mathbf{x}' \partial \mathbf{x}(\mathbf{s}, \bar{\mathbf{s}}) = S, \quad \partial^2 y / \partial \bar{\mathbf{x}}' \partial \bar{\mathbf{x}}(\mathbf{s}, \bar{\mathbf{s}}) = \bar{S}, \quad \partial^2 y / \partial \mathbf{x}' \partial \bar{\mathbf{x}}(\mathbf{s}, \bar{\mathbf{s}}) = G = G^*.$$

Далее имеем:

$$F(\theta, \theta^*) = \bar{S} \cdot \theta + \theta^* \cdot S + 2G = F^*(\theta^*, \theta) = F(\theta^*, \theta). \quad (178)$$

Эта эрмитова $n \times n$ -матричная функция от переменных θ и θ^* должна при каких-то их значениях обеспечить условные максимум и минимум 2-го дифференциала в (177). Возникает вспомогательная задача, как и в § 3.3, вычисления условных экстремумов (максимума и минимума) скалярной функции со скалярным уравнением связи внутреннего типа:

$$\left. \begin{aligned} d^2 y(\theta, \theta^*) &= d\mathbf{x}^* F(\theta, \theta^*) d\mathbf{x} = \text{extr}, \\ h(\theta, \theta^*) &= d\mathbf{x}^* (\theta \cdot \theta^* - I) d\mathbf{x} = 0. \end{aligned} \right\} \quad (179)$$

Данную вспомогательную задачу решаем, используя *формальную модификацию* метода Лагранжа с матричными переменными θ и θ^* , а также с матричным множителем $\Lambda = \Lambda^*$, так как при этом $h = \bar{h}$.

$$\mathcal{L} = \mathcal{L}(\theta, \theta^*, \Lambda) = d\mathbf{x}^* \left[F(\theta, \theta^*) - (\theta^* \cdot \theta - I) \cdot \Lambda \right] d\mathbf{x} = \bar{\mathcal{L}} = \mathcal{L}(\theta^*, \theta, \Lambda). \quad (180)$$

$$\left. \begin{aligned} \left\{ \frac{\partial \mathcal{L}}{\partial \theta^*} \right\}_L &= d\mathbf{x}^* (S - \theta \cdot \Lambda) d\mathbf{x} = 0, \\ h(\theta, \theta^*) &= d\mathbf{x}^* (\theta^* \cdot \theta - I) d\mathbf{x} = 0. \end{aligned} \right\} (\Lambda = \Lambda^*) \quad (181)$$

Отметим, что в этом случае тензор-производная от скалярной функции Лагранжа по матричной переменной θ^* берётся *слева*! Отсюда при несингулярной S непосредственно следует:

$$\begin{aligned} S &= \theta \cdot \Lambda = S', \quad \bar{S} = \bar{\theta} \cdot \bar{\Lambda} = \bar{S}' = \Lambda^* \cdot \theta^* = \Lambda \cdot \theta^*, \\ \bar{S} \cdot S &= \Lambda \cdot \theta^* \cdot \theta \cdot \Lambda = \Lambda^2 = (\bar{S} \cdot S)^* \rightarrow \Lambda_{1,2} = \pm \sqrt{\bar{S} \cdot S}; \end{aligned}$$

$$\theta_{1,2} = S \cdot \Lambda_{1,2}^{-1} = \pm S \cdot \sqrt{\bar{S} \cdot S}^{-1} = \pm \sqrt{S \cdot \bar{S}}^{-1} = \pm \sqrt{\bar{S}^{-1}} \cdot S \quad (\text{при } \det S \neq 0).$$

Из этого непосредственно видно, что $\bar{\theta} \cdot \theta = \theta \cdot \bar{\theta} = I$. Кроме того, в силу коммутативности комплексных матриц S и $\bar{S}^{-1}$ ($\bar{S}$ и S^{-1}), отсюда же видно, что $\theta^* \cdot \theta = \theta \cdot \theta^* = I$.

Обратим особое внимание на то, что при вычислении $\Lambda_{1,2}$ берётся *арифметический корень* из эрмитовой положительно определённой матрицы, т. е. так, чтобы Λ_1 и Λ_2 были маргинальными (положительным и отрицательным) корнями! Далее имеем:

$$\left. \begin{aligned} \{d\mathbf{x}^* F(\theta, \theta^*) d\mathbf{x}\}_{\max} &= d\mathbf{x}^* F(\theta_1, \theta_1^*) d\mathbf{x} = d\mathbf{x}^* F_1 d\mathbf{x}, \\ \{d\mathbf{x}^* F(\theta, \theta^*) d\mathbf{x}\}_{\min} &= d\mathbf{x}^* F(\theta_2, \theta_2^*) d\mathbf{x} = d\mathbf{x}^* F_2 d\mathbf{x}; \end{aligned} \right\} \quad (182)$$

$$\left. \begin{aligned} F_1 &= 2 \cdot (G + \sqrt{\bar{S} \cdot S}), \quad \partial^2 \mathcal{L} / \partial \theta^* \partial \theta = -d\mathbf{x}^* \Lambda_1 d\mathbf{x} < 0; \\ F_2 &= 2 \cdot (G - \sqrt{\bar{S} \cdot S}), \quad \partial^2 \mathcal{L} / \partial \theta^* \partial \theta = -d\mathbf{x}^* \Lambda_2 d\mathbf{x} > 0. \end{aligned} \right\} \quad (183)$$

В окончательном виде формулы для 2-ых дифференциалов и для их матриц Гессе универсальны и справедливы даже при сингулярной S (поскольку вычисляются произведения типа $\bar{S} \cdot \theta$). При $S = Z$: $F = 2G$, т. е. вполне однозначна. Полученные решения обладают симметрией: при замене в самом начале $d\mathbf{x}$ на $d\bar{\mathbf{x}}$ и $d\bar{\mathbf{x}}$ на $d\mathbf{x}$ получаем аналогичные эрмитовы матрицы Гессе $\bar{F}_1 = 2 \cdot (\bar{G} + \sqrt{S \cdot \bar{S}})$ и $\bar{F}_2 = 2 \cdot (\bar{G} - \sqrt{S \cdot \bar{S}})$.

Следовательно, знак d^2y в точке стационарности целевой функции $y = y(\mathbf{x}, \bar{\mathbf{x}})$ задают знакоопределённости *маргинальных* матриц F_1 и F_2 . Это есть всё тот же *эффект расширения* (§ 3.3) 2-го дифференциала скалярной вещественной функции $y = y(\mathbf{x}, \bar{\mathbf{x}})$. Эффект имеет место для ненулевого d^2y при условии $\partial^2 y / \partial \mathbf{x}' \partial \mathbf{x} \neq Z$.

Из вышеизложенного следует:

Теорема 8. Вещественная целевая функция $y = y(\mathbf{x}, \bar{\mathbf{x}})$ в бинарной точке своей стационарности $(\mathbf{s}, \bar{\mathbf{s}})$ имеет максимум 2-го уровня, если F_1 отрицательно определённая; имеет минимум 2-го уровня, если F_2 положительно определённая; имеет седловину 2-го уровня, если F_1 и F_2 обладают хотя бы двумя собственными значениями разного знака. Случаи, когда они имеют в совокупности нулевые и отрицательные (нулевые и положительные) собственные значения отвечают нестрогим максимуму (минимуму) того же 2-го уровня.

Следствие 1. Если $\partial^2 y / \partial \mathbf{x}' \partial \mathbf{x}(\mathbf{s}, \bar{\mathbf{s}}) = Z \leftrightarrow \partial^2 y / \partial \bar{\mathbf{x}}' \partial \bar{\mathbf{x}}(\mathbf{s}, \bar{\mathbf{s}}) = Z$, то характер стационарности функции y 2-го уровня однозначно задаёт ненулевая $\partial^2 y / \partial \mathbf{x}^* \partial \mathbf{x}(\mathbf{s}, \bar{\mathbf{s}})$.

Следствие 2. Если $\partial^2 y / \partial \mathbf{x}^* \partial \mathbf{x}(\mathbf{s}, \bar{\mathbf{s}}) = Z$ и $\partial^2 y / \partial \bar{\mathbf{x}}' \partial \bar{\mathbf{x}}(\mathbf{s}, \bar{\mathbf{s}}) \neq Z$, то функция y в точке $(\mathbf{s}, \bar{\mathbf{s}})$ имеет стационарный перегиб 2-го уровня.

Следствие 3. Если $\partial^2 y / \partial \mathbf{x}^* \partial \mathbf{x}(\mathbf{s}, \bar{\mathbf{s}})$ положительно (полу)определена, то функция y в точке $(\mathbf{s}, \bar{\mathbf{s}})$ не может иметь максимума.

Следствие 4. Если $\partial^2 y / \partial \mathbf{x}^* \partial \mathbf{x}(\mathbf{s}, \bar{\mathbf{s}})$ отрицательно (полу)определена, то функция y в точке $(\mathbf{s}, \bar{\mathbf{s}})$ не может иметь минимума.

Из теорем 7 и 8, как следствие, вытекает *общий принцип отсутствия максимума модуля* аналитической (по Даламберу — Эйлеру) векторной комплексной функции $\mathbf{z}(\mathbf{x})$, где $\mathbf{x}$ есть $q \times 1$ -вектор в $\mathcal{E}^q$, $\mathbf{z}$ есть $n \times 1$ -вектор в $\mathcal{E}^n$; $\text{rang}\{\mathbf{dz}/d\mathbf{x}\} = q \leq n$. (Это значит, что *линеаризация*, или касательная плоскость $\langle \mathbf{x} \rangle$ для данной векторной функции в $\mathcal{E}^n$ имеет размерность q .) Действительно, для $y = y(\mathbf{x}, \bar{\mathbf{x}}) = \mathbf{z}^*(\mathbf{x}) \cdot \mathbf{z}(\mathbf{x}) = \bar{y}$ имеют место соотношения в бинарной точке её возможной стационарности:

$$\begin{aligned} \frac{\partial y}{\partial \bar{\mathbf{x}}} &= \left(\frac{d\mathbf{z}}{d\mathbf{x}}(\mathbf{s}) \right)^* \cdot \mathbf{z}(\mathbf{s}) = \mathbf{0}, \\ d^2y &= d\mathbf{x}' \frac{\partial y^2}{\partial \mathbf{x}' \partial \mathbf{x}}(\mathbf{s}, \bar{\mathbf{s}}) d\mathbf{x} + d\bar{\mathbf{x}}' \frac{\partial y^2}{\partial \bar{\mathbf{x}}' \partial \bar{\mathbf{x}}}(\mathbf{s}, \bar{\mathbf{s}}) d\bar{\mathbf{x}} + 2 d\mathbf{x}^* \frac{\partial y^2}{\partial \mathbf{x}^* \partial \mathbf{x}}(\mathbf{s}, \bar{\mathbf{s}}) d\mathbf{x} = \\ &= d\mathbf{x}' \left[\left(\frac{d^2 \mathbf{z}}{d\mathbf{x}' d\mathbf{x}}(\mathbf{s}) \right)^{(1 \leftrightarrow 3)} \cdot \overline{\mathbf{z}(\mathbf{s})} \right]_{n \times 1} d\mathbf{x} + d\bar{\mathbf{x}}' \left[\left(\frac{d^2 \mathbf{z}}{d\mathbf{x}' d\mathbf{x}}(\mathbf{s}) \right)^{(1 \leftrightarrow 3)} \cdot \mathbf{z}(\mathbf{s}) \right]_{n \times 1} d\bar{\mathbf{x}} + \\ &+ 2 d\mathbf{x}^* \left[\left(\frac{d\mathbf{z}}{d\mathbf{x}}(\mathbf{s}) \right)^* \cdot \frac{d\mathbf{z}}{d\mathbf{x}}(\mathbf{s}) \right]_{q \times n \quad n \times q} d\mathbf{x}. \end{aligned}$$

Вариант 1. Если $\mathbf{z}(\mathbf{s}) = \mathbf{0}$, то $d^2y(\mathbf{x} = \mathbf{s}, \bar{\mathbf{x}} = \bar{\mathbf{s}}) \geq 0$, так как эрмитова матрица 2-й смешанной тензор-производной G положительно определённая или полуопределённая, что отвечает минимуму модуля $\mathbf{z}(\mathbf{x})$, в том числе нестрогому.

Вариант 2. Если $\mathbf{z}(\mathbf{s}) \neq \mathbf{0}$, то при этом $\mathbf{z}(\mathbf{s}) \in \langle \ker\{d\mathbf{z}/d\mathbf{x}(\mathbf{s})\}^* \rangle$ и, в частности, $d\mathbf{z}/d\mathbf{x}(\mathbf{s}) = \mathbf{Z}$. Тогда целевая функция y в точке $(\mathbf{s}, \bar{\mathbf{s}})$, в силу следствия 3 теоремы 6, не может иметь максимума, так как 2-я смешанная тензор-производная G положительно определённая или полуопределённая.

До сих пор в данной главе рассматривались, как основные, задачи безусловной оптимизации целевых функций от парных комплексных сопряжённых переменных, т. е. без какого-либо ограничения на область их изменения в комплексном координатном пространстве. Для большей целостности это исследование необходимо дополнить рассмотрением задач условной оптимизации тех же целевых функций с двумя типами ограничения на переменную — *внешним и внутренним*.

§ 3.5. Условные экстремумы для функций от ограниченных внешне комплексных переменных

В данном параграфе $n \times 1$ -вектор $\tilde{\mathbf{x}}$ играет роль *внешне* ограниченной переменной, подчиняющейся уравнению связи $\mathbf{h}(\tilde{\mathbf{x}}) = \mathbf{0}$, $\mathbf{h}(\mathbf{x} \notin \langle \tilde{\mathbf{x}} \rangle) \neq \mathbf{0}$. Причём $\mathbf{h} = \mathbf{h}(\mathbf{x}) \in \mathcal{E}^m$ есть комплексная аналитическая $m \times 1$ -вектор-функция ($m < n$), однозначно определённая на множестве $\mathcal{E}^n$; при этом $\text{rang}\{\mathbf{dh}/\mathbf{dx}\} = m < n$. Уравнение связи задаёт некоторое компактное подмножество в $\mathcal{E}^n$, или некую геометрическую поверхность $\langle \tilde{\mathbf{x}} \rangle \subset \mathcal{E}^n$ размерности $q = n - m$. Исходной функции ограничения $\mathbf{h} = \mathbf{h}(\mathbf{x})$ взаимно однозначно соответствует комплексно сопряжённая функция $\bar{\mathbf{h}} = \bar{\mathbf{h}}(\bar{\mathbf{x}})$. Очевидно, что $\mathbf{h}(\mathbf{x})$ и $\bar{\mathbf{h}}(\bar{\mathbf{x}})$ принимают все свои нулевые значения на бинарном компактном подмножестве $\langle \langle \tilde{\mathbf{x}} \rangle \cup \langle \tilde{\bar{\mathbf{x}}} \rangle \rangle \subset \mathcal{E}^n$. При оговорённых выше требованиях к вектор-функции ограничения это подмножество есть бинарная, *формально* гладкая и регулярная (причём порядка гладкости и регулярности не менее двух) комплексная геометрическая поверхность размерности $q < n$, вложенная в комплексное аффинное координатное пространство $\mathcal{E}^n$.

Далее ставится задача поиска и идентификации характера условного экстремума функции $y(\mathbf{x}, \bar{\mathbf{x}})$ на бинарной q -поверхности $\langle \langle \tilde{\mathbf{x}} \rangle \cup \langle \tilde{\bar{\mathbf{x}}} \rangle \rangle$, заданной посредством внешнего ограничения на переменную $\mathbf{h}(\tilde{\mathbf{x}}) = \mathbf{0}$. Суть задачи состоит в поиске экстремума целевой функции на гладкой и регулярной комплексной q -поверхности в аффинном комплексном пространстве, заданной функционально внешним ограничительным способом.

Для решения поставленной задачи применим наиболее подходящую здесь *клеточную модификацию* метода Лагранжа и метода условных тензор-производных из § 2.3, с учётом операций формального анализа. Принимая, что векторы $\mathbf{x}$ и $\bar{\mathbf{x}}$, $\mathbf{h}$ и $\bar{\mathbf{h}}$ — аналитически независимые характеристики в этих обеих парах, представим соответственно саму процедуру и само искомое решение поклеточно в *двухблочной* форме. Напомним тут (см. начало § 2.2), что при дифференцировании по $\tilde{\mathbf{x}}$ последняя понимается обычным образом как *свободная* переменная, т. е. как изменяющаяся от своего значения в любой точке множества $\langle \tilde{\mathbf{x}} \rangle$ во всевозможных направлениях в координатном пространстве $\mathcal{E}^n$. Далее для большей наглядности параллельно применяются и метод Лагранжа, и проективный метод условных тензор-производных.

Проективные формулы, как аналоги формул типа (74), для условных дифференциалов внешне ограниченных парных переменных $\tilde{\mathbf{x}}$ и $\tilde{\bar{\mathbf{x}}}$ имеют схожий вид:

$$d\tilde{\mathbf{x}} = \left[\left(\frac{d\mathbf{h}}{d\tilde{\mathbf{x}}} \right)' \cdot \frac{d\mathbf{h}}{d\tilde{\mathbf{x}}} \right] d\mathbf{x} \leftrightarrow d\tilde{\bar{\mathbf{x}}} = \left[\left(\frac{d\bar{\mathbf{h}}}{d\tilde{\bar{\mathbf{x}}}} \right)' \cdot \frac{d\bar{\mathbf{h}}}{d\tilde{\bar{\mathbf{x}}}} \right] d\bar{\mathbf{x}}. \quad (184)$$

Первый проектор осуществляет *аффинное* проецирование в $\mathcal{E}^n$ на ядро $\langle \ker d\mathbf{h}/d\tilde{\mathbf{x}} \rangle$ параллельно образу $\langle \text{im } d\mathbf{h}/d\tilde{\mathbf{x}} \rangle_{\mathbf{R}}$. Второй проектор осуществляет *аффинное* проецирование в $\mathcal{E}^n$ на ядро $\langle \ker d\bar{\mathbf{h}}/d\tilde{\bar{\mathbf{x}}} \rangle$ параллельно образу $\langle \text{im } d\bar{\mathbf{h}}/d\tilde{\bar{\mathbf{x}}} \rangle_{\mathbf{R}}$. Это по-прежнему симметричные, причём *адекватно*, комплексные проекторы (см. § 3.1). В комплексном *евклидовом* пространстве эти проекторы проецируют ортогонально на ядро матрицы Якоби, т. е. на q -поверхности ограничения $\langle \tilde{\mathbf{x}} \rangle$ и $\langle \tilde{\bar{\mathbf{x}}} \rangle$.

Формальная модификация метода условных тензор-производных (из § 2.2) даёт систему векторных уравнений — аналог системы (77):

$$\left[\frac{\partial \mathbf{y}}{\partial \mathbf{x}}(\tilde{\mathbf{s}}, \tilde{\bar{\mathbf{s}}}) \right] = \frac{\partial \mathbf{y}}{\partial \tilde{\mathbf{x}}}(\tilde{\mathbf{s}}, \tilde{\bar{\mathbf{s}}}) \cdot \left[\left(\frac{d\mathbf{h}}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \right)' \cdot \frac{d\mathbf{h}}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \right] = \mathbf{0}, \quad \left(\frac{\partial \mathbf{y}}{\partial \tilde{\mathbf{x}}}(\tilde{\mathbf{s}}, \tilde{\bar{\mathbf{s}}}) \neq \mathbf{0} \right). \quad (185)$$

$$\mathbf{h}(\tilde{\mathbf{x}}) = \mathbf{0};$$

Ей же отвечает тождественная комплексно сопряжённая система. Любая из этих 2-х систем даёт как решение бинарную точку условной стационарности $(\tilde{\mathbf{s}}, \tilde{\bar{\mathbf{s}}})$. Оба вектор-множителя Лагранжа вычисляются обычным образом как в (83), с учётом их комплексного сопряжения:

$$\lambda_1 = \frac{\partial \mathbf{y}}{\partial \tilde{\mathbf{x}}}(\tilde{\mathbf{s}}, \tilde{\bar{\mathbf{s}}}) \cdot \left(\frac{d\mathbf{h}}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \right)^{-1} = \lambda, \quad \lambda_2 = \bar{\lambda}. \quad (186)$$

Функция Лагранжа имеет вид:

$$\mathcal{L} = \mathcal{L}(\tilde{\mathbf{x}}, \tilde{\bar{\mathbf{x}}}, \lambda, \bar{\lambda}) = \mathbf{y}(\tilde{\mathbf{x}}, \tilde{\bar{\mathbf{x}}}) - [\lambda \cdot \mathbf{h}(\tilde{\mathbf{x}}) + \bar{\lambda} \cdot \bar{\mathbf{h}}(\tilde{\bar{\mathbf{x}}})] = \bar{\mathcal{L}}(\tilde{\mathbf{x}}, \tilde{\bar{\mathbf{x}}}, \lambda, \bar{\lambda}). \quad (187)$$

Формальная модификация метода Лагранжа даёт систему — аналог (85):

$$\left. \begin{aligned} \frac{\partial \mathcal{L}}{\partial \tilde{\mathbf{x}}}(\tilde{\mathbf{s}}, \tilde{\bar{\mathbf{s}}}) &= \frac{\partial \mathbf{y}}{\partial \tilde{\mathbf{x}}}(\tilde{\mathbf{s}}, \tilde{\bar{\mathbf{s}}}) - \lambda \cdot \frac{d\mathbf{h}}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) = \mathbf{0}, \\ \frac{\partial \mathcal{L}}{\partial \bar{\lambda}} &= -\bar{\mathbf{h}}(\tilde{\bar{\mathbf{x}}}) = \mathbf{0}; \end{aligned} \right\} \quad \left(\frac{\partial \mathbf{y}}{\partial \tilde{\mathbf{x}}}(\tilde{\mathbf{s}}, \tilde{\bar{\mathbf{s}}}) \neq \mathbf{0} \right). \quad (188)$$

С учётом отсутствия аналитической взаимосвязи пары переменных $\tilde{\mathbf{x}}$ и $\tilde{\bar{\mathbf{x}}}$, характер условной стационарности в бинарной точке $(\tilde{\mathbf{s}}, \tilde{\bar{\mathbf{s}}})$ можно установить простым путём — через *формальный* анализ 2-го условного дифференциала функции Лагранжа (187):

$$\begin{aligned} \left[d^2 \mathcal{L}(\tilde{\mathbf{x}} = \tilde{\mathbf{s}}, \tilde{\bar{\mathbf{x}}} = \tilde{\bar{\mathbf{s}}}, \lambda, \bar{\lambda}) \right] &= d \tilde{\mathbf{x}}' \tilde{\mathbf{S}} d \tilde{\mathbf{x}} + 2 d \tilde{\mathbf{x}}' \tilde{\mathbf{G}} d \tilde{\bar{\mathbf{x}}} + d \tilde{\bar{\mathbf{x}}} \tilde{\bar{\mathbf{S}}} d \tilde{\bar{\mathbf{x}}} = \\ &= d \tilde{\mathbf{x}}' \tilde{\mathbf{S}} d \tilde{\mathbf{x}} + 2 d \tilde{\mathbf{x}}' \tilde{\mathbf{G}} d \tilde{\mathbf{x}} + d \tilde{\bar{\mathbf{x}}} \tilde{\bar{\mathbf{S}}} d \tilde{\bar{\mathbf{x}}}, \end{aligned} \quad (189)$$

где применяются следующие обозначения для *условных* матриц $\tilde{\mathbf{S}}$ и $\tilde{\mathbf{G}}$:

$$\tilde{\mathbf{S}} = \left[\left(\frac{d\mathbf{h}}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \right)' \cdot \frac{d\mathbf{h}}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \right] \cdot \frac{\partial^2 \mathcal{L}}{\partial \tilde{\mathbf{x}}' \partial \tilde{\mathbf{x}}}(\tilde{\mathbf{s}}, \tilde{\bar{\mathbf{s}}}) \cdot \left[\left(\frac{d\mathbf{h}}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \right)' \cdot \frac{d\mathbf{h}}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \right] = \tilde{\mathbf{S}}', \quad (190)$$

$$\tilde{\mathbf{G}} = \left[\left(\frac{d\bar{\mathbf{h}}}{d\tilde{\bar{\mathbf{x}}}(\tilde{\bar{\mathbf{s}}})} \right)' \cdot \frac{d\bar{\mathbf{h}}}{d\tilde{\bar{\mathbf{x}}}(\tilde{\bar{\mathbf{s}}})} \right] \cdot \frac{\partial^2 \mathcal{L}}{\partial \tilde{\mathbf{x}}^* \partial \tilde{\bar{\mathbf{x}}}(\tilde{\mathbf{s}}, \tilde{\bar{\mathbf{s}}})} \cdot \left[\left(\frac{d\mathbf{h}}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \right)' \cdot \frac{d\mathbf{h}}{d\tilde{\mathbf{x}}}(\tilde{\mathbf{s}}) \right] = \tilde{\mathbf{G}}^*. \quad (191)$$

Далее по аналогии с преобразованием (176) $\rightarrow$ (177) 2-й условный дифференциал функции Лагранжа приводится к *эрмитовой* форме:

$$\left[d^2 \mathcal{L}(\tilde{\mathbf{x}} = \tilde{\mathbf{s}}, \tilde{\bar{\mathbf{x}}} = \tilde{\bar{\mathbf{s}}}, \lambda, \bar{\lambda}) \right] = d \tilde{\mathbf{x}}^* \left(\tilde{\mathbf{S}} \cdot \tilde{\Theta} + \tilde{\Theta}^* \cdot \tilde{\mathbf{S}} + 2 \tilde{\mathbf{G}} \right) d \tilde{\mathbf{x}}. \quad (192)$$

Два крайних значения матричного параметра отвечают маргинальным положительным и отрицательным арифметическим корням из $\{\tilde{\mathbf{S}} \cdot \tilde{\mathbf{S}}\}$:

$$\tilde{\Theta}_{1,2} = \pm \tilde{\mathbf{S}} \cdot \sqrt{\tilde{\mathbf{S}} \cdot \tilde{\mathbf{S}}}^{-1} = \pm \sqrt{\tilde{\mathbf{S}} \cdot \tilde{\mathbf{S}}^{-1}} = \pm \sqrt{\tilde{\mathbf{S}}^{-1} \cdot \tilde{\mathbf{S}}} \quad \text{при } \det \tilde{\mathbf{S}} \neq 0.$$

Выводы о характере условной стационарности делаются в соответствии с формулировками теоремы 8 анализом маргинальных функций:

$$F_1 = 2 \cdot (\tilde{\mathbf{G}} + \sqrt{\tilde{\mathbf{S}} \cdot \tilde{\mathbf{S}}}), \quad F_2 = 2 \cdot (\tilde{\mathbf{G}} - \sqrt{\tilde{\mathbf{S}} \cdot \tilde{\mathbf{S}}}).$$

Отметим, что для скалярных функций $y = y(\mathbf{x}, \bar{\mathbf{x}})$ при *внешнем* характере ограничения на бинарную комплексную переменную, как и для скалярных функций от вещественной переменной $y = y(\mathbf{x})$, истинные условные экстремумы и перегибы возможны лишь при числе измерений более 1. Это следует из вышеприведённого анализа, где *непосредственно* проявляется и используется адекватная комплексификация (§ 3.1).

§ 3.6. Условные экстремумы для функций от ограниченных внутренне комплексных переменных

В этом параграфе $n \times 1$ -вектор $\tilde{\mathbf{x}}$ играет роль *внутренне* ограниченной переменной, подчиняющейся вещественному уравнению связи особого типа $\mathbf{h}(\tilde{\mathbf{x}}, \tilde{\bar{\mathbf{x}}}) = \bar{\mathbf{h}}(\tilde{\mathbf{x}}, \tilde{\bar{\mathbf{x}}}) = \mathbf{h}(\tilde{\mathbf{x}}, \tilde{\bar{\mathbf{x}}}) = \mathbf{0}$; $\mathbf{h}(\langle \mathbf{x}, \bar{\mathbf{x}} \rangle \notin \langle \tilde{\mathbf{x}}, \tilde{\bar{\mathbf{x}}} \rangle) \neq \mathbf{0}$. Причём $\mathbf{h} = \mathbf{h}(\mathbf{x}, \bar{\mathbf{x}}) \in \mathcal{E}^m$ — неголоморфная $m \times 1$ -вектор-функция ($m \leq n$), однозначно определённая на множестве $\mathcal{E}^n$ и, по крайней мере, дважды *формально* непрерывно дифференцируемая по переменным $\mathbf{x}$ и $\bar{\mathbf{x}}$; при этом $\text{rang}\{\partial \mathbf{h} / \partial \mathbf{x}\} = \text{rang}\{\partial \mathbf{h} / \partial \bar{\mathbf{x}}\} = m < n$. Уравнение связи задаёт некоторое компактное подмножество в $\mathcal{E}^n$, или некую геометрическую поверхность $\langle \tilde{\mathbf{x}}, \tilde{\bar{\mathbf{x}}} \rangle \subset \mathcal{E}^n$ размерности $q < n$, где $q = n - m$. При оговорённых выше требованиях к вектор-функции ограничения данное подмножество есть бинарная *формально* гладкая и регулярная (порядка гладкости и регулярности не менее двух) комплексная геометрическая поверхность размерности q , вложенная в комплексное аффинное координатное пространство $\mathcal{E}^n$.

Далее ставится задача поиска и идентификации характера условного экстремума функции $y = y(\mathbf{x}, \bar{\mathbf{x}})$ на бинарной q -поверхности $\langle \tilde{\mathbf{x}}, \tilde{\bar{\mathbf{x}}} \rangle$, заданной теперь посредством внутреннего ограничения на переменную $\mathbf{h}(\tilde{\mathbf{x}}, \tilde{\bar{\mathbf{x}}}) = \bar{\mathbf{h}}(\tilde{\mathbf{x}}, \tilde{\bar{\mathbf{x}}}) = \mathbf{0}$. Суть этой задачи состоит в поиске экстремума целевой функции на гладкой и регулярной комплексной q -поверхности в аффинном комплексном пространстве $\mathcal{E}^n$, заданной функционально внутренним ограничительным способом.

Проективные формулы, как аналоги прототипов (74), для условных дифференциалов внутренне ограниченных переменных $\tilde{\mathbf{x}}$ и $\tilde{\bar{\mathbf{x}}}$ имеют схожий вид:

$$d\tilde{\mathbf{x}} = \left[\left(\frac{\partial \mathbf{h}}{\partial \tilde{\mathbf{x}}} \right)' \cdot \frac{\partial \mathbf{h}}{\partial \tilde{\bar{\mathbf{x}}}} \right] d\tilde{\bar{\mathbf{x}}} \leftrightarrow d\tilde{\bar{\mathbf{x}}} = \left[\left(\frac{\partial \mathbf{h}}{\partial \tilde{\bar{\mathbf{x}}}} \right)' \cdot \frac{\partial \mathbf{h}}{\partial \tilde{\mathbf{x}}} \right] d\tilde{\mathbf{x}}. \quad (193)$$

Причём первый проектор осуществляет *аффинное* проецирование в $\mathcal{E}^n$ на ядро $\langle \ker \partial \mathbf{h} / \partial \tilde{\mathbf{x}} \rangle$ параллельно образу $\langle \text{im } \partial \mathbf{h} / \partial \tilde{\bar{\mathbf{x}}} \rangle_{\mathbf{R}}$. Второй проектор также осуществляет *аффинное* проецирование в $\mathcal{E}^n$, но на ядро $\langle \ker \partial \bar{\mathbf{h}} / \partial \tilde{\bar{\mathbf{x}}} \rangle$ параллельно образу $\langle \text{im } \partial \bar{\mathbf{h}} / \partial \tilde{\mathbf{x}} \rangle_{\mathbf{R}}$. Это суть *адекватно* симметричные комплексные проекторы (§ 3.1). В комплексном *евклидовом* пространстве данные проекторы ортогонально проецируют на те же ядра сопряжённых матриц Якоби.

Формальная модификация метода условных тензор-производных (из § 2.2) даёт систему векторных уравнений — аналог системы (77):

$$\left[\frac{\partial y}{\partial \tilde{\mathbf{x}}}(\tilde{\mathbf{s}}, \tilde{\mathbf{s}}) \right] = \frac{\partial y}{\partial \tilde{\mathbf{x}}}(\tilde{\mathbf{s}}, \tilde{\mathbf{s}}) \cdot \left[\left(\frac{\partial \mathbf{h}}{\partial \tilde{\mathbf{x}}}(\tilde{\mathbf{s}}, \tilde{\mathbf{s}}) \right)' \cdot \frac{\partial \mathbf{h}}{\partial \tilde{\mathbf{x}}}(\tilde{\mathbf{s}}, \tilde{\mathbf{s}}) \right] = \mathbf{0}, \left\{ \begin{array}{l} \left(\frac{\partial y}{\partial \tilde{\mathbf{x}}}(\tilde{\mathbf{s}}, \tilde{\mathbf{s}}) \neq \mathbf{0} \right), \\ \mathbf{h}(\tilde{\mathbf{x}}, \tilde{\mathbf{x}}) = \mathbf{0}; \end{array} \right. \quad (194)$$

Ей же отвечает тождественная комплексно сопряжённая система. Любая из этих 2-х систем даёт как решение бинарную точку условной стационарности $(\tilde{\mathbf{s}}, \tilde{\mathbf{s}})$. Оба вектор-множителя Лагранжа вычисляются обычным образом как в (83), с учётом их комплексного сопряжения:

$$\lambda_1 = \frac{\partial y}{\partial \tilde{\mathbf{x}}}(\tilde{\mathbf{s}}, \tilde{\mathbf{s}}) \cdot \left(\frac{\partial \mathbf{h}}{\partial \tilde{\mathbf{x}}}(\tilde{\mathbf{s}}, \tilde{\mathbf{s}}) \right)_R^{-1} = \lambda, \quad \lambda_2 = \bar{\lambda}. \quad (195)$$

Функция Лагранжа имеет вид:

$$\mathcal{L} = \mathcal{L}(\mathbf{x}, \tilde{\mathbf{x}}, \lambda, \bar{\lambda}) = y(\tilde{\mathbf{x}}, \tilde{\mathbf{x}}) - (\lambda + \bar{\lambda}) \cdot \mathbf{h}(\tilde{\mathbf{x}}, \tilde{\mathbf{x}}) = \bar{\mathcal{L}}(\bar{\mathbf{x}}, \tilde{\mathbf{x}}, \lambda, \bar{\lambda}). \quad (196)$$

Формальная модификация метода Лагранжа даёт систему — аналог (85):

$$\left. \begin{array}{l} \frac{\partial \mathcal{L}}{\partial \tilde{\mathbf{x}}}(\tilde{\mathbf{s}}, \tilde{\mathbf{s}}) = \frac{\partial y}{\partial \tilde{\mathbf{x}}}(\tilde{\mathbf{s}}, \tilde{\mathbf{s}}) - \lambda \cdot \frac{\partial \mathbf{h}}{\partial \tilde{\mathbf{x}}}(\tilde{\mathbf{s}}, \tilde{\mathbf{s}}) = \mathbf{0}, \\ \frac{\partial \mathcal{L}}{\partial \lambda} = -\mathbf{h}(\tilde{\mathbf{x}}, \tilde{\mathbf{x}}) = \mathbf{0}; \end{array} \right\} \left(\frac{\partial y}{\partial \tilde{\mathbf{x}}}(\tilde{\mathbf{s}}, \tilde{\mathbf{s}}) \neq \mathbf{0} \right). \quad (197)$$

С учётом отсутствия аналитической взаимосвязи пары переменных $\tilde{\mathbf{x}}$ и $\tilde{\mathbf{x}}$, характер условной стационарности в бинарной точке $(\tilde{\mathbf{s}}, \tilde{\mathbf{s}})$ можно тут установить простым путём — через формальный анализ 2-го условного дифференциала функции Лагранжа (196):

$$\begin{aligned} \left[d^2 \mathcal{L}(\tilde{\mathbf{x}} = \tilde{\mathbf{s}}, \tilde{\mathbf{x}} = \tilde{\mathbf{s}}, \lambda, \bar{\lambda}) \right] &= d \tilde{\mathbf{x}}' S d \tilde{\mathbf{x}} + 2 d \tilde{\mathbf{x}}' G d \tilde{\mathbf{x}} + d \tilde{\mathbf{x}}' \tilde{S} d \tilde{\mathbf{x}} = \\ &= d \mathbf{x}' \tilde{S} d \mathbf{x} + 2 d \mathbf{x}' \tilde{G} d \mathbf{x} + d \bar{\mathbf{x}}' \tilde{S} d \bar{\mathbf{x}}, \end{aligned} \quad (198)$$

где применяются следующие обозначения для условных матриц S и G:

$$\tilde{S} = \left[\left(\frac{\partial \mathbf{h}}{\partial \tilde{\mathbf{x}}}(\tilde{\mathbf{s}}, \tilde{\mathbf{s}}) \right)' \cdot \frac{\partial \mathbf{h}}{\partial \tilde{\mathbf{x}}}(\tilde{\mathbf{s}}, \tilde{\mathbf{s}}) \right] \cdot \frac{\partial^2 \mathcal{L}}{\partial \tilde{\mathbf{x}}' \partial \tilde{\mathbf{x}}}(\tilde{\mathbf{s}}, \tilde{\mathbf{s}}) \cdot \left[\left(\frac{\partial \mathbf{h}}{\partial \tilde{\mathbf{x}}}(\tilde{\mathbf{s}}, \tilde{\mathbf{s}}) \right)' \cdot \frac{\partial \mathbf{h}}{\partial \tilde{\mathbf{x}}}(\tilde{\mathbf{s}}, \tilde{\mathbf{s}}) \right] = \tilde{S}', \quad (199)$$

$$\tilde{G} = \left[\left(\frac{\partial \mathbf{h}}{\partial \tilde{\mathbf{x}}}(\tilde{\mathbf{s}}, \tilde{\mathbf{s}}) \right)' \cdot \frac{\partial \mathbf{h}}{\partial \tilde{\mathbf{x}}}(\tilde{\mathbf{s}}, \tilde{\mathbf{s}}) \right] \cdot \frac{\partial^2 \mathcal{L}}{\partial \tilde{\mathbf{x}}' \partial \tilde{\mathbf{x}}}(\tilde{\mathbf{s}}, \tilde{\mathbf{s}}) \cdot \left[\left(\frac{\partial \mathbf{h}}{\partial \tilde{\mathbf{x}}}(\tilde{\mathbf{s}}, \tilde{\mathbf{s}}) \right)' \cdot \frac{\partial \mathbf{h}}{\partial \tilde{\mathbf{x}}}(\tilde{\mathbf{s}}, \tilde{\mathbf{s}}) \right] = \tilde{G}^*. \quad (200)$$

Далее по аналогии с преобразованием (176) → (177) 2-й условный дифференциал функции Лагранжа приводится к *эрмитовой* форме:

$$\left[d^2 \mathcal{L}(\tilde{\mathbf{x}} = \tilde{\mathbf{s}}, \tilde{\bar{\mathbf{x}}} = \tilde{\bar{\mathbf{s}}}, \lambda, \bar{\lambda}) \right] = d\mathbf{x}^* \left(\tilde{\bar{\mathbf{S}}} \cdot \tilde{\boldsymbol{\theta}} + \tilde{\boldsymbol{\theta}}^* \cdot \tilde{\mathbf{S}} + 2\tilde{\mathbf{G}} \right) d\mathbf{x}. \quad (201)$$

Два крайних значения матричного параметра отвечают маргинальным положительным и отрицательным арифметическим корням из $\{\tilde{\bar{\mathbf{S}}} \cdot \tilde{\mathbf{S}}\}$:

$$\tilde{\theta}_{1,2} = \pm \tilde{\mathbf{S}} \cdot \sqrt{\tilde{\bar{\mathbf{S}}} \cdot \tilde{\mathbf{S}}}^{-1} = \pm \sqrt{\tilde{\bar{\mathbf{S}}} \cdot \tilde{\mathbf{S}}^{-1}} = \pm \sqrt{\tilde{\bar{\mathbf{S}}}^{-1} \cdot \tilde{\mathbf{S}}} \quad \text{при } \det \tilde{\mathbf{S}} \neq 0.$$

Выводы о характере условной стационарности делаются в соответствии с формулировками теоремы 8 анализом маргинальных функций:

$$F_1 = 2 \cdot (G + \sqrt{\tilde{\bar{\mathbf{S}}} \cdot \tilde{\mathbf{S}}}) \quad \text{и} \quad F_2 = 2 \cdot (G - \sqrt{\tilde{\bar{\mathbf{S}}} \cdot \tilde{\mathbf{S}}}).$$

Причём особо отметим, что при *внутреннем* характере ограничения в этих задачах условный экстремум целевой функции возможен даже с *одномерной* комплексной переменной, заданной на комплексной оси или на каком-либо её замкнутом интервале, т. е. для функции $y = y(\mathbf{x}, \bar{\mathbf{x}})$ при ограничении $h(\mathbf{x}, \bar{\mathbf{x}}) = 0$. Это весьма существенно отличает решение задач на условные экстремумы на вещественных и на комплексных координатных пространствах. Причина данного парадокса — удвоение числа степеней свободы для комплексных переменных.

§ 3.7. Экстремумы для функций от вещественных и комплексных сопряжённых переменных

В самом общем случае, когда целевая функция зависит от смешанных (вещественных и комплексных сопряжённых) переменных, процедура решения задачи на экстремум функции вида $y = y(\mathbf{r}, \mathbf{x}, \bar{\mathbf{x}}) = \psi(\mathbf{w})$ на первом этапе, т. е. нахождения точки или области её стационарности, вполне обычная:

$$\frac{dy}{d\mathbf{w}} = \left(\frac{\partial y}{\partial \mathbf{r}}, \frac{\partial y}{\partial \mathbf{x}}, \frac{\partial y}{\partial \bar{\mathbf{x}}} \right) = \mathbf{0} \quad \leftrightarrow \quad \left. \begin{array}{l} \frac{\partial y}{\partial \mathbf{r}} = \mathbf{0}, \\ \frac{\partial y}{\partial \bar{\mathbf{x}}} = \mathbf{0}. \end{array} \right\} \quad (202)$$

Отсюда вычисляется точка стационарности $\mathbf{w} = (\mathbf{s}_r, \mathbf{s}_x, \bar{\mathbf{s}}_x)$.

На втором этапе решения, т. е. при идентификации стационарности в данной точке, предварительно находят *симметричную комплексную* матрицу Гессе в нижеуказанной 9-ти блочной структуре:

$$\frac{d^2 y}{d\mathbf{w}' d\mathbf{w}} = \begin{pmatrix} \frac{\partial^2 y}{\partial \mathbf{r}' \partial \mathbf{r}} & \frac{\partial^2 y}{\partial \mathbf{r}' \partial \mathbf{x}} & \frac{\partial^2 y}{\partial \mathbf{r}' \partial \bar{\mathbf{x}}} \\ \frac{\partial^2 y}{\partial \mathbf{x}' \partial \mathbf{r}} & \frac{\partial^2 y}{\partial \mathbf{x}' \partial \mathbf{x}} & \frac{\partial^2 y}{\partial \mathbf{x}' \partial \bar{\mathbf{x}}} \\ \frac{\partial^2 y}{\partial \bar{\mathbf{x}}' \partial \mathbf{r}} & \frac{\partial^2 y}{\partial \bar{\mathbf{x}}' \partial \mathbf{x}} & \frac{\partial^2 y}{\partial \bar{\mathbf{x}}' \partial \bar{\mathbf{x}}} \end{pmatrix} = \begin{pmatrix} S_{rr} & A'_{xr} & \bar{A}'_{rx} \\ A_{xr} & S_{xx} & \bar{G}_{xx} \\ \bar{A}_{rx} & G_{xx} & \bar{S}_{xx} \end{pmatrix} = \left(\frac{d^2 y}{d\mathbf{w}' d\mathbf{w}} \right)' \quad (203)$$

Коэффициенты её векового (характеристического) уравнения суть вещественные числа. Оно даёт n_r вещественных и $2n_x$ комплексных сопряжённых собственных значений μ_j . Согласно адекватной аналогии (см. § 3.1), преобразование матрицы Гессе (203) к диагональной форме осуществляет некая комплексная *адекватно ортогональная* матрица R , имеющая такую же 9-ти блочную структуру. Для упрощения процесса анализа знакоопределённости комплексной матрицы Гессе (203) нужно её привести неким модальным преобразованием R_{12} к промежуточной 2-х клеточной форме с переходом в базис $\tilde{E}_2 = R_{12} \cdot \tilde{E}_1$:

$$\left(\frac{d^2 y}{d\mathbf{w}' d\mathbf{w}} \right)_{\tilde{E}_2} = \begin{pmatrix} S_{11} & Z & \bar{Z} \\ Z & S & \bar{G} \\ Z & G & \bar{S} \end{pmatrix} = \left(\frac{d^2 y}{d\mathbf{w}' d\mathbf{w}} \right)'_{\tilde{E}_2} \quad (204)$$

где $S_{11} = S'_{11}$ составляет 1-ю клетку, $S = S'$, $G = G^*$ образуют 2-ю клетку.

Далее квадратичная форма, соответствующая *новой* матрице Гессе, разбивается на две независимые квадратичные формы по 2-м клеткам — вещественной и комплексной:

$$d^2 y = d\mathbf{w}' \frac{d^2 y}{d\mathbf{w}' d\mathbf{w}} d\mathbf{w} = d\mathbf{r}_1' \frac{\partial^2 y}{\partial \mathbf{r}_1' \partial \mathbf{r}_1} d\mathbf{r}_1 + \left(d\mathbf{x}_1' \frac{\partial^2 y}{\partial \mathbf{x}_1' \partial \mathbf{x}_1} d\mathbf{x}_1 + 2d\mathbf{x}_1^* \frac{\partial^2 y}{\partial \mathbf{x}_1^* \partial \mathbf{x}_1} d\mathbf{x}_1 + d\bar{\mathbf{x}}_1' \frac{\partial^2 y}{\partial \bar{\mathbf{x}}_1' \partial \bar{\mathbf{x}}_1} d\bar{\mathbf{x}}_1 \right) = \quad (205)$$

$$= d\mathbf{r}_1' S_{11} d\mathbf{r}_1 + \left(d\mathbf{x}_1' S d\mathbf{x}_1 + 2d\mathbf{x}_1^* G d\mathbf{x}_1 + d\bar{\mathbf{x}}_1' \bar{S} d\bar{\mathbf{x}}_1 \right). \quad (206)$$

Затем квадратичная форма 2-й (комплексной) клетки матрицы Гессе преобразуется в эрмитову форму с использованием ранее изложенной в общем виде процедуры (§ 3.4). В итоге получается представление d^2y в точке стационарности $\mathbf{w} = (\mathbf{s}_r, \mathbf{s}_x, \bar{\mathbf{s}}_x)$ в виде алгебраической суммы из квадратичной и эрмитовой форм; причём эрмитова форма — с эффектом расширения 2-го дифференциала (§§ 3.3, 3.4):

$$d^2y = d\mathbf{r}_1' S_{11} d\mathbf{r}_1 + d\mathbf{x}_1^* \left(\bar{\mathbf{S}} \cdot \boldsymbol{\theta} + \boldsymbol{\theta}^* \cdot \mathbf{S} + 2\mathbf{G} \right) d\mathbf{x}_1 = \quad (207)$$

$$= d\mathbf{r}_1' S_{11} d\mathbf{r}_1 + d\mathbf{x}_1^* \mathbf{F}(\boldsymbol{\theta}, \boldsymbol{\theta}^*) d\mathbf{x}_1. \quad (208)$$

Нижеуказанные маргинальные эрмитовы матрицы-функции H_1 и H_2 , содержащие в своей 2-х клеточной структуре полученные ранее в § 3.4 маргинальные матричные функции F_1 и F_2 , имеют ту же 2-х клеточную форму, удобную для оценки знакоопределённости 2-го дифференциала целевой функции d^2y в точке её стационарности $\mathbf{w} = (\mathbf{s}_r, \mathbf{s}_x, \bar{\mathbf{s}}_x)$, а, следовательно, и идентификации характера её возможного экстремума 2-го уровня:

$$H_1 = \begin{pmatrix} S_{11} & Z \\ Z & F_1 \end{pmatrix}, \quad H_2 = \begin{pmatrix} S_{11} & Z \\ Z & F_2 \end{pmatrix}. \quad (209, 210)$$

Таким образом, полный анализ вещественных целевых функций, задаваемых на комплексных и смешанных вещественно-комплексных координатных пространствах, с целью нахождения и идентификации характера их безусловных и условных экстремумов может выполняться достаточно наглядно и довольно эффективно методами, изложенными в данной главе.

Глава 4. Применение аналитической оптимизации в общей и линейной алгебре

§ 4.1. Генеральное неравенство для средних величин

В общей и линейной алгебре применяется ряд неравенств типа [45]: $f_1(\mathbf{x}) \geq f_2(\mathbf{x})$, $f_3(\mathbf{x}_1, \mathbf{x}_2) \geq f_4(\mathbf{x}_1, \mathbf{x}_2)$, $f_5(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) \geq f_6(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ и т. д. В них обычно заключается какой-либо алгебро-геометрический смысл. Доказательство этих, по сути, функциональных неравенств возможно осуществлять, в принципе, с применением дифференциального метода поиска и идентификации безусловного экстремума целевой функции от одной или нескольких векторных переменных. Например, для полного доказательства первого из указанных неравенств необходимо пройти следующие этапы:

- 1) доказать, что некоторое $\langle \mathbf{b} \rangle$ есть область *стационарности* функции $f_{12}(\mathbf{x}) = f_1(\mathbf{x})/f_2(\mathbf{x})$, причём $f_{12}(\mathbf{b}) = 1$;
- 2) доказать, что эта стационарность есть *минимум* функции $f_{12}(\mathbf{x})$;
- 3) доказать, что этот минимум имеет *глобальный* характер.

Пункт 3, например, можно обосновать тем, что $f_{12}(\mathbf{x})$ непрерывно дифференцируема на области $\mathcal{T}^n \subset \mathcal{A}^n$ своего определения, а $f_{12}(\mathbf{b})$ — *единственно возможное* стационарное решение на этой области.

В качестве весьма важного примера рассмотрим далее формулировку и доказательство **генерального неравенства для средних величин**. Ранее оно уже было сформулировано и доказано в целом в монографии автора [27]. Как будет показано в последующих двух параграфах, это неравенство, помимо применения в теории средних, имеет приложения к теории алгебраических уравнений степени n и к теории $n \times n$ -матриц.

Напомним, что в конце § 1.9 в качестве довольно яркого примера было приведено доказательство классического неравенства Коши для средних арифметического и геометрического при $n = 2$ аналитическим способом — *дифференциальным методом*. Этот подход в самом общем виде применяется ниже для доказательства генерального неравенства в форме иерархии полного ряда средних величин.

В теории матриц весьма важны *положительно определённые* ранга n симметричные (эрмитовы) $n \times n$ -матрицы B . Они имеют *положительные* разнообразные, но связанные друг с другом скалярные инварианты. Это *характеристические коэффициенты* $k(B, t)$ — суммы детерминантов диагональных $t \times t$ -миноров B ; это *характеристические следы* $\text{tr} B^\theta$. Они же суть **суммы Виета** порядка t и **суммы Варинга** порядка θ для собственных значений μ_i матрицы B — корней её векового уравнения. (Причём с положительным знаком слагаемых во всех этих суммах!)

Совокупность из n положительных чисел может рассматриваться либо абстрактно, либо конкретно, например, как множество корней x_i алгебраического уравнения степени n или как множество собственных значений μ_i для $n \times n$ -матрицы B (с их положительным спектром).

Для совокупности из n вещественных положительных чисел $\langle \mu_i \rangle$ определим специальные характеристики — *средние алгебраические* $\bar{m}_t$ (малые медианы) и *средние степенные* $\bar{M}_\theta$ (большие медианы):

$$\bar{m}_1 = \bar{M}_1 = \sum_{i=1}^n \mu_i / n = \text{tr } B / n, \quad (211, 212)$$

$$\bar{m}_t = \sqrt[t]{k_t(\mu_i) / C_n^t} = \sqrt[t]{k(B, t) / C_n^t}, \quad (213, 214)$$

$$\bar{M}_\theta = \sqrt[\theta]{h_\theta(\mu_i) / n} = \sqrt[\theta]{\text{tr } B^\theta / n}, \quad (215, 216)$$

где черта сверху обозначает усреднение; $k_t(\mu_i) \equiv k(B, t)$ — суммы Виета, $h_\theta(\mu_i) \equiv \text{tr } B^\theta$ — суммы Варинга, n — размер совокупности $\langle \mu_i \rangle$ или размер $n \times n$ -матрицы B , t и θ — порядки соответствующих средних величин, например, $\bar{m}_t$ и $\bar{M}_\theta$. Кроме того, C_n^t обозначают, как общепринято, биномиальные коэффициенты Ньютона. (Отметим, вместе с тем, что *среднее арифметическое* $\bar{m}_1 = \bar{M}_1$ есть пересечение множеств средних алгебраических $\bar{m}_t$ и средних степенных $\bar{M}_\theta$.)

Иногда в литературе $\bar{m}_t$ называются *средними симметрическими* (хотя, в принципе, все рассматриваемые средние по своей структуре — симметричные). Используемая здесь классификация средних величин исходит из их применения в алгебре и геометрии [27]. Заметим также, что некоторые элементы $\mu_i > 0$ могут совпадать между собой. Случай, когда некоторые из элементов нулевые, может представлять отдельный интерес для *знакополоуопределённых* матриц ($\mu_i \geq 0$).

Если же нулевые μ_i отсутствуют, то могут быть весьма полезными реверсивные аналоги малых и больших медиан, которые определяются как соответствующие *реверсивные* средние величины:

$$\bar{v}_1 = \bar{V}_1 = \left(\sum_{i=1}^n \mu_i^{-1} / n \right)^{-1} = \left(\text{tr } B^{-1} / n \right)^{-1}, \quad (217, 218)$$

$$\bar{v}_t = -\sqrt[t]{k_t(\mu_i^{-1}) / C_n^t} = -\sqrt[t]{k(B^{-1}, t) / C_n^t}, \quad (219, 220)$$

$$\bar{V}_\theta = -\sqrt[\theta]{h_\theta(\mu_i^{-1}) / n} = -\sqrt[\theta]{\text{tr } B^{-\theta} / n}. \quad (221, 222)$$

Фактически они получаются как обращённые средние от обратных исходных элементов μ_i^{-1} и также являются средними величинами. Например, с физической точки зрения их размерность инвариантна. (Отметим, вместе с тем, что *среднее геометрическое* $\bar{m}_n = \bar{V}_n$ есть пересечение множеств средних алгебраических $\bar{m}_t$ и их реверсивных аналогов, а *среднее гармоническое* $\bar{v}_1 = \bar{V}_1$ есть пересечение множеств реверсивных средних алгебраических $\bar{V}_t$ и средних степенных $\bar{V}_\theta$.)

Теорема 9 (о средних). Для какой-либо совокупности n вещественных положительных чисел $\langle x_i \rangle$, в которой хотя бы одно число отличается от другого, имеет место *генеральное неравенство для средних величин*, охватывающее всю область данной совокупности, а именно:

$$\max \langle x_i \rangle = \bar{M}_\infty > \dots > \bar{M}_\theta > \dots > \bar{M}_1 = \quad (223)$$

$$= \bar{m}_1 > \dots > \bar{m}_t > \dots > \bar{m}_n = \quad (224)$$

$$= \bar{v}_n > \dots > \bar{v}_t > \dots > \bar{v}_1 = \quad (225)$$

$$= \bar{V}_1 > \dots > \bar{V}_\theta > \dots > \bar{V}_\infty = \min \langle x_i \rangle \quad (226)$$

$$(t = 1, \dots, n; \theta = 1, \dots, \infty).$$

Знак равенства, причём сразу для всех этих средних величин, имеет место тогда и только тогда, когда $x_1 = \dots = x_n$. Если бы данная исходная совокупность $\langle x_i \rangle$ содержала $s = n - r$ нулевых чисел x_r , то тогда цепь неравенств вырождалась справа в нулевые равенства, начиная с $\bar{m}_{r+1}$, а слева все средние величины оставались ненулевыми.

Например, для множества $\langle \mu_i \rangle$ любой спектрально положительной $n \times n$ -матрицы B предельные медианы совпадают с её экстремальными собственными значениями! Поэтому, например, $\max \langle \mu_i \rangle$ называется в литературе [25] *спектральным радиусом* матрицы. В пределе имеем:

$$\max \langle \mu_i \rangle = \lim_{\theta \rightarrow \infty} \bar{M}_\theta, \quad (227)$$

$$\min \langle \mu_i \rangle = \lim_{\theta \rightarrow \infty} \bar{V}_\theta. \quad (228)$$

Далее рассмотрим полное доказательство сформулированного выше генерального неравенства для средних величин в целом и его анализ. Для этого применим аналитический дифференциальный метод поиска и идентификации характера экстремума целевых функций от нескольких скалярных переменных. (Основная идея этого доказательства уже была продемонстрирована в решениях задач на примерах 4 и 5 из § 1.9.)

Доказательство. Определим 8 целевых скалярных функций g и R , f и F для разности и для отношения соответствующих пар средних величин исходя из некоей совокупности вещественных положительных чисел x_i (где $i = 1, n$), задаваемых также геометрически в $\mathcal{A}^n$ радиус-вектором $\mathbf{x} = (x_1, \dots, x_n)$ в 1-м квадранте аффинной системы координат $\tilde{E}$:

$$r \left\{ \begin{smallmatrix} t \\ t+1 \end{smallmatrix} \right\} (\mathbf{x}) = \bar{m}_t(\mathbf{x}) - \bar{m}_{t+1}(\mathbf{x}), \quad (229)$$

$$r \left\{ \begin{smallmatrix} 1 \\ n \end{smallmatrix} \right\} (\mathbf{x}) = \bar{m}_1(\mathbf{x}) - \bar{m}_n(\mathbf{x}), \quad (230)$$

$$f \left\{ \begin{smallmatrix} t \\ t+1 \end{smallmatrix} \right\} (\mathbf{x}) = \bar{m}_t(\mathbf{x}) / \bar{m}_{t+1}(\mathbf{x}), \quad (231)$$

$$f \left\{ \begin{smallmatrix} 1 \\ n \end{smallmatrix} \right\} (\mathbf{x}) = \bar{m}_1(\mathbf{x}) / \bar{m}_n(\mathbf{x}), \quad (232)$$

$$R \left\{ \begin{smallmatrix} \theta+1 \\ \theta \end{smallmatrix} \right\} (\mathbf{x}) = \bar{M}_{\theta+1}(\mathbf{x}) - \bar{M}_\theta(\mathbf{x}), \quad (233)$$

$$R \left\{ \begin{smallmatrix} \theta \\ 1 \end{smallmatrix} \right\} (\mathbf{x}) = \bar{M}_\theta(\mathbf{x}) - \bar{M}_1(\mathbf{x}), \quad (234)$$

$$F \left\{ \begin{smallmatrix} \theta+1 \\ \theta \end{smallmatrix} \right\} (\mathbf{x}) = \bar{M}_{\theta+1}(\mathbf{x}) / \bar{M}_\theta(\mathbf{x}), \quad (235)$$

$$F \left\{ \begin{smallmatrix} \theta \\ 1 \end{smallmatrix} \right\} (\mathbf{x}) = \bar{M}_\theta(\mathbf{x}) / \bar{M}_1(\mathbf{x}). \quad (236)$$

Функции g и R , f и F имеют *общее* и *единственное* стационарное значение — соответственно 0 и 1 на заданной области в $\mathcal{A}^n$ своего определения с аргументом-решением в форме центрального луча $\langle \mathbf{b} \rangle$ — биссектрисы 1-го квадранта в аффинном базисе $\tilde{E}$ (это геометрическое место точек равенства всех положительных координат). Такое решение отвечает нулевым градиентам всех 8 целевых функций:

$$g'(\mathbf{b}) = f'(\mathbf{b}) = R'(\mathbf{b}) = F'(\mathbf{b}) = \mathbf{0}, \quad (237)$$

где $\mathbf{b}$ — любая точка этой биссектрисы, т. е. $x_1 = \dots = x_n = b$. В этих же точках имеем стационарные значения целевых функций:

$$g(\mathbf{b}) = R(\mathbf{b}) = 0, \quad f(\mathbf{b}) = F(\mathbf{b}) = 1. \quad (238)$$

Это суть минимумы всех целевых функций, так как их матрицы Гессе на биссектрисе $\langle \mathbf{b} \rangle$ положительно полуопределённые ранга $(n - 1)$:

$$\begin{aligned} g'' \left\{ \begin{smallmatrix} 1 \\ n \end{smallmatrix} \right\} (\mathbf{b}) &= (n-1) \cdot g'' \left\{ \begin{smallmatrix} t \\ t+1 \end{smallmatrix} \right\} (\mathbf{b}) = b \cdot f'' \left\{ \begin{smallmatrix} 1 \\ n \end{smallmatrix} \right\} (\mathbf{b}) = b \cdot (n-1) \cdot f'' \left\{ \begin{smallmatrix} t \\ t+1 \end{smallmatrix} \right\} (\mathbf{b}) = \\ &= R'' \left\{ \begin{smallmatrix} \theta+1 \\ \theta \end{smallmatrix} \right\} (\mathbf{b}) = \frac{1}{\theta-1} \cdot R'' \left\{ \begin{smallmatrix} \theta \\ 1 \end{smallmatrix} \right\} (\mathbf{b}) = b \cdot F'' \left\{ \begin{smallmatrix} \theta+1 \\ \theta \end{smallmatrix} \right\} (\mathbf{b}) = \frac{b}{\theta-1} \cdot F'' \left\{ \begin{smallmatrix} \theta \\ 1 \end{smallmatrix} \right\} (\mathbf{b}) = \\ &= \frac{n \cdot I - It}{n^2 \cdot b} = G, \end{aligned} \quad (239—246)$$

где It определяется как *тотально-единичная* матрица, все n^2 элементов которой равны 1. Детерминанты главных миноров G порядка $g < n$:

$$\left(\frac{1}{n \cdot b} \right)^g \cdot \frac{n-g}{n} > 0. \quad (247)$$

Нетрудно видеть, что матрицы Гессе всех 8-ми целевых функций вырождаются вдоль биссектрисы $\langle \mathbf{b} \rangle$ — линейного подпространства размерности 1. Функции f и F , т. е. отношений средних величин, на биссектрисе постоянны и равны 1. Можно легко показать, что они тоже постоянны и при этом *больше* 1 на любом другом центральном луче, исходящем из начала координат. Минимум *всех* функций глобальный! Любые центральные лучи задаются, например, через координату x_1 :

$$\left. \begin{aligned} x_2 &= k_2 \cdot x_1, \\ \dots\dots\dots \\ x_i &= k_i \cdot x_1, \\ \dots\dots\dots \\ x_n &= k_n \cdot x_1. \end{aligned} \right\} k_i > 0 \quad (k_1 = 1) \quad (248)$$

С учётом полученных стационарных (минимальных) значений всех целевых функций на биссектрисе $\langle \mathbf{b} \rangle$, получаем простые соотношения между значениями на ней их матриц Гессе:

$$r'', f'' \left\{ \begin{matrix} t \\ t+m \end{matrix} \right\}(\mathbf{b}) = m \cdot r'', f'' \left\{ \begin{matrix} t \\ t+1 \end{matrix} \right\}(\mathbf{b}); \quad (249)$$

$$R'', F'' \left\{ \begin{matrix} \theta+m \\ \theta \end{matrix} \right\}(\mathbf{b}) = m \cdot R'', F'' \left\{ \begin{matrix} \theta+1 \\ \theta \end{matrix} \right\}(\mathbf{b}). \quad (250)$$

Анализ показывает тут следующее. Во-первых, на биссектрисе $\langle \mathbf{b} \rangle$ матрицы Гессе отношений соседних средних величин не зависят от порядка t или θ ; во-вторых, они же изменяются аддитивно с ростом интервала между этими порядками; в-третьих, они же совпадают для всех функций отношений между соседними средними степенными и отношения между средним арифметическим и средним геометрическим. Для функций отношений соседних средних алгебраических эта матрица Гессе делится равномерно на $(n - 1)$ равных частей.

Причём самое парадоксальное заключается в том, что матрица Гессе целевой функции F отношения между средним степенным и средним арифметическим на биссектрисе $\langle \mathbf{b} \rangle$, в силу (250), неограниченно растёт с ростом порядка θ . Хотя при $\theta \rightarrow \infty$, в силу (227), эта же функция F стремится к дроби $x_{\max}/\bar{M}_1$, изменяется непрерывно и на биссектрисе равна в точности 1 (т. е. минимуму). Кроме того, матрица Гессе целевой функции F отношения между соседними средними степенными на биссектрисе $\langle \mathbf{b} \rangle$, в силу (250), даже при $\theta \rightarrow \infty$ сохраняет постоянное значение. Хотя, в силу (227), эта же функция F стремится к 1 независимо от аргумента, т. е. к константе, для которой и градиент, и матрица Гессе суть нулевые. Эти, казалось бы, противоречивые факты объясняются влиянием соотношения бесконечно малого (отклонения аргумента от биссектрисы) и бесконечно большого (параметра $\theta \rightarrow \infty$). Вследствие этого факта в окрестности биссектрисы $\langle \mathbf{b} \rangle$ при $\theta \rightarrow \infty$ матрица Гессе терпит разрыв и становится вне её нулевой. В свою очередь, функция $F \left\{ \begin{matrix} \theta \\ 1 \end{matrix} \right\}(x)$ при $\theta \rightarrow \infty$ имеет постоянное значение 1, но с точностью до некоей бесконечно малой зависит от аргумента, принимая абсолютный минимум 1 как раз на биссектрисе $\langle \mathbf{b} \rangle$, где функция $F \left\{ \begin{matrix} \theta+1 \\ \theta \end{matrix} \right\}(x)$ сразу принимает это минимальное значение. Отсюда отмеченные парадоксы, заключающиеся в необычном экстремальном поведении функций F .

Более наглядным образом указанные необычные закономерности можно продемонстрировать на модельных функциях от одного скалярного аргумента. Например, пусть $n = 2$, $x_1 = 1$, $x_2 = x > 0$ или $x_2 = 1/x > 0$. Имеем:

$$F_1 \left\{ \frac{\theta+1}{\theta} \right\} (x) = \theta+1 \sqrt{\frac{1+x^{\theta+1}}{2}} / \sqrt{\frac{1+x^{\theta}}{2}} = F_1 \left\{ \frac{\theta+1}{\theta} \right\} (1/x), \text{ где } \theta \geq 1;$$

$$F_2 \left\{ \frac{\theta}{1} \right\} (x) = \theta \sqrt{\frac{1+x^{\theta}}{2}} / \frac{1+x}{2} = F_2 \left\{ \frac{\theta}{1} \right\} (1/x), \text{ где } \theta > 1.$$

Здесь x играет роль аргумента, но и, вместе с тем, экстремального элемента из имеющейся совокупности двух положительных чисел $\{1, x\}$.

При конечном степенном параметре θ имеем:

$$F_1(1) = F_2(1) = 1 = \min: F_2(x \neq 1) > F_1(x \neq 1) > 1;$$

$$\frac{dF_1}{dx}(1) = \frac{dF_2}{dx}(1) = 0;$$

$$\frac{d^2 F_1}{dx^2}(1) = \frac{1}{4} > 0, \quad \frac{d^2 F_2}{dx^2}(1) = \frac{\theta-1}{4} > 0 \left(\frac{d^2 F_2}{dx^2}(x) \geq \frac{d^2 F_1}{dx^2}(x) > 0 \right).$$

При $\theta \rightarrow \infty$ и $\alpha \rightarrow +0$ имеем:

$$\lim_{\theta \rightarrow \infty} F_1(x) = 1 = \min, \quad \lim_{\theta \rightarrow \infty} F_2 \begin{pmatrix} x > 1 \\ x = 1 \\ x < 1 \end{pmatrix} = \begin{pmatrix} 2x / (1+x) \\ 1 = \min \\ 2 / (1+x) \end{pmatrix};$$

$$\lim_{\theta \rightarrow \infty} \frac{dF_1}{dx}(x) = 0, \quad \lim_{\theta \rightarrow \infty} \frac{dF_2}{dx} \begin{pmatrix} x = 1 + \alpha \\ x = 1 \\ x = 1 - \alpha \end{pmatrix} = \begin{pmatrix} +1/2 \\ 0 \\ -1/2 \end{pmatrix};$$

$$\lim_{\theta \rightarrow \infty} \frac{d^2 F_1}{dx^2} \begin{pmatrix} x > 1 \\ x = 1 \\ x < 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1/4 > 0 \\ 0 \end{pmatrix}, \quad \lim_{\theta \rightarrow \infty} \frac{d^2 F_2}{dx^2} \begin{pmatrix} x = 1 + \alpha \\ x = 1 \\ x = 1 - \alpha \end{pmatrix} = \begin{pmatrix} 0 \\ (\theta-1)/4 \rightarrow \infty > 0 \\ 0 \end{pmatrix},$$

(В последнем случае имеем *маргинальный* минимум — § 1.3.)

Ввиду разрыва матрицы Гессе F'' в окрестности биссектрисы можно сделать вывод, что трёхвалентная симметричная матрица F''' третьих производных при $\theta \rightarrow \infty$ на биссектрисе должна быть бесконечной, но только в отрицательной области. Отметим также, что для аналогичных целевых функций реверсивных средних величин все вышеизложенные закономерности остаются и тут в силе, но знак перед матрицами Гессе меняется на противоположный, а формальный их вид сохраняется. То же происходит, если в функциях отношений средних величин поменять местами числитель и знаменатель. Таким образом, с учётом предельных формул (227), (228), доказательство и анализ генерального неравенства для средних величин полностью завершены. ***

В следующих параграфах рассматриваются отдельные возможности его применения в теории решения алгебраических уравнений степени n , в том числе вековых уравнений, а также в теории точных $n \times n$ -матриц. Процесс рассмотрения всех этих вопросов будет, в частности, связан с постановкой и решением разнообразных задач на экстремум.

§ 4.2. Экстремальные корни алгебраического уравнения

Генеральное неравенство для средних величин содержит в себе, как частные случаи, неравенство Коши для средних арифметического и геометрического и реверсивный аналог для средних гармонического и геометрического, неравенство Маклорена для средних алгебраических и реверсивный аналог и, кроме того, неравенство Гёльдера для средних арифметического и степенных и реверсивный аналог [45]. В целом оно объединяет их в единой, полной и завершённой цепи неравенств. Для *спектрально положительной* матрицы B (т. е. при $\mu_i > 0$) определим арифметическую, геометрическую и гармоническую медианы:

$$\bar{m}_1 = \text{tr} B / n = \bar{M}_1, \quad (251)$$

$$\bar{m}_n = \sqrt[n]{\det B} = \bar{v}_n, \quad (252)$$

$$\bar{v}_1 = (\text{tr} B^{-1} / n)^{-1} = \bar{V}_1. \quad (253)$$

Согласно генеральному неравенству средних, справедливы оценки:

$$\max \langle \mu_i \rangle \geq \text{tr} B / n \geq (\text{tr} B / n)^n \geq \det B \geq (\text{tr} B^{-1} / n)^{-n} \geq (\text{tr} B^{-n} / n)^{-1} \geq \min \langle \mu_i \rangle. \quad (254)$$

Ставится задача дискретной оптимизации: *вычислить максимальное и минимальное собственные значения положительной $n \times n$ -матрицы B : $\max \langle \mu_i \rangle$ и $\min \langle \mu_i \rangle$, т. е. для $n \times n$ -матрицы с априори положительным спектром собственных значений $\langle \mu_i \rangle$. Согласно (227) и (228), они оба являются как его предельными средними, так и его экстремальными элементами. Между ними в полной иерархии, согласно генеральному неравенству, располагаются *иерархические инварианты* матрицы B , вычисляемые по формулам (216), (251), (214), (252), (220), (253), (222). Все иерархические инварианты сохраняются при линейных модальных преобразованиях либо $n \times n$ -матрицы B , либо её базиса, т. е. активных и пассивных. Причём средним алгебраическим тут отвечают именно *иерархические геометрические инварианты, или нормы B* [27].*

В общем аспекте отметим: если все n корней $x_i > 0$ алгебраического уравнения степени n или собственных значений $\mu_i > 0$ $n \times n$ -матрицы B различаются между собой, то любые n инвариантов как функции от x_i или μ_i независимы между собой, а остальные инварианты выражаются через них. Соответственно, если только $g < n$ значений $x_i > 0$ или $\mu_i > 0$ различны, то любые g инвариантов как функции от x_i или μ_i , идущие подряд в цепи генерального неравенства, независимы между собой, а остальные ненулевые инварианты выражаются через них.

Как хорошо известно, корни x_i или μ_i различны между собой тогда и только тогда, когда *дискриминант* алгебраического уравнения, в том числе векового, отличается от нуля. Эта важнейшая характеристика определяется тождественным образом либо через квадрат произведения всевозможных разностей всех корней, либо с точностью до знака через *результант* многочленов $y(x)$ и $y'(x)$ [39]:

$$\Delta^2 = \prod_{i>k} (x_i - x_k)^2 = (-1)^{\frac{n \cdot (n-1)}{2}} \cdot \text{Res}[y(x), y'(x)] = \quad (255, 256)$$

$$= \left\{ \begin{array}{cccc} \det & & & \\ n & h_1 & h_2 & \dots & h_{n-1} \\ h_1 & h_2 & h_3 & \dots & h_n \\ h_2 & h_3 & h_4 & \dots & h_{n+1} \\ \dots & \dots & \dots & \dots & \dots \\ h_{n-1} & h_n & h_{n+1} & \dots & h_{2n-2} \end{array} \right\} = \left\{ \begin{array}{cccc} \det & & & \\ n & \text{tr} B & \text{tr} B^2 & \dots & \text{tr} B^{n-1} \\ \text{tr} B & \text{tr} B^2 & \text{tr} B^3 & \dots & \text{tr} B^n \\ \text{tr} B^2 & \text{tr} B^3 & \text{tr} B^4 & \dots & \text{tr} B^{n+1} \\ \dots & \dots & \dots & \dots & \dots \\ \text{tr} B^{n-1} & \text{tr} B^n & \text{tr} B^{n+1} & \dots & \text{tr} B^{2n-2} \end{array} \right\}.$$

Значимость дискриминанта алгебраического уравнения даёт возможность установить: *простое* оно или нет (т. е. не содержит *кратные* корни или содержит их). В принципе, он также подлежит вычислению и через коэффициенты уравнения k_p , связанные линейно с суммами Варинга h_θ .

Если $\det B \neq 0$ и $B \neq cI$, то n инвариантов, идущие подряд в цепи генерального неравенства средних и взятые в своей степени t или θ , *линейно* независимы друг от друга, а остальные инварианты, взятые в своей степени t или θ , выражаются через них *линейными* формулами. Например, через n сумм Виета или n сумм Варинга можно в предельной последовательности вычислять экстремальные корни или собственные значения при условии, что их спектр положительный и разнообразен!

Именно на этой изначальной идее базируется излагаемый далее *предельный метод* вычисления экстремальных корней алгебраического уравнения степени n и также экстремальных собственных значений $n \times n$ -матрицы B (с их вещественным положительным спектром). Этот метод был ранее изложен в монографии [27]. Он и родственен ему по предельной идее **метод Лобачевского — Грегфе** [40], как затем будет отдельно показано, — параллельно, но каждый по своему, выполняет основное назначение: последовательное приближение вычисляемого результата к предельной характеристике. Естественно, что в конце 1-го этапа вычислений после сокращения в многочлене алгебраического уравнения $y(x)$ одночлена или одночленов $(x - x_i)$ с найденным *точным* экстремальным корнем x_i процесс можно продолжить дальше вплоть до вычисления всех корней.

Сразу же отметим, что вначале целесообразно отделить фрагменты алгебраического уравнения с кратными корнями, если они имеются, и вычислить затем последние. При этом исходное уравнение разбивается на ряд простых (т. е. без кратных корней) алгебраических уравнений меньшей степени и, в частности, степени 1. В процессе используется алгоритм Евклида с целью вычисления *наибольшего общего делителя* многочлена уравнения и его 1-й производной (при необходимости он повторяется). Таким образом, исходное *сложное* алгебраическое уравнение всегда легко приводится к тождественной в целом системе простых уравнений.

Кроме того, для большего удобства в дальнейших вычислениях алгебраическое уравнение (с положительными корнями) целесообразно представить в *знакопередающей* форме:

$$y(x) = x^n - k_1 \cdot x^{n-1} + \dots + (-1)^t \cdot k_t \cdot x^{n-t} + \dots + (-1)^n \cdot k_n = 0. \quad (257)$$

В аналогичной *знакопередающей* форме целесообразно представлять и вековое уравнение для спектрально положительной $n \times n$ -матрицы B :

$$\det(B - \mu \cdot I) = (-\mu)^n + \text{tr} B \cdot (-\mu)^{n-1} + \dots + k(B, t) \cdot (-\mu)^{n-t} + \dots + \det B = 0. \quad (258)$$

В данной форме записи (при $x_i > 0$ или $\mu_i > 0$) все коэффициенты алгебраического уравнения суть обязательно *положительные* числа. Причём в уравнении (257) k_t — суммы Виета порядка t ; в уравнении (258) $k(B, t)$ — суммы детерминантов C_n^t диагональных $t \times t$ -миноров матрицы, но без чередования знака перед ними. Все суммы Варинга h_θ или все характеристические следы матрицы $\text{tr} B^\theta$ тоже *положительные*.

Для лучшей реализации *предельного метода на основе генерального неравенства средних величин* необходимо осуществить запуск процесса вычисления средних степенных в цепи (223). При этом исходно можно располагать либо n суммами Виета, т. е. коэффициентами уравнения в форме (257), (258); либо первыми n суммами Варинга, т. е. $h_\theta(\mu_i) \equiv \text{tr} B^\theta$, которые как множества взаимно однозначно связаны друг с другом:

$$\langle k_1, k_2, \dots, k_n \rangle \leftrightarrow \langle h_1, h_2, \dots, h_n \rangle.$$

Суммы Виета k_t и суммы Варинга h_θ суть симметричные и степенные многочлены от корней уравнения. Абстрактно любые симметричные и степенные суммы, как известно [39, 63], связаны **формулами Ньютона**:

$$h_t + \sum_{j=1}^{t-1} (-1)^j \cdot k_j \cdot h_{t-j} + (-1)^t \cdot t \cdot k_t = 0, \quad (t = \overline{1, n}), \quad (259)$$

$$h_\theta + \sum_{j=1}^n (-1)^j \cdot k_j \cdot h_{\theta-j} = 0. \quad (\theta > n). \quad (260)$$

Исторически они были открыты Варингом. Отсюда легко получаются **рекуррентные формулы Варинга прямого и обратного типа**:

$$k_t = \frac{1}{t} \cdot \sum_{j=1}^t (-1)^{j-1} \cdot h_j \cdot k_{t-j}, \quad (t = \overline{1, n}), \quad (261)$$

$$h_t = \sum_{j=1}^{t-1} (-1)^{j-1} \cdot k_j \cdot h_{t-j} + (-1)^{t-1} \cdot t \cdot k_t, \quad (t = \overline{1, n}), \quad (262)$$

$$h_\theta = \sum_{j=1}^n (-1)^{j-1} \cdot k_j \cdot h_{\theta-j}, \quad (\theta > n). \quad (263)$$

Формулы прямого типа при $t \leq n$ представляются в явном виде (264):

$$k_t = \frac{1}{t!} \cdot \det \left\{ \begin{array}{ccccc} h_1 & 1 & 0 & \dots & 0 \\ h_2 & h_1 & 2 & \dots & 0 \\ h_3 & h_2 & h_1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ h_{t-2} & h_{t-3} & h_{t-4} & \dots & 0 \\ h_{t-1} & h_{t-2} & h_{t-3} & \dots & t-1 \\ h_t & h_{t-1} & h_{t-2} & \dots & h_1 \end{array} \right\}.$$

Формулы обратного типа при $t \leq n$ представляются в явном виде (265):

$$h_t = \det \begin{pmatrix} +1 \cdot k_1 & -1 & 0 & \dots & 0 \\ -2 \cdot k_2 & +k_1 & -1 & \dots & 0 \\ +3 \cdot k_3 & -1 \cdot k_2 & +k_1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ (-1)^{t-3} \cdot (t-2) \cdot k_{t-2} & (-1)^{t-4} \cdot k_{t-3} & (-1)^{t-5} \cdot k_{t-4} & \dots & 0 \\ (-1)^{t-2} \cdot (t-1) \cdot k_{t-1} & (-1)^{t-3} \cdot k_{t-2} & (-1)^{t-4} \cdot k_{t-3} & \dots & -1 \\ (-1)^{t-1} \cdot t \cdot k_t & (-1)^{t-2} \cdot k_{t-1} & (-1)^{t-3} \cdot k_{t-2} & \dots & +k_1 \end{pmatrix}.$$

Разумеется, вид формул (259)—(265) отвечает знакопередающей форме алгебраического уравнения типа (257), (258).

Суммы Виета k_t , где t принимает значения от 2-х до n , согласно (264), суть *изопараметрические многочлены* валентности 2 от $h_j^d = \text{tr}^d B^j$ (см. § 1.6.1):

$$k_t = k(B, t) = \sum_{Q(t, q)} \frac{(-1)^{t+q}}{d_1! d_2! \dots d_m!} \cdot \left(\frac{h_{j_1}}{j_1} \right)^{d_1} \cdot \left(\frac{h_{j_2}}{j_2} \right)^{d_2} \cdot \dots \cdot \left(\frac{h_{j_m}}{j_m} \right)^{d_m} =$$

$$= \sum_{Q(t, q)} N_{\substack{d_1, d_2, \dots, d_m \\ j_1, j_2, \dots, j_m}}^{(t, q)} \cdot (h_{j_1})^{d_1} \cdot (h_{j_2})^{d_2} \cdot \dots \cdot (h_{j_m})^{d_m},$$

$$\text{где: } \left. \begin{array}{l} d_1 + d_2 + \dots + d_m = q \leq t, \\ j_1 \cdot d_1 + j_2 \cdot d_2 + \dots + j_m \cdot d_m = t, \end{array} \right\} (t = \text{const}).$$

Причём коэффициенты N подчиняются простым соотношениям:

$$\sum_{Q(t, q)} N_{\substack{d_1, d_2, \dots, d_m \\ j_1, j_2, \dots, j_m}}^{(t, q)} = t!; \quad \sum_{Q(t, q)} N_{\substack{d_1, d_2, \dots, d_m \\ j_1, j_2, \dots, j_m}}^{(t, q)} \cdot (-1)^q = 0.$$

Коэффициенты k_t получаются здесь с *дополнительным* множителем $(-1)^t$, что отвечает алгебраическому уравнению в знакопередающей форме (257), (258). Коэффициенты N целесообразно тут сравнить по структуре с аналогичными коэффициентами в изопараметрических многочленах для дифференциалов $d^p f$ той же валентности 2 (т. е. количества независимых индексов) — см. § 1.6.1.

Аналогичным образом суммы Варинга $h_t = \text{tr} B^t$, где t пробегает значения от 2-х до n , согласно (265), также суть *изопараметрические многочлены* валентности 2, но от $k_j^d = k^d(B, j)$.

Если $\text{rang } B = r < n$, то $k(B, t > r) = 0$. Тогда последующие $h_\theta = \text{tr} B^\theta$ при $\theta > r$ *линейно* зависимы только от первых r характеристических следов (или эквивалентно — только от первых r характеристических коэффициентов).

Рекуррентные формулы Варинга (262) и (263) позволяют запустить алгоритм вычисления средних степенных от положительных корней уравнения (в том числе векового), т. е. $\bar{M}_1, \bar{M}_2, \dots, \bar{M}_\theta$ — вплоть до сколь угодно большого θ , неуклонно приближаясь к конечной цели $x_{\max} > 0$ именно *снизу*, согласно цепи (223) генерального неравенства.

Более общо имеем предельные соотношения и неравенства:

$$\lim_{\theta \rightarrow \infty} \sqrt[\theta]{h_\theta \langle x_i \rangle / n} = x_{\max} = \lim_{\theta \rightarrow \infty} \sqrt[\theta]{h_\theta \langle x_i \rangle}, \quad (266)$$

$$\sqrt[\theta]{h_\theta \langle x_i \rangle / n} < x_{\max} < \sqrt[\theta]{h_\theta \langle x_i \rangle}. \quad (267)$$

В частности, исходно имеем: $k_1/n < x_{\max} < k_1$, $\text{tr}B/n < \mu_{\max} < \text{tr}B$. Очевидно, что скорость приближения результата к $x_{\max}$ тем больше, чем более отличаются корни между собой. Подставив в рекуррентную формулу (263) предельное значение $x_{\max}$ и сократив множитель $x_{\max}^{\theta-h}$, получаем исходное алгебраическое уравнение уже в виде тождества. На каком-то этапе вычисления обрываются из-за неминуемой ошибки округления. При этом упорядоченность последовательности итераций и неравенство (267) могут нарушаться, что должно останавливать процесс. Корень $x_{\max}$ идентифицируется проверкой по исходному уравнению. Заметим, что возможная кратность корня q только замедляет процесс вычисления, но не влияет на конечный результат, так как:

$$\lim_{\theta \rightarrow \infty} \sqrt[\theta]{q \cdot x_{\max}^\theta} = \lim_{\theta \rightarrow \infty} \sqrt[\theta]{q \cdot x_{\max}^\theta / n} = x_{\max}.$$

Корень $x_{\min} > 0$, в принципе, согласно (228), можно вычислять таким же образом, если использовать *инверсионную форму* алгебраического уравнения, т. е. поделив исходное уравнение на $(-x)^n$ и на старший коэффициент k_n , перейдя к аргументу $1/x$. Тогда имеем аналогичные инверсионные общие соотношения:

$$\lim_{\theta \rightarrow \infty} -\sqrt[\theta]{h_\theta \langle x_i^{-1} \rangle} = x_{\min} = \lim_{\theta \rightarrow \infty} -\sqrt[\theta]{h_\theta \langle x_i^{-1} \rangle / n}, \quad (268)$$

$$-\sqrt[\theta]{h_\theta \langle x_i^{-1} \rangle} < x_{\min} < -\sqrt[\theta]{h_\theta \langle x_i^{-1} \rangle / n}. \quad (269)$$

В частности, $k_n/k_{n-1} < x_{\min} < n \cdot k_n/k_{n-1}$, $\text{tr}^{-1}B^{-1} < \mu_{\min} < n \cdot \text{tr}^{-1}B^{-1}$.

Если корни уравнения — точные рациональные числа, то в процессе последовательного приближения с требуемой точностью у результата неизбежно проявится рациональная природа. Иррациональные корни вычисляются с заданной точностью. Процесс может быть продолжен и далее с вычислением предшествующих $x_{\max}$ корней в их иерархическом спектре вплоть до вычисления минимального корня.

Обратим внимание на то, что вышеизложенный метод, как и близкий ему по той же предельной идее метод Лобачевского — Греффе (1834 г.), имеет *глобальный* характер. Все расчётные параметры в нём строго предопределены. Кроме того, оба предельных метода базируются на возрастании вклада $x_{\max}$ (относительно прочих корней x_i) в суммах Варинга и в средних степенных с увеличением порядка θ .

Особо отметим, что используемый в изложенном предельном методе подход позволяет *выразить оба экстремальных корня алгебраического уравнения — с положительным спектром через его коэффициенты в виде итоговых предельных формул*. (Например, ряд Лагранжа не даёт возможности выразить формулой конкретные корни этого уравнения.)

$$x_{\max} = \lim_{\theta \rightarrow \infty} \sqrt[\theta]{\det \left\{ -K_{(n+\theta) \times (n+\theta)}^{(1)} \right\}}, \quad (270)$$

где квадратная матрица $K^{(1)}$ под знаком корня выражается так:

$$\left\{ \begin{array}{cccccccc} -1 \cdot k_1 & +1 & 0 & \dots & 0 & 0 & \dots & 0 & 0 \\ +2 \cdot k_2 & -k_1 & +1 & \dots & 0 & 0 & \dots & 0 & 0 \\ -3 \cdot k_3 & +k_2 & -k_1 & \dots & 0 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ (-1)^n n k_n & (-1)^{n-1} k_{n-1} & (-1)^{n-2} k_{n-2} & \dots & \dots & \dots & \dots & 0 & 0 \\ 0 & (-1)^n k_n & (-1)^{n-1} k_{n-1} & \dots & \dots & \dots & \dots & 0 & 0 \\ 0 & 0 & (-1)^n k_n & \dots & \dots & \dots & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & (-1)^{n-1} k_{n-1} & (-1)^{n-2} k_{n-2} & \dots & +1 & 0 \\ 0 & 0 & 0 & \dots & (-1)^n k_n & (-1)^{n-1} k_{n-1} & \dots & -k_1 & +1 \\ 0 & 0 & 0 & \dots & 0 & (-1)^n k_n & \dots & +k_2 & -k_1 \end{array} \right\}$$

(все нулевые элементы стоят в 2-х треугольниках размера θ и $n + \theta - 2$, верхнем и нижнем; все прочие элементы обязательно ненулевые).

$$x_{\min} = \lim_{\theta \rightarrow \infty} -\theta \sqrt{\det \left\{ -\frac{1}{k_n} \cdot K_{(n+\theta) \times (n+\theta)}^{(2)} \right\}}, \quad (271)$$

где квадратная матрица $K^{(2)}$ под знаком корня выражается так:

$$\left\{ \begin{array}{cccccccc} -1 \cdot k_{n-1} & +k_n & 0 & \dots & 0 & 0 & \dots & 0 & 0 \\ +2 \cdot k_{n-2} & -k_{n-1} & +k_n & \dots & 0 & 0 & \dots & 0 & 0 \\ -3 \cdot k_{n-3} & +k_{n-2} & -k_{n-1} & \dots & 0 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ (-1)^n k_n & (-1)^{n-1} k_1 & (-1)^{n-2} k_2 & \dots & \dots & \dots & \dots & 0 & 0 \\ 0 & (-1)^n & (-1)^{n-1} k_1 & \dots & \dots & \dots & \dots & 0 & 0 \\ 0 & 0 & (-1)^n & \dots & \dots & \dots & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & (-1)^{n-1} k_1 & (-1)^{n-2} k_2 & \dots & +k_n & 0 \\ 0 & 0 & 0 & \dots & (-1)^n & (-1)^{n-1} k_1 & \dots & -k_{n-1} & +k_n \\ 0 & 0 & 0 & \dots & 0 & (-1)^n & \dots & +k_{n-2} & -k_{n-1} \end{array} \right\}$$

(все нулевые элементы стоят в тех же самых 2-х треугольниках, а все прочие элементы также обязательно ненулевые). Отметим, что при транспонировании матриц $K^{(1)}$ и $K^{(2)}$ результаты не изменяются.

* * *

Последовательность скалярных характеристических коэффициентов $k(B, t) = k_t$ через формулы Ньютона (259) взаимнооднозначно связана с последовательностью характеристических следов $\text{tr} B^\theta = h_\theta$ той же длины — теоретически вплоть до порядка $n = \text{rang } B$ для несингулярных $n \times n$ -матриц B . (Для не вещественнозначных и знаконеопределённых матриц B характеристические коэффициенты и следы до их порядка n могут эпизодически обнуляться.) Решение каких-либо задач, связанных изначально с k_t , можно рассматривать исходя из значений h_θ ! Этот факт используется в теории матриц, например, в методе Леверье вычисления $k_t = k(B, t)$ через n следов $\text{tr} B^\theta$ ($t, \theta = \overline{1, n}$), где $\det B \neq 0$:

$$k(B, t) = \frac{1}{t} \cdot \sum_{j=1}^t (-1)^{j-1} \cdot k(B, t-j) \cdot \text{tr} B^j. \quad (272)$$

Это классическая **рекуррентная формула Варинга—Леверье**. Она является полным аналогом формулы (261). Отсюда по **методу Леверье** получается вековое уравнение матрицы B исходя из следов её степеней. (Данный метод вычисления коэффициентов векового уравнения весьма способствовал Леверье для его великого открытия *на кончике пера* планеты Нептун.) *Точные* вычисления всех $k(B, t)$ должны обрываться при порядке $t = n$. Соответственно исходное вековое уравнение для применения предельного метода вычисления собственных значений для знакоположительной B имеет порядок n . Ранее указанное замечание по поводу возможности преобразования исходного уравнения к простой форме, т. е. к форме без кратных корней, сохраняется.

Если же исходное алгебраическое уравнение (в том числе вековое) в принятой здесь знакочередующейся форме имеет вещественные, но *знаконеопределённые* корни и соответственно не все положительные коэффициенты k_j , то далее наиболее целесообразно сместить аргумент исходя из априорной границы или отрицательных, или положительных вещественных корней по известным **правилам Ролля—Маклорена**:

$$\left[-1 - d_1 \sqrt{-\min k_j} \right] = \Delta^{(-)} < x_i < \Delta^{(+)} = \left[-1 + d_2 \sqrt{-\min (-1)^j \cdot k_j} \right]. \quad (273, 274)$$

Здесь $\Delta^{(-)}$ и $\Delta^{(+)}$ — априорные границы вещественных отрицательных и положительных корней x_i , d_1 и d_2 — индексы первых отрицательных коэффициентов k_j и $(-1)^j \cdot k_j$. Для обеспечения наибольшей скорости сходимости предельного метода выбирают наименьший по абсолютной величине Δ . Если таковым оказывается $\Delta^{(-)}$, то аргумент x смещают в сторону его положительных значений с подстановкой в уравнение типа: $x = u + \Delta^{(-)} \rightarrow f(u) = 0$; $u_i = x_i - \Delta^{(-)} > 0$. Если таковым оказывается $\Delta^{(+)}$, то аргумент x смещают в сторону его отрицательных значений и затем меняют знак: $x = u + \Delta^{(+)}$, $v = -u \rightarrow f(v) = 0$; $v_i = -x_i + \Delta^{(+)} > 0$. В итоге в обоих случаях получают новое уравнение со всеми положительными коэффициентами и соответственно с предполагаемыми вещественными положительными корнями u_i или v_i .

Для матрицы B с её знаконеопределённым вещественным спектром собственных значений указанное преобразование смещения трактуется также или с применением $\Delta^{(-)} < 0$, или с применением $\Delta^{(+)} > 0$:

$$\begin{aligned} B_2 &= \{B_1 - \Delta^{(-)} \cdot I\}, \{B_1\}_i = B_1 - \mu_i \cdot I, \{B_2\}_i = B_2 - v_i \cdot I = \{B_1 - \Delta^{(-)} \cdot I - v_i \cdot I\} = \\ &= \{B - (v_i + \Delta^{(-)}) \cdot I\} \rightarrow v_i = \mu_i - \Delta^{(-)} > 0; \\ B_2 &= \{B_1 - \Delta^{(+)} \cdot I\}, \{B_1\}_i = B_1 - \mu_i \cdot I, \{B_2\}_i = B_2 - v_i \cdot I = \{B_1 - \Delta^{(+)} \cdot I - v_i \cdot I\} = \\ &= \{B - (v_i + \Delta^{(+)}) \cdot I\}; B_3 = -B_2 \rightarrow v_i = -(\mu_i - \Delta^{(+)}) = -\mu_i + \Delta^{(+)} > 0. \end{aligned}$$

Особо отметим: для вековых алгебраических уравнений вывод о вещественности и положительности их корней — собственных значений в ряде частных случаев можно сделать по элементам $n \times n$ -матрицы B . Априори вещественные собственные значения имеют вещественные *симметричные* матрицы $S = S'$ и мнимые *кососимметричные* матрицы $iK = -(iK)'$, где $K = -K'$ вещественные. Например, для вещественной $n \times n$ -матрицы B это могут быть её характеристические матрицы типа: $S_B = (B + B')/2$ и $K_B = (B - B')/2$ ($B = S_B + K_B$). В важном частном случае, когда $S_B \cdot K_B = K_B \cdot S_B \leftrightarrow B \cdot B' = B' \cdot B$, исходная $n \times n$ -матрица *нормальная*: $B \in \langle N \rangle$. Для неё S_N и K_N приводятся к диагональной форме совместно в одном и том же базисе. Поэтому их собственные значения в парных суммах дают таковые для исходной матрицы N . Следовательно, решая отдельно вековые уравнения для S_N и для $-iK_N$ (последнее обязательно биквадратное), можно получить по отдельности вещественные и сопряжённые мнимые части в парах комплексных сопряжённых собственных значений матрицы N . Далее только остаётся сделать подбор этих пар путём проверки на вековом уравнении. Такой же, по сути, подход, но с применением эрмитова транспонирования распространяется на комплексные нормальные матрицы $B \cdot B^* = B^* \cdot B$. Что же касается вещественных симметричных и комплексных эрмитовых матриц, то о положительности их априорно вещественных собственных значений судят по классическим правилам Сильвестра, рассмотренным ранее в § 1.9.

Таким образом, множество $n \times n$ -матриц, которые *априори* подходят для реализации предельного метода, включает в себя в самом общем случае вещественные нормальные матрицы и комплексные эрмитовы нормальные матрицы.

Альтернатива процедуре смещения знаконеопределённых корней в положительную область имеется. Для знаконеопределённой матрицы с вещественными собственными значениями μ_i может применяться операция возведения её в квадрат. Для алгебраического уравнения с вещественными корнями x_i может применяться однократная операция *квадрирования* (см. далее). Вследствие таких операций корни уравнения и собственные значения $n \times n$ -матрицы B также возводятся в квадрат, переходя при этом в положительную область. На конечной стадии остаётся подобрать их знаки проверкой по исходному уравнению.

Но вещественность корней *конкретного* алгебраического уравнения вообще — с вещественными коэффициентами устанавливают при $n < 4$ по его дискриминанту (255), а при $n > 4$, например, классическим **методом Штурма** [39] в априорном интервале аргумента $(-\infty, +\infty)$.

В заключение этого параграфа дадим небольшой *сравнительный анализ* предельного метода, основанного на генеральном неравенстве для средних величин, и классического метода Лобачевского — Греффе, с учётом простоты их реализации, объёма всех вычислений и скорости сходимости, например, к максимальному вещественному корню.

Обратим внимание на то, что при вычислении предельным методом максимального собственного значения μ_1 *спектрально положительной* $n \times n$ -матрицы B , согласно формулам (266) и (267), аналогично имеем:

$$\lim_{\theta \rightarrow \infty} \sqrt[\theta]{\text{tr} B^\theta / n} = \mu_{\max} = \lim_{\theta \rightarrow \infty} \sqrt[\theta]{\text{tr} B^\theta}, \quad \sqrt[\theta]{\text{tr} B^\theta / n} < \mu_{\max} < \sqrt[\theta]{\text{tr} B^\theta}. \quad (276)$$

Количество итераций, связанных с возведением в степень матрицы B , тут существенно сокращают, когда увеличение её степени достигают в процессе вычислений по схеме *квадрирования матрицы* [6]:

$$B \cdot B = B^2, \quad B^2 \cdot B^2 = B^4, \quad \dots, \quad B^{2^{d-1}} \cdot B^{2^{d-1}} = B^{2^d} \quad (2^d = \theta). \quad (277)$$

Следовательно, степень θ в такой схеме возрастает экспоненциально. Общее количество требуемых итераций d существенно сокращается. Однако видно, что объёмы вычислений при каждой итерации также существенно возрастают. Тут недостаточно простого повторения одной *примитивной* итерации (263) с увеличением каждый раз θ на единицу. С аналогичной ситуацией неминуемо сталкиваются именно тогда, когда, вычисляя корень алгебраического уравнения предельным методом, для эквивалентного сокращения общего количества итераций d используют операцию поэтапного *квадрирования уравнения*. На подобном приёме и предельной идее базируется **метод Лобачевского — Греффе**:

$$\left. \begin{aligned} y(x) &= 0, \\ y_1(x^2) &= y(x) \cdot y(-x) = 0, \\ y_2(x^4) &= y_1(x^2) \cdot y_1(-x^2) = 0, \\ &\dots\dots\dots \\ y_d(x^{2^d}) &= y_{d-1}(x^{2^{d-1}}) \cdot y_{d-1}(-x^{2^{d-1}}) = 0. \end{aligned} \right\} \quad (278)$$

В этом методе при знакопередающей форме уравнения имеем:

$$|x_{\max}| = \lim_{d \rightarrow \infty} \sqrt[d]{k_{1(d)}} = \lim_{\theta \rightarrow \infty} \sqrt[\theta]{\sum x_i^\theta}, \quad (279)$$

где $k_{1(d)}$ есть 1-й коэффициент d -го уравнения, $\theta = 2^d$.

Резкое возрастание здесь объёма попутных вычислений связано с необходимостью нахождения n коэффициентов нового уравнения после каждой операции квадрирования. Коэффициенты *последующего* квадрированного уравнения q_j вычисляются исходя из значений всех коэффициентов *предыдущего* уравнения k_j по формулам:

$$\left. \begin{aligned} q_1 &= k_1^2 - 2k_2, \\ q_2 &= k_2^2 - 2k_1k_3 + 2k_4, \\ q_3 &= k_3^2 - 2k_2k_4 + 2k_1k_5 - 2k_6, \\ &\dots\dots\dots \\ q_{n-3} &= k_{n-3}^2 - 2k_{n-2}k_{n-4} + 2k_{n-1}k_{n-5} - 2k_nk_{n-6}, \\ q_{n-2} &= k_{n-2}^2 - 2k_{n-1}k_{n-3} + 2k_nk_{n-4}, \\ q_{n-1} &= k_{n-1}^2 - 2k_nk_{n-2}, \\ q_n &= k_n^2. \end{aligned} \right\} \quad (280)$$

По смыслу операции квадрирования и по виду данных соотношений следует, что принимаемая исходная форма алгебраического уравнения (знакопостоянная или знакопеременная) на формулы типа (280) для коэффициентов не оказывает влияния.

Однако известно, что процесс квадрирования приводит к быстрому накоплению ошибки вычислений. С учётом сказанного, можно сделать вывод, что однозначного критерия в пользу выбора из 2-х вариантов предельного метода нет. Ответ на вопрос данного выбора лежит, скорее, в практической области вычислений, нежели в теоретической.

* * *

В общем плане отметим, что к степеням положительных чисел x_i также применимо генеральное неравенство для средних величин. Но при этом, например, для средних алгебраических производятся весьма неочевидные неравенства типа:

$$p\sqrt[p]{\bar{m}_t(x_1^p, \dots, x_n^p)} < q\sqrt[q]{\bar{m}_t(x_1^q, \dots, x_n^q)}, \quad (281)$$

$$p\sqrt[p]{\bar{v}_t(x_1^p, \dots, x_n^p)} > q\sqrt[q]{\bar{v}_t(x_1^q, \dots, x_n^q)}, \quad (282)$$

где $1 \leq p < q$, $n-1 \geq t \geq 1$ и хотя бы два элемента x_i различны, а количество ненулевых элементов больше t . Эти неравенства следуют из цепей (223) и (226) генерального неравенства для средних величин применительно к суммам Виета от $x_i^{\pm p}$ и от $x_i^{\pm q}$.

Также между прямыми и реверсивными средними алгебраическими имеется неочевидная взаимосвязь через среднее геометрическое:

$$\bar{m}_n^n = \bar{m}_t^t \cdot \bar{v}_{n-t}^{n-t} = \bar{m}_{n-t}^{n-t} \cdot \bar{v}_t^t = \bar{v}_n^n. \quad (283)$$

§ 4.3. Инверсия генерального неравенства

Обращение элементов x_i в совокупности положительных чисел $\langle x_i \rangle$ никак не влияет на их знаки. Но для элементов x_i^{-1} есть свои средние. Для совокупности n вещественных положительных чисел типа $\langle x_i^{-1} \rangle$, в которой хотя бы одно число отличается от другого числа, имеет место *инверсионная форма* генерального неравенства для средних величин (203)–(206), охватывающая всю область данной совокупности:

$$\max \langle x_i^{-1} \rangle = \bar{V}_\infty^{-1} > \dots > \bar{V}_\theta^{-1} > \dots > \bar{V}_1^{-1} = \quad (284)$$

$$= \bar{v}_1^{-1} > \dots > \bar{v}_t^{-1} > \dots > \bar{v}_n^{-1} = \quad (285)$$

$$= \bar{m}_n^{-1} > \dots > \bar{m}_t^{-1} > \dots > \bar{m}_1^{-1} = \quad (286)$$

$$= \bar{M}_1^{-1} > \dots > \bar{M}_\theta^{-1} > \dots > \bar{M}_\infty^{-1} = \min \langle x_i^{-1} \rangle \quad (287)$$

$$(t = 1, \dots, n; \theta = 1, \dots, \infty).$$

Все *инверсионные средние величины* получаются в результате простого обращения определяющих формул для средних величин (211)–(222). Данные неравенства могут применяться, например, при решении или анализе предельным методом *инверсионного алгебраического уравнения* с вычислением корня $x_{\min}$. Аналогично (257) представим это уравнение в знакочередующейся форме:

$$y(x^{-1}) = x^{-n} - \hat{k}_1 \cdot x^{-(n-1)} + \dots + (-1)^t \cdot \hat{k}_t \cdot x^{-(n-t)} + \dots + (-1)^n \cdot \hat{k}_n = 0, \quad (288)$$

где $\hat{k}_1 = k_{n-1} / k_n, \dots, \hat{k}_t = k_{n-t} / k_n, \dots, \hat{k}_n = k_1 / k_n$ — положительные коэффициенты уравнения (инверсионные коэффициенты). Они же суть положительные суммы Виета от корней этого уравнения, взаимно-однозначно связанные с их положительными суммами Варинга:

$$\hat{k}_t = k_t \langle x_i^{-1} \rangle, \quad \hat{h}_t = h_t \langle x_i^{-1} \rangle; \quad \langle \hat{k}_1, \hat{k}_2, \dots, \hat{k}_n \rangle \leftrightarrow \langle \hat{h}_1, \hat{h}_2, \dots, \hat{h}_n \rangle.$$

Инверсионные формулы Ньютона имеют вид, аналогичный (259), (260):

$$\hat{h}_t + \sum_{j=1}^{t-1} (-1)^j \cdot \hat{k}_j \cdot \hat{h}_{t-j} + (-1)^t \cdot t \cdot \hat{k}_t = 0, \quad (t = \overline{1, n}) \quad (289)$$

$$\hat{h}_\theta + \sum_{j=1}^n (-1)^j \cdot \hat{k}_j \cdot \hat{h}_{\theta-j} = 0. \quad (\theta > n) \quad (290)$$

Далее имеем инверсионные аналоги формул (266), (267), используемых в предельном методе вычисления максимального корня уравнения:

$$\lim_{\theta \rightarrow \infty} \sqrt[t]{h_\theta(x_i^{-1})} / n = x_{\min}^{-1} = \lim_{\theta \rightarrow \infty} \sqrt[t]{h_\theta(x_i^{-1})}; \sqrt[t]{h_\theta(x_i^{-1})} / n < x_{\min}^{-1} < \sqrt[t]{h_\theta(x_i^{-1})},$$

т. е. это те же самые формулы (268), (269) для минимального корня.

Отметим также, что малые и большие медианы для совокупностей положительных чисел $\langle x_i \rangle$ (или $\langle x_i^{-1} \rangle$) при $t = \overline{1, n}$ связаны друг с другом модифицированными формулами Ньютона:

$$C_{n-1}^{t-1} \cdot \bar{m}_t^t - C_n^{t-1} \cdot \bar{M}_1 \cdot \bar{m}_{t-1}^{t-1} + C_n^{t-2} \cdot \bar{M}_2^2 \cdot \bar{m}_{t-2}^{t-2} + \dots \\ \dots + (-1)^{t-1} \cdot C_n^1 \cdot \bar{M}_{t-1}^{t-1} \cdot \bar{m}_1 + (-1)^t \cdot \bar{M}_t^t = 0, \quad (291)$$

$$C_{n-1}^{t-1} \cdot \bar{v}_t^{-t} - C_n^{t-1} \cdot \bar{v}_1^{-1} \cdot \bar{v}_{(t-1)}^{-(t-1)} + C_n^{t-2} \cdot \bar{v}_2^{-2} \cdot \bar{v}_{t-2}^{-(t-2)} + \dots \\ \dots + (-1)^{t-1} \cdot C_n^1 \cdot \bar{v}_{t-1}^{-(t-1)} \cdot \bar{v}_1^{-1} + (-1)^t \cdot \bar{v}_t^{-t} = 0; \quad (292)$$

$$\text{где } C_{n-1}^{t-1} = C_n^{t-1} - C_n^{t-2} + \dots + (-1)^{t-2} \cdot C_n^1 + (-1)^{t-1}.$$

§ 4.4. Полные требования к коэффициентам алгебраического уравнения для положительности всех его корней

Ставится следующая задача: *каким необходимым и достаточным требованиям должны удовлетворять коэффициенты вещественного алгебраического уравнения степени n (в знакочередующейся форме), для вещественности и положительности или, ещё более широко, — для вещественности его корней?* Оказывается, и эта алгебраическая проблема в самом общем виде есть задача на максимумы и минимумы! Следует, разумеется, различать решение поставленной задачи для отдельного вещественного алгебраического уравнения и для векового уравнения $n \times n$ -матрицы B , в частности, вещественной симметричной S и комплексной эрмитовой H .

Во-первых, спектр собственных значений для последних априори всегда вещественный. *Во-вторых*, для положительной определённости именно матриц S или H , согласно признаку Сильвестра, необходимо и достаточно, чтобы детерминанты их главных угловых миноров были положительные, или, что тождественно, с учётом признака Декарта, — чтобы все коэффициенты их векового уравнения были положительные. (Или, чтобы n характеристических следов S и H были положительные!) *В-третьих*, несингулярные вещественные симметричные матрицы типа $S = A' \cdot A$, или $A \cdot A'$ или типа S^2 и комплексные эрмитовы матрицы типа $H = A^* \cdot A$, или $A \cdot A^*$ или типа H^2 априори имеют всегда положительные собственные значения. Более того, как было показано выше, элементы нормальных матриц N содержат достаточно изначальной информации для того, чтобы решать задачу об отыскании их собственных значений, сводя её к решению алгебраического (векового) уравнения с заведомо вещественными положительными корнями.

Решение поставленной задачи для векового уравнения $n \times n$ -матрицы общего вида и для отдельного алгебраического уравнения степени n при $n > 4$ определяется ответом на вопрос: «Имеет ли вещественное алгебраическое уравнение комплексные сопряжённые корни или нет?» Ранее уже было сказано, что точный ответ на данный вопрос для *конкретного* уравнения всегда можно получить однозначно, используя классический метод Штурма. Однако этот метод, даже теоретически, не даёт общих необходимых и достаточных требований, которым должны отвечать коэффициенты вещественного алгебраического уравнения для вещественности всех его корней или, с учётом процедуры смещения аргумента в положительном направлении, — для их положительности.

Первоочередное необходимое требование для вещественности и положительности корней, согласно классическому **признаку Декарта**, состоит в положительности коэффициентов алгебраического уравнения в знакопередающей форме (257) или (258). Однако данный признак не гарантирует того, что у него нет пар комплексных сопряжённых корней. Например, при выборе даже самого большого априорного параметра положительного смещения $\Delta^{(-)} = -(1 + \max |k_j|)$ гарантируется только то, что вещественные части комплексных корней будут положительные.

Согласно цепи (224) генерального неравенства для средних величин, алгебраическое уравнение с вещественными положительными корнями имеет равные медианы (образуемые из коэффициентов) тогда и только тогда, когда уравнение имеет *биномиальную* форму:

$$y(x) = (x - \mu)^n = \sum_{j=0}^n (-1)^j \left(C_n^j \cdot \mu^j \right) \cdot x^{n-j} = 0 \leftrightarrow \bar{m}_j = \mu = \text{const.} \quad (293)$$

При отличии хотя бы двух корней друг от друга коэффициенты уравнения никак не могут соответствовать биномиальной форме (293), причём действует неравенство (224). Например, совпадение каких-либо отдельных или соседних медиан (213), (214), нарушение их иерархии с изменением порядка t от 1 до n , обнуление отдельных коэффициентов k_t до $t = n$ — всё это те отклонения, которые свидетельствуют о том, что алгебраическое уравнение с неотрицательными коэффициентами имеет комплексные сопряжённые корни! Наиболее общо любое вещественное алгебраическое уравнение с положительными корнями неотъемлемо характеризуется цепью иерархических инвариантов из всех медиан.

Отсюда более строгое *необходимое* требование к коэффициентам вещественного алгебраического уравнения в знакочередующейся форме нежели классический признак Декарта есть **иерархический признак**, заключающийся в необходимости выполнения цепи (224) генерального неравенства или любого другого отрезка из n медиан (разумеется, при несовпадении всех корней) [27]. Например, в качестве медиан берутся средние алгебраические, выражаемые через коэффициенты уравнения k_t по формулам (213), (214); или первые n средних степенных, выражаемые через суммы Варинга или характеристические следы по (215), (216). Но и этот признак недостаточен. Его недостаточность видна из примера:

$$x^3 - 29x^2 + 278x - 868 = (x - 7) \cdot \left[x - (11 + \sqrt{3} \cdot i) \right] \cdot \left[x - (11 - \sqrt{3} \cdot i) \right] = 0,$$

где $\bar{m}_1 = 29/3$, $\bar{m}_2 = \sqrt{278/3}$, $\bar{m}_3 = \sqrt[3]{868}$; причём $\bar{m}_1 > \bar{m}_2 > \bar{m}_3 > 0$. Причиной недостаточности признака является то, что генеральное неравенство средних величин устанавливает лишь иерархию медиан, но не возможные интервалы для них.

Пусть x_1 и x_2 — решения вещественного квадратного уравнения $x^2 - k_1 \cdot x + k_2 = 0$ при переменных-коэффициентах $k_1 > 0$ и $k_2 > 0$. Ранее в примере 2 из § 1.9 и в примере из § 3.3 было показано, что скалярные функции u_1 и u_2 для разности и для отношения средних арифметических и геометрических от пары x_1 и x_2 имеют нестрогую глобальную стационарность на биссектрисе 1-го квадранта (при $x_1 = x_2$), причём *минимумы* 0 и 1 для вещественных x_1 и x_2 и *максимумы* 0 и 1 для комплексных сопряжённых x_1 и x_2 . Однако при $n > 2$ в случае смешивания положительных вещественных корней и пар комплексных сопряжённых корней в аналогичных скалярных функциях u_1 и u_2 от них вещественные и комплексные сопряжённые корни уравнения степени n конкурируют друг с другом в определении характера стационарности на биссектрисе 1-го квадранта (т. е. при $x_1 = x_2 = \dots = x_n$), что приводит к *перегибу* функций u_1 и u_2 . Отсюда именно для этого случая генеральное неравенство средних может и выполняться, и не выполняться!

Однозначные условные экстремумы слева и справа для k_t или для $\bar{m}_t$ при положительных корнях уравнения здесь обязательно существуют, в силу имеющихся ограничений сверху и снизу в цепи неравенств (224). Если каждый последующий коэффициент k_t входит в вычисляемое двойное неравенство, то все коэффициенты алгебраического уравнения в данном смысле корректные. Теоретически аналогичные экстремумы вычисляются исходя из сумм Варинга (характеристических следов).

Укажем пример, весьма близкий к предыдущему, но в котором все коэффициенты уравнения в данном смысле корректные:

$$x^3 - 29x^2 + 275x - 847 = (x - 7) \cdot (x - 11)^2 = 0;$$

где $\bar{m}_1 = 29/3$, $\bar{m}_2 = \sqrt{275/3}$, $\bar{m}_3 = \sqrt[3]{847}$ и также $\bar{m}_1 > \bar{m}_2 > \bar{m}_3 > 0$.

Далее преобразуем исходное алгебраическое уравнение степени n (257) с предполагаемыми вещественными положительными корнями x_i к весьма важной для последующего анализа *приведённой* форме:

$$\begin{aligned} y_n(x) = & x^n - C_n^1 \cdot \bar{m}_1 \cdot x^{n-1} + C_n^2 \cdot \bar{m}_2^2 \cdot x^{n-2} - \dots + (-1)^j \cdot C_n^j \cdot \bar{m}_j^j \cdot x^{n-j} + \dots \\ & \dots + (-1)^{n-1} \cdot C_n^{n-1} \cdot \bar{m}_{n-1}^{n-1} \cdot x + (-1)^n \cdot \bar{m}_n^n = 0, \end{aligned} \quad (298)$$

где $\bar{m}_j^j$ — *приведённые* коэффициенты, $\bar{m}_j$ — средние алгебраические, или малые медианы от положительных корней x_i , согласно (213) и (214); $\bar{m}_1 > \bar{m}_2 > \dots > \bar{m}_{n-1} > \bar{m}_n$ при неравенстве тут хотя бы одного корня другому, согласно цепи неравенств (224). Именно в приведённой форме алгебраического уравнения непосредственно работает необходимый *иерархический признак* положительности всех его корней!

* * *

Рассмотрим отдельно алгебраическое уравнение той же степени n , но с неотрицательными корнями $x_i \geq 0$. Поясним, зачем всё это нужно. Дело в том, что иногда находимая нижняя допустимая граница для $\bar{m}_t$ в (295) нулевая, что отвечает одному или нескольким нулевым корням уравнения. Например, это имеет место всегда в неравенстве (294), т. е. при $t = 2$; тут нижняя нулевая граница отвечает $(n - 1)$ *нулевому* корню. При $t > 2$ знание нижней границы, обусловленной нулевыми корнями, может быть полезным для более полной информации о вычисляемой истинной нижней границе (как она может достигаться).

Пусть уравнение (298) имеет $s = n - r$ нулевых корней. Числа s и r , по определению, суть степень вырожденности и ранг алгебраического уравнения. Иерархическая цепь неравенств для $\bar{m}_j$ обрывается здесь на $j = r$ (так как при $j > r$ все медианы обнуляются). Из уравнения (298) выделим фрагмент, отвечающий только положительным корням:

$$y_r(x) = x^r - C_r^1 \cdot \bar{m}_1 \cdot x^{r-1} + C_r^2 \cdot \bar{m}_2^2 \cdot x^{r-2} - \dots + (-1)^r \cdot C_r^j \cdot \bar{m}_j^j \cdot x^{r-j} + \dots$$

$$\dots + (-1)^{r-1} \cdot C_r^{r-1} \cdot \bar{m}_{r-1}^{r-1} \cdot x + (-1)^r \cdot \bar{m}_r^r = 0, \quad (299)$$

коэффициенты и медианы в этих уравнениях связаны формулами

$$k_j(n) = k_j(r) = C_n^j \cdot \bar{m}_j^j(n) = C_r^j \cdot \bar{m}_j^j(r), \quad (300)$$

где $j = \overline{1, r}$; $\bar{m}_j(n)$ — j -е медианы в (298), $\bar{m}_j(r)$ — j -е медианы в (299).

В частности, в уравнении (298) степени вырожденности $s = n - r$ медианы подчиняются более строгой цепи *критических* неравенств:

$$\frac{C_n^1}{C_r^1} \cdot \bar{m}_1(n) \geq \sqrt{\frac{C_n^2}{C_r^2}} \cdot \bar{m}_2(n) \geq \dots \geq r - j \sqrt{\frac{C_n^{r-1}}{C_r^{r-1}}} \cdot \bar{m}_{r-1}(n) \geq \sqrt[r]{C_n^r} \cdot \bar{m}_r(n), \quad (301)$$

тождественной цепи $\bar{m}_1(r) \geq \bar{m}_2(r) \geq \dots \geq \bar{m}_{r-1}(r) \geq \bar{m}_r(r)$ в (299). Знаки равенства, причём все сразу, имеют место тогда и только тогда, когда все ненулевые корни x_i равны. При этом само уравнение степени n имеет *квазибиномиальную* форму. Например, если фрагмент (299) уравнения имеет биномиальную форму типа (293), то малые и большие медианы от корней уравнения (298) степени n и ранга r изменяются с ростом порядка j как функции:

$$\bar{m}_j(n) = \mu \cdot \sqrt[j]{C_r^j / C_n^j}, \quad \bar{M}_j(n) = \mu \cdot \sqrt[j]{r / n}. \quad (302, 303)$$

Обратим внимание на то, что внутри (301) выполняется неравенство

$$\left[\sqrt[i]{C_r^i / C_n^i} / \sqrt[j]{C_r^j / C_n^j} \right] < 1, \quad (1 < i < j < r < n), \quad (304)$$

которое и обуславливает его большую строгость в сравнении с менее строгой исходной цепью

$$\bar{m}_1(n) > \bar{m}_2(n) > \dots > \bar{m}_{r-1}(n) > \bar{m}_r(n) > 0.$$

Критические неравенства (301) и редукционное неравенство (304) принимают разнообразные частные случаи. Рассмотрим особо важный частный случай, отвечающий параметрам: $n = t$, $r = t - 1$. Он имеет значение, например, в полном анализе вещественного алгебраического уравнения в знакопередающей форме на предмет вещественности и положительности всех его корней. Рассматриваемый случай отвечает алгебраическому уравнению степени t с одним нулевым корнем:

$$y_t(x) = x^t - C_t^1 \cdot \bar{m}_1(t) \cdot x^{t-1} + C_t^2 \cdot \bar{m}_2^2(t) \cdot x^{t-2} - \dots + (-1)^t \cdot C_t^j \cdot \bar{m}_j^j(t) \cdot x^{t-j} + \dots \\ \dots + (-1)^{t-2} \cdot C_t^{t-2} \cdot \bar{m}_{t-2}^{t-2}(t) \cdot x + (-1)^{t-1} \cdot C_t^{t-1} \cdot \bar{m}_{t-1}^{t-1}(t) \cdot x = 0, \quad (305)$$

(в нём свободный член обязательно нулевой $\bar{m}_t^t(t) = x_1 \cdot x_2 \cdot \dots \cdot x_t = 0$).

Формулы (300), (301) приобретают вид:

$$k_j(t) = k_j(t-1) = C_t^j \cdot \bar{m}_j^j(t) = C_{t-1}^j \cdot \bar{m}_j^j(t-1), \quad (306)$$

$$\frac{C_t^1}{C_{t-1}^1} \cdot \bar{m}_1(t) \geq \sqrt{\frac{C_t^2}{C_{t-1}^2}} \cdot \bar{m}_2(t) \geq \dots \geq t-2 \sqrt{\frac{C_t^{t-2}}{C_{t-1}^{t-2}}} \cdot \bar{m}_{t-2}(t) \geq t-1 \sqrt{C_t^{t-1}} \cdot \bar{m}_{t-1}(t) \geq 0, \quad (307)$$

что тождественно $\bar{m}_1(t-1) \geq \bar{m}_2(t-1) \geq \dots \geq \bar{m}_{t-2}(t-1) \geq \bar{m}_{t-1}(t-1)$.

Если уравнение (305) имеет квазибиномиальную форму, то как тот же частный случай формул (302) и (303) имеем:

$$\bar{m}_j(t) = \mu \cdot \sqrt[j]{C_{t-1}^j / C_t^j} = \mu \cdot \sqrt[j]{(t-j)/t}, \quad \bar{M}_j(t) = \mu \cdot \sqrt[j]{(t-1)/t}. \quad (308, 309)$$

Аналогичные связи имеются между алгебраическими медианами для систем положительных корней $\langle x_1, x_2, \dots, x_{t-1} \rangle$ и неотрицательных корней $\langle x_1, x_2, \dots, x_t \rangle$ с одним нулевым корнем. Они используются именно тогда, когда в результате предварительной проверки устанавливается, что в системе (297) все коэффициенты k_j порядка $1 \leq j \leq t-1$ корректные, а далее нужно проверить корректность коэффициента порядка t . Если в (305) для $\bar{m}_j(t)$ выполняется цепь неравенств (307), то $\min\{m_t\} = 0$ (при этом, например, $x_t = 0$). В противном случае в (305) $\min\{m_t\} > 0$ ($x_t > 0$). Это позволяет в первом случае вычислять только $\max\{m_t\}$. Тогда корректная медиана m_t должна находиться обязательно в интервале от 0 до $\max\{m_t\}$.

* * *

Отметим ещё одну (помимо *иерархической*) важную особенность приведённой формы алгебраического уравнения типа (298). А именно: последовательное дифференцирование многочлена этого уравнения с точностью до постоянного множителя сохраняет для производных ту же самую приведённую форму, но только с понижением её степени! Для многочлена $y_n(x)$ в приведённой форме имеем производные по x :

$$\left. \begin{aligned} \frac{dy_n}{dx} &= (-1)^{n-1} \cdot n \cdot y_{n-1}(x), \\ \frac{d^2 y_n}{dx^2} &= (-1)^{n-2} \cdot n \cdot (n-1) \cdot y_{n-2}(x), \\ &\dots\dots\dots \\ \frac{d^{n-1} y_n}{dx^{n-1}} &= -n! \cdot y_1(x), \\ \frac{d^n y_n}{dx^n} &= +n! = \text{const.} \end{aligned} \right\} \quad (310)$$

В процессе дифференцирования многочлена $y_n(x)$ в приведённой форме последовательно отсекаются именно коэффициенты высшего порядка (что будет далее использоваться). Аналогичная редукционная картина имеет место и при последовательном дифференцировании многочлена приведённой формы *инверсионного* алгебраического уравнения (288):

$$\begin{aligned} y_n(x^{-1}) &= x^{-n} - C_n^1 \cdot \bar{v}_1^{-1} \cdot x^{-(n-1)} + C_n^2 \cdot \bar{v}_2^{-2} \cdot x^{-(n-2)} - \dots \\ &\dots + (-1)^n \cdot C_n^j \cdot \bar{v}_j^{-j} \cdot x^{-(n-j)} + \dots + (-1)^{n-1} \cdot C_n^{n-1} \cdot \bar{v}_{n-1}^{-(n-1)} \cdot x^{-1} + (-1)^n \cdot \bar{v}_n^{-n} = 0. \end{aligned} \quad (311)$$

Здесь $\bar{v}_j^{-1}$ — алгебраические медианы из цепи (285) инверсионного генерального неравенства средних; $\bar{v}_j^{-j}$ — приведённые инверсионные коэффициенты. Медианы выражаются как средние алгебраические от обратных корней по обращённым формулам (219), (220). Причём и тут соблюдение иерархии медиан, как в цепи (285), — исходное требование к коэффициентам уравнения (281) с вещественными положительными корнями. Последовательное дифференцирование многочлена уравнения $y_n(x^{-1}) = y_n(u)$ по $u = x^{-1}$ точно также сохраняет его приведённую форму с точностью до постоянного множителя. При этом уравнение (311) далее анализируется именно в прямой форме, но с новой переменной u .

Однако, несмотря на эту, казалось бы, простую взаимосвязь прямой и инверсионной форм алгебраического уравнения сопутствующие им генеральные неравенства для медиан локально различны. Причём их отдельные элементы взаимосвязаны между собой формулами (283). Например, из *прямого* окаймляющего неравенства для медианы $\bar{v}_{n-t}^{-1}$ уравнения (311), с учётом формул (283), получаем дополнительное неравенство для медиан уравнения (268):

$$\begin{aligned} \bar{v}_{n-t-1}^{-1} > \bar{v}_{n-t}^{-1} > \bar{v}_{n-t+1}^{-1} &\leftrightarrow \bar{v}_{n-t-1} < \bar{v}_{n-t} < \bar{v}_{n-t+1} \leftrightarrow \\ \leftrightarrow \bar{v}_{n-t-1}^{n-t} < \bar{v}_{n-t}^{n-t} < \bar{v}_{n-t+1}^{n-t} &\leftrightarrow \\ \leftrightarrow \bar{v}_{n-t-1}^{n-t-1} \cdot \bar{v}_{n-t-1} < \bar{v}_{n-t}^{n-t} < \bar{v}_{n-t+1}^{n-t+1} / \bar{v}_{n-t+1} &\leftrightarrow \\ \leftrightarrow \bar{m}_n^n \cdot \bar{v}_{n-t-1} / \bar{m}_{t+1}^{t+1} < \bar{m}_n^n / \bar{m}_t^t < \bar{m}_n^n / \bar{m}_{t-1}^{t-1} \cdot \bar{v}_{n-t+1} &\leftrightarrow \\ \leftrightarrow \bar{m}_{t+1}^{t+1} / \bar{v}_{n-t-1} > \bar{m}_t^t > \bar{m}_{t-1}^{t-1} \cdot \bar{v}_{n-t+1} . \end{aligned}$$

Отсюда для $\bar{m}_t$ следует *обратное* окаймляющее неравенство

$$\bar{m}_{t+1} \cdot \sqrt[t]{\bar{m}_{t+1} / \bar{v}_{n-(t+1)}} > \bar{m}_t > \bar{m}_{t-1} \cdot \sqrt[t]{\bar{v}_{n-(t-1)} / \bar{m}_{t-1}} , \quad (312)$$

которое как бы дополняет прямое *иерархическое* неравенство из (224) ... $< \bar{m}_{t+1} < \bar{m}_t < \bar{m}_{t-1} \dots$, т. е. фрагмент генерального неравенства. Из цепей (224) и (225) генерального неравенства средних видно, что в (312) левое подкоренное выражение больше 1, а правое подкоренное выражение меньше 1, как и должно быть здесь по смыслу. (Обратим внимание на то, что знаки неравенств направлены противоположно.)

* * *

Для установления искомого *полного признака вещественности и положительности корней алгебраического уравнения* остаётся найти последовательный способ вычисления парных экстремумов в (295) и (296). Например, для уравнения 2-й степени полный признак есть (294), или, что эквивалентно, $\bar{m}_1, \bar{m}_2 > 0, \Delta^2 \geq 0$, где Δ^2 — дискриминант. Вначале более подробно в этом аспекте изучим уравнение 3-й степени. Далее совершенно естественным путём придём к общему признаку для уравнения степени n , решив поставленную экстремальную задачу (297) с применением дифференциального подхода.

В приведённой форме уравнение 3-й степени имеет вид:

$$y_3(x) = x^3 - 3\bar{m}_1 \cdot x^2 + 3\bar{m}_2 \cdot x - \bar{m}_3 = 0. \quad (313)$$

Дальнейший анализ этого уравнения на предмет вещественности и положительности всех его 3-х корней имеет смысл только, если на 1-м этапе проверки выполняется признак Декарта (т. е. положительность всех трёх коэффициентов k_j), а на 2-м этапе проверки выполняется иерархический признак $\bar{m}_1 > \bar{m}_2 > \bar{m}_3 > 0$ (при неравенстве хотя бы одного корня другому) или $\bar{m}_1 = \bar{m}_2 = \bar{m}_3 > 0$ (при равенстве всех корней $x_1 = x_2 = \dots = x_3 > 0$). Для уравнения 3-й степени в форме (313) его дискриминант выражается так:

$$\Delta^2 = (x_1 - x_2)^2 \cdot (x_1 - x_3)^2 \cdot (x_2 - x_3)^2 = \\ = \left(3\bar{m}_1^2 \cdot \bar{m}_2^4 + 6\bar{m}_1 \cdot \bar{m}_2^2 \cdot \bar{m}_3^3 - \bar{m}_3^6 - 4\bar{m}_1^3 \cdot \bar{m}_3^3 \right) / 4;$$

$$\text{где } \bar{m}_1 = k_1/3, \quad \bar{m}_2^2 = k_2/3, \quad \bar{m}_3^3 = k_3.$$

Уравнение 3-й степени с вещественными коэффициентами имеет все *вещественные* корни тогда и только тогда, когда $\Delta^2 \geq 0$. С учётом ещё и признака Декарта: для *вещественности и положительности* всех корней необходимо и достаточно, чтобы выполнялись требования: $\bar{m}_1, \bar{m}_2, \bar{m}_3 > 0, \Delta^2 \geq 0$. Это пока вполне аналогично требованию для уравнения 2-й степени. Но зато при $n > 3$ полный признак на основе дискриминанта алгебраического уравнения не применим!

Для полноты изложения вопроса проиллюстрируем сначала на том же уравнении 3-й степени возможности, которые предоставляет в этом аспекте метод Штурма (§ 4.2). Напомним, что для его применения анализируемое алгебраическое уравнение должно быть простым.

Поскольку сам метод Штурма, по сути, базируется на алгоритме Евклида применительно к процессу вычисления наибольшего общего делителя-многочлена для $y(x)$ и $y'(x)$, то для простого уравнения 3-й степени процесс (275) должен включать два шага:

$$\left. \begin{aligned} y(x) &= l_1(x) \cdot y'(x) - f_1(x), \\ y'(x) &= l_2(x) \cdot f_1(x) - f_2, \\ f_2 &= \text{const.} \end{aligned} \right\} \quad l_i(x) = a_i \cdot x + b_i,$$

Конкретно для уравнения (313) дополнительно имеем:

$$y'(x) / 3 = x^2 - 2\bar{m}_1 \cdot x + \bar{m}_2^2,$$

$$f_1(x) = 2 \cdot (\bar{m}_1^2 - \bar{m}_2^2) \cdot x - (\bar{m}_1 \cdot \bar{m}_2^2 - \bar{m}_3^3),$$

$$f_2 = \frac{1}{4} \cdot \left(\frac{\bar{m}_1 \cdot \bar{m}_2^2 - \bar{m}_3^3}{\bar{m}_1^2 - \bar{m}_2^2} \right)^2 - \bar{m}_1 \cdot \frac{\bar{m}_1 \cdot \bar{m}_2^2 - \bar{m}_3^3}{\bar{m}_1^2 - \bar{m}_2^2} + \bar{m}_2^2.$$

Далее составляем табл. 3 для подсчёта общего числа перемен знаков в априорных интервалах аргумента в ряду $y(x)$, $-y'(x)$, $f_1(x)$, f_2 .

Таблица 3. Перемены знака многочленов в методе Штурма при $n = 3$.

$f(x) \backslash x$	$x = +\infty$	$x = 0$	$x = -\infty$
$y(x)$	$+\infty$	$-$	$-\infty$
$-y'(x)$	$-\infty$	$-$	$-\infty$
$f_1(x)$	$+\infty$	$-$	$-\infty$
f_2	f_2	f_2	f_2

Согласно теореме Штурма, отсюда следует: для того, чтобы все корни простого уравнения (313) с вещественными положительными коэффициентами k_j были вещественными и при этом положительными, необходимо и достаточно, чтобы выполнялось знаковое условие $f_2 < 0$. Тогда в *априорном* для них интервале $(0, +\infty)$ по числу перемен знака в 1-й колонке имеем три различных вещественных корня: $x_3 > x_2 > x_1 > 0$.

Введём следующие обозначения:

$$p = (\bar{m}_1 \cdot \bar{m}_2^2 - \bar{m}_3^3) / (\bar{m}_1^2 - \bar{m}_2^2),$$

$$\alpha^3 = (3\bar{m}_1 \cdot \bar{m}_2^2 - 2\bar{m}_1^3) + 2 \cdot (\bar{m}_1^2 - \bar{m}_2^2)^{3/2},$$

$$\beta^3 = (3\bar{m}_1 \cdot \bar{m}_2^2 - 2\bar{m}_1^3) - 2 \cdot (\bar{m}_1^2 - \bar{m}_2^2)^{3/2}.$$

Из требования $f_2 < 0$ следуют эквивалентные неравенства:

$$\begin{aligned}
 4f_0 = p^2 - 4\bar{m}_1 \cdot p + 4\bar{m}_2^2 &= (p - p_1) \cdot (p - p_2) < 0 \leftrightarrow p_1 < p < p_2 \leftrightarrow \\
 \leftrightarrow 2 \cdot \left(\bar{m}_1 - \sqrt{\bar{m}_1^2 - \bar{m}_2^2} \right) &< \frac{\bar{m}_1 \cdot \bar{m}_2^2 - \bar{m}_3^3}{\bar{m}_1^2 - \bar{m}_2^2} < 2 \cdot \left(\bar{m}_1 + \sqrt{\bar{m}_1^2 - \bar{m}_2^2} \right) \leftrightarrow \\
 \leftrightarrow \beta^3 < \bar{m}_3^3 < \alpha^3 &\leftrightarrow \beta < \bar{m}_3 < \alpha.
 \end{aligned} \tag{314}$$

В силу того, что исходно $\bar{m}_1 > \bar{m}_2 > 0$, имеем $\alpha > 0$. Но знак числа β тут зависит от соотношения алгебраических медиан $\bar{m}_1$ и $\bar{m}_2$:

1) при $3\bar{m}_1^2 = 4\bar{m}_2^2$ имеем $\beta = 0$,

2) при $3\bar{m}_1^2 > 4\bar{m}_2^2$ имеем $\beta < 0$,

3) при $3\bar{m}_1^2 < 4\bar{m}_2^2$ имеем $\beta > 0$.

Эти три варианта легко интерпретируются с привлечением неравенства (307) при значениях параметров $t = 3$ и $r = 2$. А именно, первые два варианта отвечают в (295) $\min\{\bar{m}_3\} = 0$ при $\beta \leq 0$; третий же вариант отвечает в (295) $\min\{\bar{m}_3\} = \beta$ при $\beta > 0$.

Но гораздо бо́льшая конкретика достигается в результате решения этой задачи на условные экстремумы (максимум и минимум) высшего коэффициента $\bar{m}_3$ в её постановке (297), например, методом Лагранжа. (Здесь уравнение может иметь и кратные корни!) Соответственно тогда при значении $n = 3$ имеем ту же самую пару экстремальных решений, отвечающих максимуму и минимуму высшего коэффициента $k_3 = \bar{m}_3^3$.

$$1) \ x_1 = x_2 = \frac{k_1 - \sqrt{k_1^2 - 3k_2}}{3}, \quad x_3 = \frac{k_1 + 2\sqrt{k_1^2 - 3k_2}}{3}; \tag{315}$$

$$a = (x_1)^2 \cdot x_3 = \max\{k_3\} = \frac{1}{27} \cdot \left[(9k_1k_2 - 2k_1^3) + 2 \cdot (k_1^2 - 3k_2)^{3/2} \right]. \tag{316}$$

Отсюда при $k_1 = 3\bar{m}_1$ и $k_2 = 3\bar{m}_2^2$, т. е. для уравнения (313), имеем:

$$x_1 = x_2 = \bar{m}_1 - \sqrt{\bar{m}_1^2 - \bar{m}_2^2}, \quad x_3 = \bar{m}_1 + 2\sqrt{\bar{m}_1^2 - \bar{m}_2^2}; \tag{317}$$

$$a = \alpha^3 = \max\{\bar{m}_3^3\} = (3\bar{m}_1 \cdot \bar{m}_2^2 - 2\bar{m}_1^3) + 2 \cdot (\bar{m}_1^2 - \bar{m}_2^2)^{3/2}. \tag{318}$$

$$2) \ x_1 = x_2 = \frac{k_1 + \sqrt{k_1^2 - 3k_2}}{3}, \ x_3 = \frac{k_1 - 2\sqrt{k_1^2 - 3k_2}}{3}; \quad (319)$$

$$b = (x_1)^2 \cdot x_3 = \frac{1}{27} \cdot \left[(9k_1k_2 - 2k_1^3) - 2 \cdot (k_1^2 - 3k_2)^{3/2} \right]. \quad (320)$$

Отсюда при $k_1 = 3\bar{m}_1$ и $k_2 = 3\bar{m}_2^2$, т. е. для уравнения (313), имеем:

$$x_1 = x_2 = \bar{m}_1 + \sqrt{\bar{m}_1^2 - \bar{m}_2^2}, \ x_3 = \bar{m}_1 - 2\sqrt{\bar{m}_1^2 - \bar{m}_2^2}; \quad (321)$$

$$b = \beta^3 = \left(3\bar{m}_1 \cdot \bar{m}_2^2 - 2\bar{m}_1^3 \right) - 2 \cdot \left(\bar{m}_1^2 - \bar{m}_2^2 \right)^{3/2}. \quad (322)$$

В итоге для β имеем те же вышеуказанные три варианта, но теперь с конкретной интерпретацией, с учётом значений корней x_1 , x_2 и x_3 :

$$\max\langle 0, \beta^3 \rangle \leq \bar{m}_3^3 \leq \alpha^3 \leftrightarrow \max\langle 0, \beta \rangle \leq \bar{m}_3 \leq \alpha. \quad (323, 324)$$

В данном случае при $x_3 = 0$ ($\bar{m}_3 = 0$) имеем $3\bar{m}_1^2 = 4\bar{m}_2^2$ и обратно, где $\beta = 0$; при $x_3 > 0$ ($\bar{m}_3 > 0$) имеем $3\bar{m}_1^2 > 4\bar{m}_2^2$ и обратно, где $\beta < 0$; при $x_3 < 0$ (что не допустимо!) имеем $3\bar{m}_1^2 < 4\bar{m}_2^2$ и обратно, где $\beta > 0$. Варианты $\beta \leq 0$ отвечают в (324) $\min\{\bar{m}_3\} = 0$. Вариант $\beta > 0$ отвечает в (324) $\min\{\bar{m}_3\} = \beta$. Это объясняется просто тем, что при $\bar{m}_3 = 0$, согласно (307), тут должно обязательно выполняться более строгое *критическое* неравенство $3\bar{m}_1^2 \geq 4\bar{m}_2^2$. Однако, если последнее не выполняется, то тогда обязательно $\bar{m}_3 \geq \beta > 0$. Очевидно, что знаки равенств в (323), (324) отвечают парам равных корней $x_1 = x_2$ в (317), (319), но при этом неравному им x_3 .

Итак, выше были рассмотрены три способа проверки корректности всех коэффициентов для уравнения 3-й степени (313): с применением дискриминанта, с применением метода Штурма и с применением метода Лагранжа. Все они приводят к одному и тому же результату. Первый способ пригоден только при $n \leq 3$. Второй и третий способы с увеличением n более 3-х неизбежно требуют резко возрастающих объёмов вычислений, а также проработки множества возникающих при этом вариантов.

* * *

Поэтому далее установим общий *аналитико-геометрический метод* для последовательной проверки корректности медиан алгебраического уравнения степени n в приведённой форме (298). Он позволяет дать *теоретическую картину* корректности коэффициентов алгебраического уравнения при любом значении степени n , т. е., наконец, сформулировать необходимые и достаточные требования к ним для вещественности и положительности всех корней уравнения.

Начнём также с уравнения (313). Выразим из него коэффициент $\bar{m}_3^3$ как функцию от x :

$$y_3(x) = 0 \leftrightarrow \bar{m}_3^3(x) = x^3 - 3\bar{m}_1 \cdot x^2 + 3\bar{m}_2^2 \cdot x, \quad (325)$$

где $\bar{m}_1 > \bar{m}_2 > 0$. (При $\bar{m}_1 = \bar{m}_2 = \bar{m}_3 > 0$ уравнение (313) имеет биномиальную форму и $x_1 = x_2 = x_3 > 0$.) Все возникающие варианты графиков функции $\bar{m}_3^3(x)$ при корректных значениях $\bar{m}_1 > \bar{m}_2 > 0$ отображены в виде *геометрической интерпретации* на рис. 7. Находим аналитически точки максимума x_1^+ и минимума x_2^- этой функции через *сопутствующее уравнение* порядка $n - 1 = 2$:

$$\frac{1}{3} \cdot \frac{d\bar{m}_3^3}{dx^3} = y_2(x) = x^2 - 2\bar{m}_1 \cdot x + \bar{m}_2^2 = 0. \quad (326)$$

Отсюда имеем:

$$x_1^+ = \bar{m}_1 - \sqrt{\bar{m}_1^2 - \bar{m}_2^2}, \quad \frac{d^2\bar{m}_3^3}{dx^2}(x_1^+) = -6 \cdot \sqrt{\bar{m}_1^2 - \bar{m}_2^2} < 0;$$

$$x_2^- = \bar{m}_1 + \sqrt{\bar{m}_1^2 - \bar{m}_2^2}, \quad \frac{d^2\bar{m}_3^3}{dx^2}(x_2^-) = +6 \cdot \sqrt{\bar{m}_1^2 - \bar{m}_2^2} > 0.$$

Начальные значения: $\bar{m}_3^3(0) = 0, \quad \frac{d\bar{m}_3^3}{dx^3}(0) = 3\bar{m}_2^2 > 0.$

Подставляя эти значения аргумента x в функцию (325), получаем те же границы интервала $\bar{m}_3^3(x_1^+) = \alpha^3$ и $\bar{m}_3^3(x_2^-) = \beta^3$, что были получены ранее по формулам (318) и (322). Но теперь на рис. 7 можно весьма наглядно и чисто геометрически интерпретировать допустимые интервалы для корректных значений высшего коэффициента $\bar{m}_3^3$ по неравенству (294), выведенные именно дифференциальным способом.

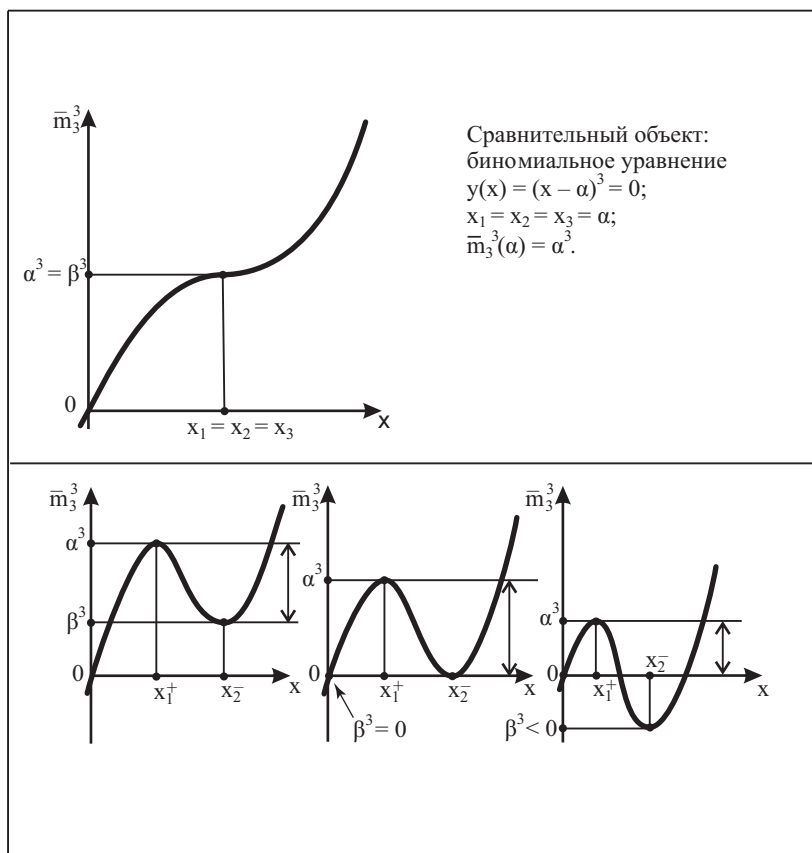


Рис. 7. Варианты допустимых интервалов для корректных значений высшего коэффициента $\bar{m}_3^3$ алгебраического уравнения 3-й степени при корректных значениях $\bar{m}_1$ и $\bar{m}_2$.

Изменение высшего коэффициента $\bar{m}_3^3$ в исходном уравнении (313) геометрически эквивалентно тут смещению оси x по вертикали либо вверх, либо вниз. При этом точки пересечения этой оси с кривой $\bar{m}_3^3(x)$ являются тремя вещественными положительными корнями уравнения (313) с изменяемым коэффициентом $\bar{m}_3^3$ до тех пор, пока смещение по ординате находится в границах неравенства (323).

Причём экстремальные верхнее или нижнее ненулевые допустимые значения $\bar{m}_3^3(x)$ отвечают появлению *двукратных* корней уравнения, т. е. либо $x_1 = x_2 = x_1^+$, согласно (317), либо $x_1 = x_2 = x_2^-$, согласно (321). Минимальное нулевое значение отвечает двукратному нулевому корню. Эти геометрически наглядные маргинальные свойства имеют общий смысл при выборе и интерпретации допустимых границ интервала для высшего коэффициента алгебраического уравнения любой степени n .

Кроме того, обратим внимание ещё на три обстоятельства, которые тоже имеют место при любом порядке высшего коэффициента $\bar{m}_t^t = k_t$. *Во-первых*, формы кривых $u_3(x)$ в (313) и $\bar{m}_3^3(x)$ в (325) тождественные. *Во-вторых*, сопутствующее уравнение (326) имеет ту же приведённую форму и с теми же коэффициентами, что и исходное уравнение (313), но степень его понижена на 1. *В-третьих*, точки экстремумов функции $\bar{m}_3^3(x)$ суть решения сопутствующего уравнения пониженной степени.

Если границы α^3 и β^3 (рис. 7) сближаются, то интервал для $\bar{m}_3^3$ сужается. Вариант $\alpha^3 = \beta^3$ отвечает биномиальной форме исходного уравнения (313). В этом случае допускается единственное корректное значение $\bar{m}_3^3 = \alpha^3 = \beta^3$ ($\bar{m}_1 = \bar{m}_2 = \bar{m}_3 = \alpha = \beta$). Два дополнительных одинаковых корня получаются из-за совпадения $\bar{m}_3^3$ и с α^3 , и с β^3 . Любые смещения оси x от уровня $\alpha^3 = \beta^3$ вверх или вниз приводят к появлению пары комплексных сопряжённых корней. Так, уравнение $u_3(x) = (x - a)^3 \pm c = 0$ при $c \neq 0$ обязательно имеет пару комплексных сопряжённых корней и один вещественный корень.

Уже при $n = 3$ видно, что 1-е ограничительное неравенство $\bar{m}_2 < \bar{m}_1$ есть необходимое требование для выполнения последующего 2-го ограничительного неравенства (324). Причём само (324) более сильное требование к $\bar{m}_3$, нежели иерархическое неравенство $0 < \bar{m}_3 < \bar{m}_2$.

Обратим внимание и на то, что (α^3, β^3) есть интервал для $\bar{m}_3^3$, при котором корни уравнения (283) суть *вещественные и положительные*, т. е. при $0 < \bar{m}_2 < \bar{m}_1$. Если же последнее требование снять и перейти к корректным значениям коэффициентов $\bar{m}_1$ и $\bar{m}_2^2$ при ограничении $\bar{m}_2^2 < \bar{m}_1^2$, то отсюда имеем интервал для $\bar{m}_3^3$, при котором корни уравнения 3-й степени в знакопередающей форме просто все *вещественные*!

Перейдём к общему алгебраическому уравнению степени $n \geq 3$, выраженному в приведённой форме (298). Согласно признаку Декарта, все его коэффициенты k_j должны быть здесь положительные. Согласно же *иерархическому* признаку, для небиномиального уравнения должны выполняться неравенства $\bar{m}_1 > \bar{m}_2 > \dots > \bar{m}_n > 0$. Причём 1-е из них слева есть и 1-е ограничительное неравенство для $\bar{m}_2$. Последующие, разумеется, будут слабее ограничительных неравенств для $\bar{m}_3, \dots, \bar{m}_n$. Они являются необходимыми условиями выполнения этих неравенств. Поэтому, хотя с формальной точки зрения исходное требование к этой иерархии медиан проверять не обязательно, но с практической точки зрения предварительная проверка иерархии медиан является простым и мощным фильтром для отсева уравнений с заведомо некорректными коэффициентами! (При $\bar{m}_1 = \bar{m}_2 = \dots = \bar{m}_n > 0$ тривиально следует, что уравнение имеет биномиальную форму, т. е. $x_1 = x_2 = \dots = x_n = \mu > 0$.) Но при $\bar{m}_j = \bar{m}_t > 0$, например $\bar{m}_1 = \bar{m}_2$, и при неравных им некоторых $\bar{m}_i$ исходное уравнение обязательно содержит комплексные сопряжённые корни. Тот же факт имеет место при любом ином нарушении иерархии положительных медиан, в силу иерархического признака. В качестве медиан здесь фигурируют средние алгебраические от корней уравнения. *Геометрическая интерпретация* отображена на рис. 8 (1) и (2) при $n = 4$ и на рис. 9 (1) и (2) при $n = 5$.

Выразим из уравнения (298) коэффициент $\bar{m}_n^n$ как функцию от x :

$$\bar{m}_n^n(x) = (-1)^{n-1} \cdot x^n + (-1)^{n-2} \cdot C_n^1 \cdot \bar{m}_1 \cdot x^{n-1} + \dots$$

$$\dots + (-1)^{n-1-j} \cdot C_n^j \cdot \bar{m}_j \cdot x^{n-j} + \dots - C_n^{n-2} \cdot \bar{m}_{n-2}^{n-2} \cdot x^2 + C_n^{n-1} \cdot \bar{m}_{n-1}^{n-1} \cdot x. \quad (327)$$

Применив к ней теорему Ферма (§ 1.1), получаем сопутствующее уравнение степени $(n - 1)$ в той же приведённой форме, т. е. с теми же самыми коэффициентами, но вплоть до порядка $(n - 1)$:

$$\frac{1}{n} \cdot \frac{d\bar{m}_n^n}{dx} = (-1)^{n-1} \cdot y_{n-1}(x) = (-1)^{n-1} \cdot x^{n-1} + (-1)^{n-2} \cdot C_{n-1}^1 \cdot \bar{m}_1 \cdot x^{n-2} + \dots$$

$$\dots + (-1)^{n-1-j} \cdot C_{n-1}^j \cdot \bar{m}_j \cdot x^{n-1-j} + \dots - C_{n-1}^{n-2} \cdot \bar{m}_{n-2}^{n-2} \cdot x + \bar{m}_{n-1}^{n-1} = 0. \quad (328)$$

Начальные значения при $x = 0$: $\bar{m}_n^n(0) = 0$, $\frac{d\bar{m}_n^n}{dx}(0) = n \cdot \bar{m}_{n-1}^{n-1} > 0$.

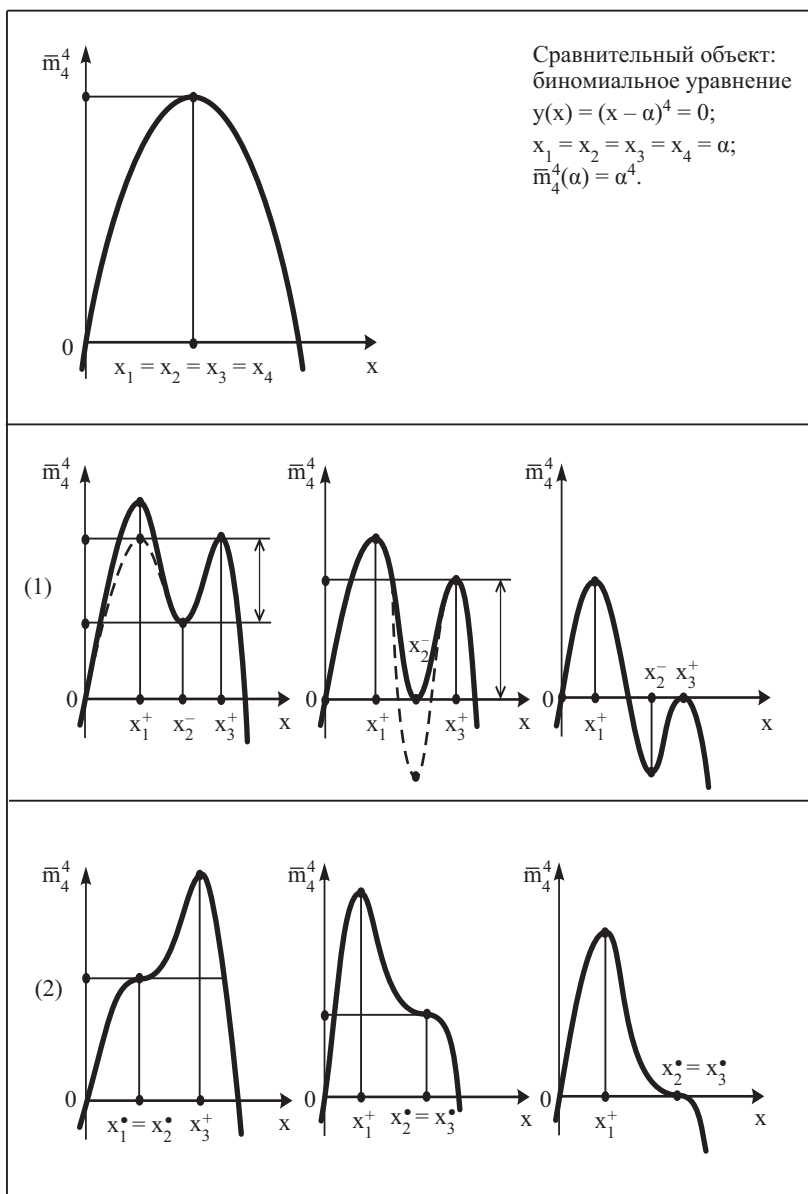


Рис. 8 (1) и (2). Варианты допустимых интервалов для корректных значений высшего коэффициента $\bar{m}_4^4$ алгебраического уравнения 4-й степени при корректных значениях $\bar{m}_1$, $\bar{m}_2$ и $\bar{m}_3$:

(1) — без кратных корней, (2) — с кратными корнями.

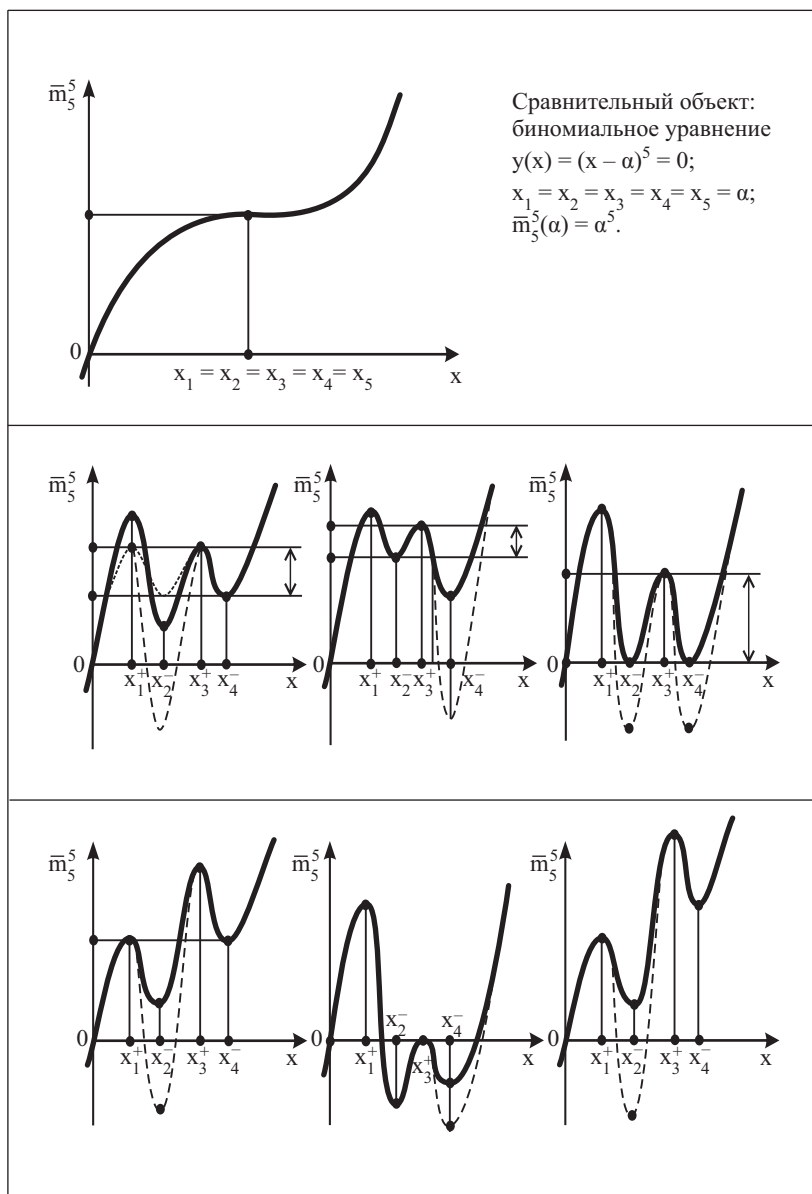


Рис. 9 (1). Характерные варианты допустимых интервалов и отсутствия таковых вообще для корректных значений высшего коэффициента $\bar{m}_5^5$ алгебраического уравнения 5-й степени при корректных значениях $\bar{m}_1, \bar{m}_2, \bar{m}_3, \bar{m}_4$ — без кратных корней.

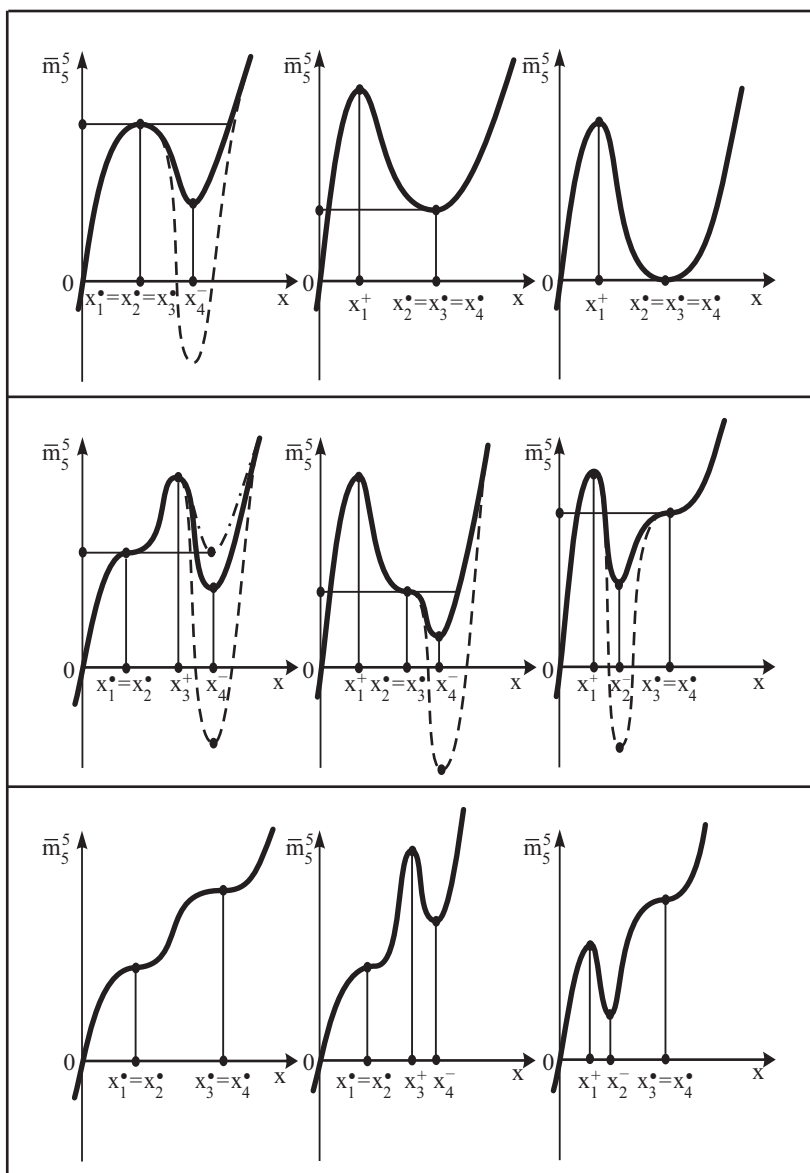


Рис. 9 (2). Характерные варианты одиночных корректных значений и отсутствия таковых для высшего коэффициента $\bar{m}_5^5$ алгебраического уравнения 5-й степени при корректных значениях $\bar{m}_1, \bar{m}_2, \bar{m}_3, \bar{m}_4$ — с кратными корнями.

Вначале обратимся к *простому* алгебраическому уравнению (298). Для уравнения с корректными коэффициентами его корни образуют иерархию $0 < x_1 < x_2 < \dots < x_n$. С учётом вышеуказанных начальных значений и из чисто геометрических соображений (теорема Ролля) — см. рис. 8(1) и 9(1) — заключаем, что активная часть кривой $\bar{m}_n^n(x)$, т. е. (327), с возрастанием $x > 0$ должна обязательно в силу непрерывности последовательно проходить череду максимумов $\bar{m}_n^n(x^+)$ и минимумов $\bar{m}_n^n(x^-)$, причём всего $n - 1$ раз, т. е. между 0 и x_1 , между x_1 и x_2 , между x_2 и x_3 , ..., между x_{n-1} и x_n . Кроме того, ось абсцисс x , смещённая вверх на величину $\bar{m}_n^n$, должна пересекать кривую $\bar{m}_n^n(x)$ в n точках $x_1, x_2, x_3, \dots, x_{n-1}, \dots, x_n$, т. е. корнях исходного простого уравнения степени n . Причём при $x = 0$ и в корневых точках последовательно чередуются положительные и отрицательные первые производные функции $\bar{m}_n^n(x)$. С другой стороны, в корневых точках $x_1^+, x_2^-, x_3^+, \dots$ *сопутствующего* уравнения (328) степени $n - 1$, в которых первые производные функции $\bar{m}_n^n(x)$ нулевые, чередуются отрицательные и положительные первые производные функции (328), т. е. функции $(-1)^{n-1} \cdot y_{n-1}(x)$. Они же, с точностью до множителя n , суть вторые производные функции $\bar{m}_n^n(x)$, определяющие в этих корневых точках характер её $n - 1$ экстремумов.

Из тех же геометрических соображений — рис. 8(1) и 9(1) следует, что для корректности коэффициента $\bar{m}_n^n$ и медианы $\bar{m}_n$ необходимо и достаточно, чтобы они отвечали тождественным неравенствам:

$$\max\langle 0, \max[\min \bar{m}_n^n(x_i^-)] \rangle \leq \bar{m}_n^n \leq \min[\max \bar{m}_n^n(x_i^+)] \leftrightarrow \quad (329)$$

$$\leftrightarrow \max\langle 0, \max[\min [\bar{m}_n(x_i^-)] \rangle \leq \bar{m}_n \leq \min[\max[\bar{m}_n(x_i^+)]]. \quad (330)$$

Причём $\min[\max \bar{m}_n^n(x_i^+)] = \alpha$, $\max[\min \bar{m}_n^n(x_i^-)] = \beta$.

Эти геометрические неравенства для корректных значений высших коэффициента и алгебраической медианы обобщают ранее полученные частные неравенства (323), (324) для $n = 3$. Здесь также знак равенства верхнему или нижнему экстремуму, в том числе нулевому при $\beta = 0$ (но не просто нулю при $\beta < 0$) отвечает появлению кратного корня в исходном уравнении (298), т. е. отвечает нарушению его простоты. Поэтому для рассматриваемого *простого* уравнения знаки равенства в (329), (330) не должны иметь места.

Отсюда же следует, что для корректности $\bar{m}_n^n$ в простом уравнении (298) чисто геометрически необходимо, чтобы сопутствующее ему приведённое уравнение (328) степени $n-1$ также было простым и имело вещественные положительные корни $0 < x_1^+ < x_2^- < x_3^+, \dots$ — всего $n-1$. Иначе не обеспечивается необходимая, по сути, волнообразность функции $\bar{m}_n^n(x)$ и её n раз пересекаемость со смещённой вверх осью x на величину корректного $\bar{m}_n^n$. Поэтому для корректности высшего коэффициента $\bar{m}_n^n$ необходима корректность предшествующих ему $\bar{m}_1$, $\bar{m}_2^2, \dots, \bar{m}_{n-1}^{n-1}$, или коэффициентов сопутствующего уравнения (328) в той же самой приведённой и простой форме, но степени $n-1$. Как видно, ситуация с проверкой корректности высшего коэффициента $\bar{m}_n^n$ простого уравнения степени n вернулась к исходной, но теперь применительно к высшему коэффициенту $\bar{m}_{n-1}^{n-1}$ для сопутствующего уравнения (328)! Повторяя данную процедуру $n-1$ раз, приходим к изначально корректному простому сопутствующему уравнению 1-й степени $y_1(x) = x - \bar{m}_1 = 0$. Оно имеет решение $x_1 = \bar{m}_1 > 0$. При этом строго обосновываем здесь формулу (294) для корректных значений $\bar{m}_2$, анализируя следующее за ним сопутствующее и обязательно простое уравнение 2-й степени $y_2(x) = x^2 - 2 \cdot \bar{m}_1 \cdot x + \bar{m}_2^2 = 0$ (где $x_1 \neq x_2$).

Обратно, корректность $\bar{m}_1$ и $\bar{m}_2$ необходима для корректности $\bar{m}_3$; корректность $\bar{m}_1$, $\bar{m}_2$ и $\bar{m}_3$ необходима для корректности $\bar{m}_4$ и т. д. вплоть до медианы $\bar{m}_n$. Проверка корректности высшего коэффициента в *прямом* порядке, начиная с $\bar{m}_3^3$, выполняется, согласно требованиям всех геометрических неравенств типа (329), (330). Причём получаемые последовательно в данном процессе сопутствующие уравнения степени 3, 4, ..., t , ..., $(n-1)$ содержат те же многочлены $y_t(x)$, что и каскад (310).

Изменение коэффициента $\bar{m}_n^n$, см. рис. 8(1) и 9(1), эквивалентно смещению оси x по вертикали или вверх или вниз. Исходное положение этой оси отвечает значению $\bar{m}_n^n = 0$. Оно может быть и корректным, и некорректным — в зависимости от того, отвечает ли неравенству (329) или нет.

Точки пересечения оси x на её уровне $\bar{m}_n^n$ с кривой $\bar{m}_n^n(x)$ суть n вещественных положительных корней простого уравнения (298) — до тех пор, пока $\bar{m}_n^n$ находится внутри границ неравенства (329), т. е. между минимальным максимумом и максимальным минимумом $\bar{m}_n^n(x)$, но не ниже 0. Здесь используется то обстоятельство, что форма степенной кривой $y_n(x)$ всегда есть инвариант относительно коэффициента $\bar{m}_n^n$. Причём для родственной ей кривой $\bar{m}_n^n(x)$ форма $y_n(x)$ при чётных n попросту отражается относительно оси x , а при нечётных n — нет.

Заметим также, что эти разнообразные варианты степенных кривых порядка n , условно отображаемые на рис. 8 (при $n = 4$) и на рис. 9 (при $n = 5$), теоретически воспроизводимы в аналитической степенной форме с использованием интерполяционных многочленов Лагранжа порядка n , реализуемых каждый раз по $(n + 1)$ точке. (В частности, эти многочлены могут быть и кратными!)

Далее рассмотрим особенности, которые могут вносить в процедуру проверки корректности *непростого* уравнения его *кратные* корни. Пусть при некотором порядке t в процессе проверки корректности $\bar{m}_t^t$ сопутствующего уравнения степени $t - 1$ выясняется, что он корректен, но ось x на уровне $\bar{m}_t^t$ *касается* либо минимального максимума, либо максимального минимума кривой $\bar{m}_t^t(x)$ — см. рис. 7, 8(1) и 9(1).

Алгебраически это означает, что абсцисса точки касания $x_1^\bullet$ есть как однократный корень сопутствующего уравнения степени $t - 1$, так и двукратный корень сопутствующего уравнения степени t . Затем при переходе к следующим сопутствующим уравнениям степеней $t + 1, \dots$ — вплоть до исходного уравнения степени n кратность корня возрастает каждый раз на 1 и в уравнении $y_n(x) = 0$ она составляет $(n - t + 2)$. Действительно, многочлен $y_t(x)$ получается в результате поэтапного дифференцирования многочлена $y_n(x)$, как в каскаде (310), $n - t$ раз.

Геометрически факт касания оси x на уровне $\bar{m}_t^t$ кривой $\bar{m}_t^t(x)$, т. е. одного из экстремумов в неравенстве (329), отвечает *алгебраически* знаку равенства в нём же либо сверху (справа) минимаксу, либо снизу (слева) максимуму (при $x_1^\bullet \geq 0$).

При дальнейшем увеличении степени сопутствующего уравнения высший коэффициент теряет степень свободы для изменения вверх или вниз, так как при отклонении от своего *единственного* уровня $\bar{m}_j^j(x_1^\bullet)$, где $j > t$, уравнение (298) и j -е сопутствующее уравнение теряет эти кратные корни, что компенсируется появлением неких пар комплексных сопряжённых корней. Но корректное значение высшего коэффициента $\bar{m}_j^j(x_1^\bullet)$ в варианте с кратными корнями $x_1^\bullet$ должно формально также отвечать геометрическим неравенствам (329), (330) — рис. 8(2) и 9(2). Отсюда, с учётом изложенных особенностей, следует, что неравенства типа (329), (330) формально действуют и для алгебраических уравнений с кратными корнями, т. е. непростых! Отсюда исходя из внешнего вида некоторых алгебраических уравнений можно сразу сделать вывод, что они имеют комплексные сопряжённые корни. Приведём отдельные характерные примеры таких уравнений:

$$y_n(x) = (x - a)^n \pm c = 0, \text{ где } n > 2, c \neq 0;$$

$$y_n(x) = (x - a_1) \cdot \dots \cdot (x - a_i)^{d_i} \cdot \dots \cdot (x - a_t) \pm c = 0, \text{ где } d_i > 2, c \neq 0;$$

$$y_n(x) = (x - a_1) \cdot \dots \cdot (x - a_i)^{d_i} \cdot \dots \cdot (x - a_j)^{d_j} \cdot \dots \cdot (x - a_t) \pm c = 0,$$

где d_i и/или $d_j > 2$, $c \neq 0$. При наличии нескольких кратных корней, например, $x_1^\bullet$ и $x_j^\bullet$ см. рис. 9(2), единственное возможное корректное значение высшего коэффициента должно находиться только на одном уровне $\bar{m}_t^t(x_1^\bullet) = \bar{m}_t^t(x_j^\bullet)$. Именно при этом высшем коэффициенте могут сосуществовать совместно все различные *вещественные* кратные корни алгебраического уравнения с вещественным спектром!

В иной трактовке: когда соседние некратные экстремумы кривой $\bar{m}_n^n(x)$, постепенно приближаясь друг к другу, сливаются в одну точку стационарности, число слившихся точек экстремумов и его чётность или нечётность определяют кратность d_i корня $x_1^\bullet$, а также характер и уровень стационарности обычным образом (§ 1.1.1). При этом ось x на уровне $\bar{m}_n^n(x_1^\bullet)$ должна пересекать кривую $\bar{m}_n^n(x)$ ещё $(n - d_i)$ раз, а в случае нескольких кратных корней — ещё $(n - \sum d_i)$ раз.

Однако в схемах чередования знаков первых и вторых производных функции $\bar{m}_n^n(x)$, см. выше, получаемые в точках стационарности x_i^* нулевые производные гораздо удобнее для универсальности этой схемы засчитывать по кратности корней d_i .

В общем плане, сравнивая всевозможные варианты кривых $\bar{m}_n^n(x)$ на рис. 7, 8 и 9, приходим к весьма существенному и общему выводу: *«Только до $n = 4$ всегда можно гарантировать то, что при корректных значениях предшествующих коэффициентов (медиан) найдётся некий интервал, в том числе, возможно, и вырожденный в одно число, для корректных значений высшего коэффициента (медианы). Но при $n > 4$ и корректных значениях предшествующих коэффициентов (медиан) вообще нельзя гарантировать наличие корректного значения высшего коэффициента (медианы), даже вырожденного в одно число.»*

Весьма примечательно здесь то, что критическое значение степени алгебраического уравнения $n = 4$ совпадает с таковым в знаменитых теоремах Абеля и Галуа! (Возможно, это объяснимо топологически.)

При $n > 4$ всегда возможно указать такие особые формы кривых для степенных многочленов алгебраического уравнения $y_n(x)$, при которых имеет место *геометрическое контрнеравенство*:

$$\max\langle 0, \max[\min \bar{m}_n^n(x^-)] \rangle > \min[\max \bar{m}_n^n(x^+)] . \quad (331)$$

Если выполнено данное контрнеравенство, то корректное значение высшего коэффициента отсутствует. Следовательно, тогда же исходное уравнение (298) обязательно имеет комплексные сопряжённые корни. Понятно, что подобный результат может проявиться ещё раньше на каком-то сопутствующем уравнении степени $t < n$.

Теоретически, чтобы проверить корректность всех коэффициентов простого алгебраического уравнения вида (298), вначале целесообразно проверить соответствие его алгебраических медиан $\bar{m}_t^t$ иерархическому неравенству (224), когда проверяется и корректность коэффициентов k_1 и k_2 , согласно неравенству (294); затем последовательно при t от 3-х до n найти максимумы и минимумы функций $\bar{m}_t^t(x)$ с поэтапной проверкой соответствия значений всех высших приведённых коэффициентов $\bar{m}_t^t$ геометрическим неравенствам типа (329). В анализе сопутствующих уравнений, может быть, эффективно применять *компьютерные модели* функций $\bar{m}_t^t(x)$. Причём достаточно исследовать априорные интервалы для аргумента $0 \div [1 + \max |k_j|]$.

Суммируя все вышеизложенные выводы, сформулируем теорему о корректности коэффициентов (медиан) *простого* алгебраического уравнения степени n , что иллюстрирует для него полные требования для вещественности и положительности всех его n различных корней!

Теорема 10 (о корректности коэффициентов и медиан). Для того чтобы все коэффициенты $\bar{m}_t^t$ (медианы $\bar{m}_t$) простого небиномиального вещественного алгебраического уравнения степени $n > 2$ в приведённой форме были корректные, т. е. для того чтобы все корни уравнения были вещественные и положительные, в целом необходимо и достаточно: во-первых, чтобы его малые медианы удовлетворяли иерархическому неравенству $\bar{m}_1 > \bar{m}_2 > \dots > \bar{m}_{n-1} > \bar{m}_n > 0$ (иерархический признак), и, во-вторых, чтобы коэффициенты $\bar{m}_t^t$ отвечали частным геометрическим неравенствам

$$\max\langle 0, \max[\min \bar{m}_t^t(x_i^-)] \rangle < \bar{m}_t^t < \min[\max \bar{m}_t^t(x_i^+)] \quad (332)$$

в последовательности $t = 3, \dots, n$; где x_i^+ и x_i^- — точки максимумов и минимумов для функций $\bar{m}_t^t(x)$.

В частности, из каждого геометрического неравенства (332) следует тождественное ограничительное неравенство для корректных *медиан*:

$$\max\langle 0, \max[\min [\bar{m}_t(x_i^-)] \rangle < \bar{m}_t < \min[\max[\bar{m}_t(x_i^+)] \rangle.$$

Здесь $\min[\max \bar{m}_n(x^+)] = \alpha$, $\max[\min \bar{m}_n(x^-)] = \beta$ — верхний и нижний пределы, обобщающие параметры α и β из формулы (324).

Отметим, что при нарушении *простоты* в цепи уравнений в случае равенства коэффициента $\bar{m}_t^t$ либо слева $\max[\min \bar{m}_t^t(x_i^-)]$, либо справа $\min[\max \bar{m}_t^t(x_i^+)]$, либо $\bar{m}_t^t(x)$ в точке $x_i^\pm$ z -образного или в точке $x_i^\mp$ s -образного стационарного перегиба та же точка стационарности x_i играет роль и *кратного* корня уравнения. Она определяет однозначно единственно возможные значения высших коэффициентов $\bar{m}_j^j(x_i)$, где $j = t, t+1, t+2, \dots, n$, во всех дальнейших непростых сопутствующих уравнениях и в исходном непростом алгебраическом уравнении.

И ещё один важный вывод. Корректный ряд алгебраических медиан $\bar{m}_1, \dots, \bar{m}_t$ всегда можно дополнить корректной медианой $\bar{m}_{t+1}$ только при $t < 4$. Например, корректные $\bar{m}_1, \bar{m}_2, \bar{m}_3$ дополняются всегда корректной $\bar{m}_4$. Но при $t \geq 4$ такое можно сделать не всегда. Например, имеются корректные $\bar{m}_1, \bar{m}_2, \bar{m}_3, \bar{m}_4$, недополняемые корректной $\bar{m}_5$. Поэтому для них количество производящих элементов $x_i > 0$ может равняться только 4-м. То же правило, согласно имеющейся взаимосвязи, распространяется на степенные и реверсивные медианы.

§ 4.5. Нормальное решение и квазирешение вырожденного линейного уравнения — вещественного и комплексного

Векторное линейное уравнение $A\mathbf{x} = \mathbf{a}$, где $\mathbf{x}$ — $n \times 1$ -вектор-аргумент, $\mathbf{a}$ — $m \times 1$ -вектор-постоянная, A — $m \times n$ -матрица-коэффициент, может либо иметь одно единственное решение, либо иметь множественное решение, либо вовсе не иметь решения. Однако в последнем варианте всегда возможно какое-либо *квазирешение*, которое в общем случае может быть единственным или множественным. Причём квазирешение, отвечающее минимуму евклидовой нормы $\|A\mathbf{x} - \mathbf{a}\|$, определяется как *квадратичное*.

С другой стороны, *единственное* решение или квазирешение из их же множества, отвечающее минимуму евклидовой нормы $\|\mathbf{x}\|$ или её квадрата $\|\mathbf{x}\|^2$, определяется как *нормальное*. Кроме того, для решения или квазирешения исходного векторного уравнения $A\mathbf{x} = \mathbf{a}$ методами минимизации базовым понятием является *невязка* $\delta(\mathbf{x}) = A\mathbf{x} - \mathbf{a}$. Решение или квазирешение *вырожденного* уравнения рассматривают в едином ключе, решая задачу минимизации целевой функции вида

$$y(\mathbf{x}) = \|\delta(\mathbf{x})\|^2 = \|A\mathbf{x} - \mathbf{a}\|^2 = (A\mathbf{x} - \mathbf{a})' \cdot (A\mathbf{x} - \mathbf{a}) = \min \quad (333)$$

— в случае *вещественного* аргумента $\mathbf{x}$. Применяя теорему Эйлера — Ферма (§ 1.9), получаем общее решение данной задачи, отвечающее стационарности функции $y(\mathbf{x})$ в декартовом базисе:

$$1/2 \cdot dy/d\mathbf{x} = A'A\mathbf{x} - A'\mathbf{a} = \mathbf{0} \rightarrow \langle \tilde{\mathbf{x}} \rangle \equiv \tilde{\mathbf{x}} + \langle \ker A \rangle. \quad (334)$$

Поскольку здесь матрица Гессе $d^2y/d\mathbf{x}'d\mathbf{x} = A'A$ всегда неотрицательно определённая, то стационарность функции $y(\mathbf{x})$ есть её минимум, что отвечает *минимодульной невязке* уравнения $A\mathbf{x} = \mathbf{a}$.

С целью геометрической интерпретации в евклидовом пространстве введённых характеристик, примем, что $\mathbf{x} \in \mathcal{E}^n$, $\mathbf{a} \in \mathcal{E}^m$, $A \in \mathcal{E}^{m \times n}$ и, следовательно, $\delta(\mathbf{x}) \in \mathcal{E}^m$. Если $\mathbf{a} \in \langle \text{im } A \rangle$, то любой элемент $\tilde{\mathbf{x}}$ из множества $\langle \tilde{\mathbf{x}} \rangle$ есть решение уравнения $A\mathbf{x} = \mathbf{a}$. Но, если $\mathbf{a} \notin \langle \text{im } A \rangle$, то любой элемент $\tilde{\mathbf{x}}$ из множества $\langle \tilde{\mathbf{x}} \rangle$ есть квадратичное квазирешение уравнения $A\mathbf{x} = \mathbf{a}$, отвечающее минимуму евклидовой длины вектора невязки $\delta(\mathbf{x})$. Ортопроекция множества $\langle \tilde{\mathbf{x}} \rangle$ на $\langle \text{im } A' \rangle \equiv \langle \ker A \rangle^\perp$ есть единственное нормальное решение или квазирешение уравнения $A\mathbf{x} = \mathbf{a}$, отвечающее минимуму евклидовой длины вектора $\|\tilde{\mathbf{x}}\|$.

Далее для формального алгебраического выражения возникающих ортопроективных понятий применяются четыре характеристических ортопроектора для A и квазиобратная матрица Мура—Пенроуза A^+ .

$$A^+ A \cdot \langle \tilde{\mathbf{x}} \rangle = A^+ A \cdot (\tilde{\mathbf{x}} + \langle \ker A \rangle) = A^+ A \tilde{\mathbf{x}} = A^+ \cdot A \tilde{\mathbf{x}} = \tilde{\mathbf{x}}^\perp. \quad (335)$$

Это и есть нормальное решение или квазирешение линейного уравнения в зависимости от того, принадлежит или нет вектор $\mathbf{a}$ образу A .

1. Пусть $\mathbf{a} \in \langle \text{im } A \rangle \leftrightarrow \mathbf{a} = A A^+ \mathbf{a}$. Тогда в точности имеем $A \tilde{\mathbf{x}} = \mathbf{a}$. Умножая это слева на A^+ , выражаем формально нормальное решение:

$$A^+ \cdot A \tilde{\mathbf{x}} = A^+ A \tilde{\mathbf{x}} = A^+ \mathbf{a} = \tilde{\mathbf{x}}^\perp. \quad (336)$$

В данном случае минимодульная невязка нулевая:

$$\delta^* = A \tilde{\mathbf{x}} - \mathbf{a} = A \tilde{\mathbf{x}}^\perp - \mathbf{a} = A \cdot A^+ \mathbf{a} - \mathbf{a} = (A A^+ - I) \mathbf{a} = -A A^+ \mathbf{a} = \mathbf{0}. \quad (337)$$

2. Пусть $\mathbf{a} \notin \langle \text{im } A \rangle \leftrightarrow \mathbf{a} \neq A A^+ \mathbf{a}$. Но из (335) имеем $A^+ \mathbf{a} \in \langle \text{im } A' \rangle \leftrightarrow A^+ \mathbf{a} = A^+ A A^+ \mathbf{a}$. Тогда в точности имеем $A^+ A \tilde{\mathbf{x}} = A^+ \mathbf{a}$. Умножая это слева на $(A^+ A)^+$, выражаем нормальное квазирешение:

$$(A^+ A)^+ \cdot A^+ A \tilde{\mathbf{x}} = (A^+ A)^+ \cdot A^+ \mathbf{a} = A^+ \cdot (A^+ A)^+ \cdot A^+ \mathbf{a} = A^+ \mathbf{a} = A^+ A \tilde{\mathbf{x}} = \tilde{\mathbf{x}}^\perp. \quad (338)$$

Следовательно, нормальные решения и квазирешения выражаются одинаково. Но в последнем случае минимодульная невязка ненулевая:

$$\delta^* = A \tilde{\mathbf{x}} - \mathbf{a} = A \tilde{\mathbf{x}}^\perp - \mathbf{a} = A \cdot A^+ \mathbf{a} - \mathbf{a} = (A A^+ - I) \mathbf{a} = -A A^+ \mathbf{a} \neq \mathbf{0}. \quad (339)$$

По геометрической сути, она есть ортоантипроекция $\mathbf{a}$ на ядро $\langle \ker A' \rangle$. Поэтому имеет место общее алгебро-геометрическое неравенство:

$$\|A\mathbf{x} \pm \mathbf{a}\|^2 = (A\mathbf{x} \pm \mathbf{a})' \cdot (A\mathbf{x} \pm \mathbf{a}) \geq \mathbf{a}' \cdot A A^+ \cdot \mathbf{a} \geq 0. \quad (340)$$

Или вообще для невязок имеем: $\|\delta(\pm \mathbf{x})\|^2 \geq \mathbf{a}' \cdot A A^+ \cdot \mathbf{a} \geq 0$.

В свою очередь, для линейного уравнения $Ax = a$ с комплексными элементами на основе эрмитовой комплексификации (§ 3.1) задача (333) приобретает форму:

$$\begin{aligned} y &= y(x, \bar{x}) = y(\bar{x}, x) = \bar{y} = \\ &= \|\delta(x)\|^2 = \|Ax - a\|^2 = (Ax - a)^* \cdot (Ax - a) = \min. \end{aligned} \quad (341)$$

Применяя к данной частной задаче минимизации теорему 7 (§ 3.4), получаем её общее решение, отвечающее стационарности $y(x, \bar{x})$:

$$\partial y / \partial \bar{x} = A^* Ax - A^* a = 0 \rightarrow \langle \tilde{x} \rangle \equiv \tilde{x} + \langle \ker A \rangle. \quad (342)$$

По той же аналогии выражаем нормальное решение или квазирешение в $\mathcal{H}^n$ и минимодульную невязку линейного уравнения в $\mathcal{H}^m$:

$$\tilde{x}^\perp = A^{\leftarrow *} A \tilde{x} = A^+ \cdot A \tilde{x} = A^+ a, \quad (343)$$

$$\delta^\bullet = A \tilde{x} - a = A \tilde{x}^\perp - a = A \cdot A^+ a - a = (AA^* - I) a = -AA^* a. \quad (344)$$

Аналогично (340) имеем алгебро-геометрическое неравенство:

$$\|Ax \pm a\|^2 = (Ax \pm a)^* \cdot (Ax \pm a) \geq a^* \cdot AA^* a \geq 0. \quad (345)$$

Далее, не нарушая общности решения задачи (с учётом возможного использования эрмитовой аналогии) имеем 5 вариантов.

1) $n = m = \text{rang } A$; $\tilde{x}^\perp = A^{-1}a$ — здесь обычное решение уравнения с невырожденной квадратной матрицей A ; при этом формально $\delta^\bullet = 0$.

2) $n > m = \text{rang } A$; $\tilde{x}^\perp = A^+ a$ — здесь нормальное решение уравнения; $\delta^\bullet = 0$, так как $AA' = Z_m$, т. е. $\langle \text{im } A \rangle \equiv \mathcal{L}^m$. Причём для вычисления A^+ применяется рациональная формула: $A^+ = A' \cdot (AA')^{-1} = A^{-1}_R$.

3) $m > n = \text{rang } A$; $\tilde{x}^\perp = A^+ a$ — здесь единственное решение или квазирешение, так как $A'A = I_n$, т. е. $\langle \ker A \rangle = 0$; при этом $\delta^\bullet = -AA'a$.

Для вычисления A^+ и AA' также применяются рациональные формулы $A^+ = (A'A)^{-1} \cdot A' = A^{-1}_L$, $AA' = I - A^+ A = I - A \cdot A^+$.

4) $n > m > \text{rang } A$; $\tilde{x}^\perp = A^+ a$ — нормальное решение или квазирешение в зависимости от значения $\delta^\bullet = -AA'a$.

5) $m > n > \text{rang } A$; $\tilde{x}^\perp = A^+ a$ — нормальное решение или квазирешение в зависимости от значения $\delta^\bullet = -AA'a$.

В вариантах 4 и 5 формульные выражения для ортопроективных характеристик через элементы A и a производятся *точно* через самые общие формулы для характеристических проекторов и квазиобратной матрицы Мура — Пенроуза. Эти формулы и структуры фигурирующих в них специальных матриц установлены автором ещё в начале 1981 г., но были опубликованы гораздо позднее в его монографии [27]. Для их вывода, применительно к частным задачам типа (333) и (341), автор изначально использовал аналитический *предельный метод решения*, изложенный в § 2.4 в трёх различных формах. Выбор композиционной функции — либо (108), либо (110), либо (113) тут на окончательные результаты не влияет. Отправной точкой для исследования послужила тогда известная публикация А. Н. Тихонова 1965 г. [36]. Например:

$$y(x) = 1/2 \cdot x' \cdot x, \quad h(x) = Ax - a, \quad \varphi(x) = 1/2 \cdot h'(x) \cdot h(x);$$

$$K_{\alpha}(x) = \varphi(x) + \alpha \cdot y(x), \quad \alpha \rightarrow 0;$$

$$\left\{ \frac{dK_{\alpha}}{dx} \right\}' = (A'A + \alpha \cdot I) \cdot x - A'a = 0 \rightarrow$$

$$\rightarrow \bar{x}^{\perp} = \lim_{\alpha \rightarrow \infty} \left[(A'A + \alpha \cdot I)^{-1} A' \right] \cdot a = A^{+}a. \quad (346)$$

Отсюда далее были получены предельные формулы вида:

$$A^{+} = \lim_{\alpha \rightarrow \infty} \left[(A'A + \alpha \cdot I)^{-1} A' \right] = \lim_{\alpha \rightarrow \infty} \left[A' \cdot (AA' + \alpha \cdot I)^{-1} \right], \quad (347, 348)$$

$$\overleftarrow{A'A} = A^{+} \cdot A = \lim_{\alpha \rightarrow \infty} \left[(A'A + \alpha \cdot I)^{-1} A'A \right] = \lim_{\alpha \rightarrow \infty} \left[A'A \cdot (A'A + \alpha \cdot I)^{-1} \right], \quad (349, 350)$$

$$\overleftarrow{AA'} = A \cdot A^{+} = \lim_{\alpha \rightarrow \infty} \left[AA' \cdot (AA' + \alpha \cdot I)^{-1} \right] = \lim_{\alpha \rightarrow \infty} \left[(AA' + \alpha \cdot I)^{-1} \cdot AA' \right], \quad (351, 352)$$

$$\overrightarrow{A'A} = I - \overleftarrow{A'A} = \lim_{\alpha \rightarrow \infty} \left[\alpha \cdot (A'A + \alpha \cdot I)^{-1} \right], \quad (353)$$

$$\overrightarrow{AA'} = I - \overleftarrow{AA'} = \lim_{\alpha \rightarrow \infty} \left[\alpha \cdot (AA' + \alpha \cdot I)^{-1} \right]. \quad (354)$$

Пределы вычислялись через *резольвенту* матрицы $B = A'A$ или AA' :

$$(B + \alpha \cdot I)^{-1} = \frac{(B + \alpha \cdot I)^{\vee}}{\det(B + \alpha \cdot I)}.$$

После вычисления этих пределов были получены *точные* формулы для ряда характеристических матриц, производимых из них проекторов и квазиобратной матрицы Мура — Пенроуза.

$$\vec{B} = K_1(B, r) / k(B, r), \quad \overset{\leftarrow}{B} = K_2(B, r) / k(B, r). \quad (355, 356)$$

Здесь $k(B, r)$ — *скалярные* характеристические коэффициенты матрицы порядка r (§ 4.1); $K_1(B, r)$ и $K_2(B, r)$ — *матричные* характеристические коэффициенты матрицы B порядка r (1-го и 2-го рода), вычисляемые через её элементы, согласно их специальной структуре. Именно через них в общем случае выражаются аффинные проекторы (355) и (356).

$$\begin{aligned} A^+ &= A' \cdot K_1(AA', r-1) / k(AA', r) = K_1(A'A, r-1) \cdot A' / k(A'A, r) = \\ &= \overset{\bullet}{A}(r) / k(AA', r) = \overset{\bullet}{A}(r) / k(A'A, r), \end{aligned} \quad (357—360)$$

где $\overset{\bullet}{A}(r)$ — характеристическая матрица для прямоугольной (в общем случае) матрицы A , вычисляемая либо через вещественные, либо через комплексные элементы, согласно её специальной структуре.

Для комплексного уравнения $Ax = a$ в формулах (346)—(354) и (357)—(360) фигурирует эрмитово транспонирование. (Соответственно в предельном методе аналитической оптимизации применялось именно *формальное дифференцирование*!)

Согласно (335), нормальное решение линейного уравнения $Ax = a$ геометрически есть основание перпендикуляра, опущенного из центра декартова базиса на плоскость $\langle \tilde{x} \rangle$. По сути, формулы (336), (357)—(360) вместе дают точное и общее решение задачи аналитической геометрии: «Найти основание перпендикуляра, опущенного из центра координат O на некую плоскость, заданную линейным уравнением $Ax = a$.» Это есть задача на условную минимизацию целевой функции $y = y(x) = \|x\|^2$, где x — векторная переменная, *ограниченная* уравнением $h(x) = A_1x - a = 0$. Однако ту же плоскость можно задать и *параметрическим* способом — через линейно зависимую переменную $\tilde{x} = x(u) = A_2u$, как, например, это в самом общем случае применялось ранее в § 2.1. Тогда имеем:

$$\langle \tilde{x} \rangle \equiv \tilde{x} + \langle \text{im } A_2 \rangle \equiv \tilde{x} + \langle \ker A_1 \rangle, \text{ где } A_1 \overset{\leftarrow}{\rightarrow} A_2 \text{ и } A_1 \overset{\leftarrow}{\rightarrow} A_2' \leftrightarrow A_1' \overset{\leftarrow}{\rightarrow} A_2'.$$

Умножая тут слева любое частное решение $\tilde{x}$ на $A_1'A_1$ или $A_2'A_2'$, в итоге в обоих вариантах задания плоскости получаем одно и то же нормальное решение, или нормальную проекцию (перпендикуляр) $\tilde{x}^\perp$.

Нормальное решение $\tilde{x}^\perp$ не инвариантно к масштабным и вообще к линейным преобразованиям, за исключением лишь ортогональных, что прямо следует из формул для квазиобратной матрицы типа (357), (358).

§ 4.6. Экстремумы отношения Релея

Отношение Релея для вещественной *симметричной* матрицы S или более общо — для комплексной эрмитовой матрицы H определяется как скалярная функция от $\mathbf{x} \in \mathcal{E}^n$ (в декартовом базисе) вида [25, 32]:

$$\mathcal{R}e(\mathbf{x}) = \frac{\mathbf{x}' \cdot S \cdot \mathbf{x}}{\mathbf{x}' \cdot \mathbf{x}} \quad (\mathbf{x} \neq \mathbf{0}). \quad (361)$$

Нетрудно видеть, что в однозначной трактовке отношение Релея есть функция от направления радиус-вектора $\mathbf{x}$ из начала координат, так как оно инвариантно к прямо пропорциональному преобразованию $\mathbf{x} \rightarrow c \cdot \mathbf{x}$, где $c \neq 0$ — свободный скаляр. Вектор-аргумент в (361) может быть весьма разнообразной природы. Например, в формуле (58) это были дифференциалы независимого аргумента. (Здесь отношение Релея оценивало степень растяжения-сжатия координат.) Главный интерес в линейной алгебре к отношению Релея состоит в том, что его экстремумы суть максимальное и минимальное собственные значения матрицы S , т. е. $\max\langle\mu_i\rangle$ и $\min\langle\mu_i\rangle$ (см. также в этой главе §§ 4.1, 4.2).

Весьма важно то, что *монарное* отношение Релея (361) — инвариант ортогональных модальных преобразований R ($R'R = RR' = I$). Поэтому его как инвариантную функцию можно изучать в любом подходящем ортогонально преобразованном базисе $\tilde{E}$. Если исходные координаты заданы в декартовом базисе $\tilde{E}$, то в некотором другом декартовом базисе $\tilde{E}_0 = R \cdot \tilde{E}$ вещественная матрица S из (361) представляется в форме *диагональной матрицы* собственных значений:

$$S = R \cdot D_\mu \cdot R' \leftrightarrow D_\mu = R' \cdot S \cdot R.$$

Отношение (361) в базисе $\tilde{E}_0$ от аргумента $\mathbf{u} = R' \mathbf{x}$, сохраняя свои абсолютные значения, приобретает упрощённый функциональный вид:

$$\mathcal{R}e(\mathbf{x}) = \mathcal{R}e(\mathbf{u}) = \frac{\mathbf{u}' \cdot D_\mu \cdot \mathbf{u}}{\mathbf{u}' \cdot \mathbf{u}} = \frac{\sum_{t=1}^n \mu_t u_t^2}{\sum_{t=1}^n u_t^2} \quad (\mathbf{u} \neq \mathbf{0}). \quad (362, 363)$$

Собственные значения μ_t могут быть любыми вещественными числами. В частности, если $\mu_j = 0$, то $\det S = 0$. Но, как правило, здесь $\det S \neq 0$.

Найдём в новом декартовом базисе $\tilde{E}_0 = R \cdot \tilde{E}$ области и значения всех стационарностей отношения Релея и, главное, — его *экстремумов*.

Применяя теорему Эйлера — Ферма, получаем *градиентное уравнение* для искомых областей стационарности:

$$\begin{aligned} \frac{1}{2} \cdot \frac{d\mathcal{R}e}{du} &= \frac{\left(u_1^2 + \dots + u_n^2\right) \cdot |\mu_1 \cdot u_1, \dots, \mu_n \cdot u_n| - \left(\mu_1 \cdot u_1^2 + \dots + \mu_n \cdot u_n^2\right) \cdot |u_1, \dots, u_n|}{u_1^2 + \dots + u_n^2} = \\ &= \frac{|\mu_1 \cdot u_1, \dots, \mu_n \cdot u_n|}{u_1^2 + \dots + u_n^2} - \mathcal{R}e(\mathbf{u}) \cdot \frac{|u_1, \dots, u_n|}{u_1^2 + \dots + u_n^2} = 0. \end{aligned} \quad (364)$$

Отсюда следует итоговое уравнение:

$|\mu_1 \cdot u_1, \dots, \mu_n \cdot u_n| = \mathcal{R}e(\mathbf{u}) \cdot |u_1, \dots, u_n|$, или в матричной форме записи

$$D_\mu \cdot \mathbf{u} = \mathcal{R}e(\mathbf{u}) \cdot \mathbf{u} = \mu \cdot \mathbf{u}. \quad (365)$$

При $\mu_i \neq \mu_j$ оно даёт решения в виде n координатных осей декартова базиса $\tilde{E}_0$, т. е. в виде n *попарно ортогональных* собственных векторов $\mathbf{u}_t = \mathbf{c}_t \cdot \mathbf{e}_t$ матрицы D_μ с точностью до свободных множителей $\mathbf{c}_t \neq 0$, где $\mathbf{e}_t$ — единичные орты декартовых координат: $\mathbf{e}_t' \cdot \mathbf{e}_t = 1$, $\mathbf{e}_j' \cdot \mathbf{e}_i = 0$ ($\mathbf{e}_j \cdot \mathbf{e}_i' = Z$). В направлениях из центра базиса $\tilde{E}_0$ отношение Релея, согласно (365), принимает значения $\mu_t = \mathcal{R}e(\mathbf{u}_t)$, где $t = 1, 2, \dots, n$ (по осям базиса).

В исходном базисе $\tilde{E}$ те же направления задают собственные векторы матрицы S , получаемые здесь преобразованием $\mathbf{x}_t = R\mathbf{u}_t$; соответственно $\mathcal{R}e(\mathbf{x}_t) = \mathcal{R}e(\mathbf{u}_t) = \mu_t$. Уравнение (365) в $\tilde{E}$ имеет форму:

$$S \cdot \mathbf{x} = \mathcal{R}e(\mathbf{x}) \cdot \mathbf{x} = \mu \cdot \mathbf{x}. \quad (366)$$

Его решения при $\mu_i \neq \mu_j$ есть n попарно ортогональных собственных векторов $\mathbf{x}_t = \mathbf{c}_t \cdot \mathbf{r}_t$ матрицы S с точностью до свободных множителей $\mathbf{c}_t \neq 0$, где $\mathbf{r}_t$ — единичные собственные векторы S : $\mathbf{r}_t' \cdot \mathbf{r}_t = 1$, $\mathbf{r}_j' \cdot \mathbf{r}_i = 0$ ($\mathbf{r}_j \cdot \mathbf{r}_i' = Z$), но $\mathbf{r}_t \cdot \mathbf{r}_t' = \vec{S}_t$, где $S_t = \{S - \mu_t \cdot I\}$ — t -я собственная матрица ранга $(n - 1)$, $\vec{S}_t$ — ортопроектор на $\langle \ker S_t \rangle \equiv \langle \text{im } \mathbf{x}_t \rangle \equiv \langle \text{im } \mathbf{c}_t \cdot \mathbf{r}_t \rangle$.

Но в случаях вырожденности некоторые собственные значения μ_t имеют кратность $k_t > 1$. Тогда их линейно независимые собственные векторы $\mathbf{x}_t$ в кластерных прямых суммах составляют систему m *попарно ортогональных* собственных $n \times k$ -*линеоров* $A_t = |\mathbf{x}_{t/1}, \dots, \mathbf{x}_{t/k}|$, которые задают собственные линейные подпространства $\langle \text{im } A_t \rangle$ размерности k_t . На этих подпространствах отношение Релея (361) стационарно и имеет значение μ_t . (В прямой ортогональной сумме они составляют $\mathcal{E}^n$.)

Если каждый собственный $n \times k$ -линеор в базисе $\tilde{E}$ составляется из k_t ортонормированных векторов $\mathbf{r}_t$, то он принимает структуру и имеет свойства квазиортogonalной матрицы $Rq_t = |\mathbf{r}_{t,1}, \dots, \mathbf{r}_{t,k}|$, для которой имеем: $Rq_t' \cdot Rq_t = I$, $Rq_j' \cdot Rq_i = Z$ ($Rq_j \cdot Rq_i' = Z$), но $Rq_t \cdot Rq_t' = \vec{S}_t$, где $S_t = \{S - \mu_t \cdot I\}$ — t -я собственная матрица для S ранга $(n - k_t)$; $\vec{S}_t$ — ортопроектор на $\langle \ker S_t \rangle \equiv \langle \text{im } Rq_t \rangle \equiv \langle \text{im } A_t \rangle$, применяемый в общем спектральном разложении матрицы типа $S = \sum_{t=1}^m \mu_t \cdot \vec{S}_t$ $\left(\sum_{t=1}^m \vec{S}_t = I \right)$, где m — количество различных μ_t . Полное множество собственных линеоров с значением μ_t получается, например, из ортонормированного линеора через свободную несингулярную $k_t \times k_t$ -матрицу C_t :

$$\langle A_t \rangle \equiv \langle Rq_t \cdot C_t \rangle. \quad (367)$$

Итак, выше были полностью охарактеризованы как все области стационарностей, так и стационарные значения отношения Релея (361). Для идентификации этих стационарностей и, в частности, выявления экстремумов отношения Релея (максимума и минимума) вычисляем его матрицу Гессе с использованием декартова базиса $\tilde{E}_0$:

$$\begin{aligned} \frac{1}{2} \cdot \frac{d^2 \mathcal{R}e}{du' du} = & \frac{\left(u_1^2 + \dots + u_n^2 \right) \cdot D_\mu - 2 \cdot \begin{vmatrix} \mu_1 \cdot u_1 \\ \dots \\ \mu_n \cdot u_n \end{vmatrix} \cdot |u_1, \dots, u_n|}{u_1^2 + \dots + u_n^2} - \\ & - \left[\frac{\begin{vmatrix} u_1 \\ \dots \\ u_n \end{vmatrix}}{u_1^2 + \dots + u_n^2} \cdot 2 \cdot \left(\frac{|\mu_1 \cdot u_1, \dots, \mu_n \cdot u_n|}{u_1^2 + \dots + u_n^2} - \mathcal{R}e(\mathbf{u}) \cdot \frac{|u_1, \dots, u_n|}{u_1^2 + \dots + u_n^2} + \right. \right. \\ & \left. \left. + \mathcal{R}e(\mathbf{u}) \cdot \frac{\left(u_1^2 + \dots + u_n^2 \right) \cdot I - 2 \cdot \begin{vmatrix} u_1 \\ \dots \\ u_n \end{vmatrix} \cdot |u_1, \dots, u_n|}{u_1^2 + \dots + u_n^2} \right] \right]. \quad (368) \end{aligned}$$

Отсюда следует, что при $\mu_i \neq \mu_j$ на n координатных осях $\mathbf{u}_t$ базиса $\tilde{\mathbf{E}}_0$ имеем соответственно n различных собственных значений матрицы Гессе μ_t в той же её диагональной форме D_μ :

$$\frac{d^2 \mathcal{R}e}{d\mathbf{u}'d\mathbf{u}}(\mathbf{u}_t) = \frac{2}{\mathbf{u}_t' \cdot \mathbf{u}_t} \cdot (D_\mu - \mu_t \cdot \mathbf{I}) = \frac{2}{u_t^2} \cdot (D_\mu - \mu_t \cdot \mathbf{I}). \quad (369)$$

Пусть $\mu_{\max} = \max\langle \mu_t \rangle$ и $\mu_{\min} = \min\langle \mu_t \rangle$. Тогда, согласно (369), на координатной оси $\mathbf{u}(\mu_{\max})$ и только на ней целевая функция (361) принимает *нестрогий максимум*, вырожденный вдоль именно этой оси; а на координатной оси $\mathbf{u}(\mu_{\min})$ и только на ней целевая функция (361) принимает *нестрогий минимум*, вырожденный именно вдоль этой оси. Это видно по знакам диагональных элементов матрицы Гессе (369): в первом случае $(\mu_i - \mu_{\max}) \leq 0$ и во втором случае $(\mu_i - \mu_{\min}) \geq 0$. Соответственно на других координатных осях $\mathbf{u}_t$ знаки диагональных элементов $(\mu_i - \mu_t)$ матрицы Гессе (369), кроме нулевого при $i = t$, обязательно различаются. Поэтому на этих областях стационарности целевая функция (361) принимает *нестрогие стационарные седловины*. (О подобных критериях характера стационарности говорилось в § 1.9.) Причём при $\mu_i \neq \mu_j$ степени вырожденности всех этих стационарностей равны 1. Однако, когда какое-либо собственное значение μ_t имеет тут кратность $k_t > 1$, тогда степень вырожденности t -й стационарности возрастает до k_t . В частности, это может относиться также к $\mu_{\max}$ и $\mu_{\min}$, т. е. к максимуму и минимуму функции (361).

В исходном базисе $\tilde{\mathbf{E}}$ матрица Гессе отношения Релея при $\mathbf{x} = \mathbf{x}_t$ имеет форму:

$$\frac{d^2 \mathcal{R}e}{d\mathbf{x}'d\mathbf{x}}(\mathbf{x}_t) = \frac{2}{\mathbf{x}_t' \cdot \mathbf{x}_t} \cdot (S_\mu - \mu_t \cdot \mathbf{I}) = \frac{2S_t}{\mathbf{x}_t' \cdot \mathbf{x}_t}. \quad (370)$$

Итак, для вещественной симметричной матрицы $S = S'$ отношение Релея (361) принимает стационарные значения μ_t всегда, когда аргумент является собственным вектором $\mathbf{x}_t$ (при $k_t = 1$) или, наиболее общо, отвечающим ему собственным линеором A_t (при $k_t \geq 1$). Экстремумы — максимум и минимум отношения Релея достигаются на векторах $\mathbf{x}_t$, которым соответствуют максимальное и минимальное собственное значение μ_t матрицы S .

Кроме того, на основании эрмитовой комплексной аналогии, с учётом операций формального анализа (см. §§ 3.2 — 3.4), те же результаты решения этой экстремальной задачи распространяются на отношение Релея для комплексной эрмитовой матрицы $H = H^*$:

$$\mathcal{R}e(\mathbf{x}) = \frac{\mathbf{x}^* \cdot H \cdot \mathbf{x}}{\mathbf{x}^* \cdot \mathbf{x}} \quad (\mathbf{x} \neq \mathbf{0}). \quad (371)$$

Заметим также, что общие формулы (366) и (370) в базисе $\tilde{E}$ могут быть получены тензорным дифференцированием функции (361) и её градиента по векторному аргументу $\mathbf{x}$. Но при таком непосредственном дифференцировании, например, в базисе $\tilde{E}$ нужно руководствоваться тензорными формулами для дифференциалов и тензор-производных:

$$d(\mathbf{x}' \cdot S \cdot \mathbf{x}) = 2d\mathbf{x}' \cdot S \cdot \mathbf{x} = 2(d\mathbf{x}' \cdot S \cdot \mathbf{x})' = 2\mathbf{x}' \cdot S d\mathbf{x}, \quad (372)$$

$$\frac{d(\mathbf{x}' \cdot S \cdot \mathbf{x})}{d\mathbf{x}} = 2\mathbf{x}' \cdot S \quad \leftrightarrow \quad \frac{d(\mathbf{x}' \cdot S \cdot \mathbf{x})}{d\mathbf{x}'} = 2S \cdot \mathbf{x}; \quad (373, 374)$$

$$d(\mathbf{x}^* \cdot H \cdot \mathbf{x}) = 2d\mathbf{x}^* \cdot H \cdot \mathbf{x} = 2(d\mathbf{x}^* \cdot H \cdot \mathbf{x})^* = 2\mathbf{x}^* \cdot H d\mathbf{x}, \quad (375)$$

$$\frac{d(\mathbf{x}^* \cdot H \cdot \mathbf{x})}{d\mathbf{x}} = 2\mathbf{x}^* \cdot H \quad \leftrightarrow \quad \frac{d(\mathbf{x}^* \cdot H \cdot \mathbf{x})}{d\mathbf{x}^*} = 2H \cdot \mathbf{x}; \quad (376, 377)$$

а также стандартными формулами дифференцирования произведения и отношения 2-х скалярных функций.

Отметим, что в самом общем плане для изучения таких скалярных функций полезны следующие формулы:

$$\mathbf{x}' \cdot B \cdot \mathbf{x} = \mathbf{x}' \cdot \frac{B+B'}{2} \cdot \mathbf{x} = \mathbf{x}' \cdot S \cdot \mathbf{x}, \quad \left(\mathbf{x}' \cdot \frac{B-B'}{2} \cdot \mathbf{x} = \mathbf{x}' \cdot K \cdot \mathbf{x} = 0 \right); \quad (378)$$

$$\mathbf{x}^* \cdot B \cdot \mathbf{x} = \mathbf{x}^* \cdot \frac{B+B^*}{2} \cdot \mathbf{x} = \mathbf{x}^* \cdot H \cdot \mathbf{x}, \quad \left(\mathbf{x}^* \cdot \frac{B-B^*}{2} \cdot \mathbf{x} = \mathbf{x}^* \cdot Q \cdot \mathbf{x} = 0 \right). \quad (379)$$

Соответственно матрица B в скалярных функциях отношений типа

$$y(\mathbf{x}) = \frac{\mathbf{x}' \cdot B \cdot \mathbf{x}}{\mathbf{x}' \cdot \mathbf{x}}, \quad y(\mathbf{x}) = \frac{\mathbf{x}^* \cdot B \cdot \mathbf{x}}{\mathbf{x}^* \cdot \mathbf{x}} \quad (\mathbf{x} \neq \mathbf{0}) \quad (380)$$

заменяется везде *без изменения их значений (!)* матрицами S и H . (Все вышеуказанные формулы доказываются непосредственно в матричном представлении, для чего их нужно выписать поэлементно.)

Квадратичные $\mathbf{x}' \cdot \mathbf{S} \cdot \mathbf{x}$ и эрмитовы $\mathbf{x}^* \cdot \mathbf{H} \cdot \mathbf{x}$ формы при *линейных* преобразованиях базиса и соответственно аргумента преобразуются до канонических форм типа $\mathbf{u}' \cdot \mathbf{I}^\pm \cdot \mathbf{u}$ и $\mathbf{u}^* \cdot \mathbf{I}^\pm \cdot \mathbf{u}$. По закону инерции форм Сильвестра — квадратичных и эрмитовых, как известно, при подобных (*конгруэнтных*) преобразованиях знаки собственных значений матрицы не изменяются, а нулевые значения остаются нулевыми. Отметим, что на монарные отношения Релея (361), (371) линейные преобразования не распространяются, так как они теряют свою инвариантность. Но для (378)–(380) допускаются опять-таки только ортогональные (унитарные) преобразования, приводящие 1-ю компоненту В к диагональной форме!

Далее рассмотрим решение задачи на экстремум (стационарность) *бинарного* отношения Релея. Оно определяется как скалярная функция от $\mathbf{x}_1$ и $\mathbf{x}_2 \in \mathcal{A}^n$ (в аффинном базисе) вида:

$$\mathcal{R}e(\mathbf{x}_1, \mathbf{x}_2) = \frac{\mathbf{x}_2' \cdot \mathbf{P} \cdot \mathbf{x}_1}{\mathbf{x}_2' \cdot \mathbf{x}_1} \quad (\mathbf{x}_1, \mathbf{x}_2 \neq \mathbf{0}), \quad (381)$$

где $\mathbf{P}$ — *вещественная* простая матрица с *вещественным* спектром (!) собственных значений. Однозначно бинарное отношение Релея (381) есть функция от направлений по $\mathbf{x}_1$ и $\mathbf{x}_2$ из начала координат, в силу того, что оно инвариантно к прямо пропорциональным преобразованиям типа $\mathbf{x}_1 \rightarrow c_1 \cdot \mathbf{x}_1$ и $\mathbf{x}_2 \rightarrow c_2 \cdot \mathbf{x}_2$, где $c_{1,2} \neq 0$ — свободные скаляры.

Проясним вначале вопрос: при каких дополнительных требованиях бинарное отношение Релея будет инвариантом линейных модальных преобразований? Тогда его как инвариантную функцию можно будет изучать в любом более подходящем аффинно преобразованном базисе. Если исходно координаты заданы в каком-то аффинном базисе $\tilde{\mathbf{E}}$, то в некотором другом базисе $\tilde{\mathbf{E}}_0 = \mathbf{V} \cdot \tilde{\mathbf{E}}$ исходная простая матрица $\mathbf{P}$ представляется в форме *диагональной матрицы* своих собственных значений: $\mathbf{P} = \mathbf{V} \cdot \mathbf{D}_\mu \cdot \mathbf{V}^{-1} \leftrightarrow \mathbf{D}_\mu = \mathbf{V}^{-1} \cdot \mathbf{P} \cdot \mathbf{V}$. Бинарное отношение Релея инвариантно к аффинным преобразованиям базиса тогда и только тогда, когда исходно координаты вектора $\mathbf{x}_1$ выражаются в контравариантном базисе $\tilde{\mathbf{E}}$, координаты вектора $\mathbf{x}_2$ выражаются в ковариантном базисе $\tilde{\mathbf{E}}$. Они суть взаимные аффинные базисы для одного того же вектора $\mathbf{x}$. Бинарное отношение Релея упрощается при переходе в $\tilde{\mathbf{E}}_0$:

$$\mathcal{R}e(\mathbf{x}_1, \mathbf{x}_2) = \frac{\mathbf{x}_2' \cdot \mathbf{P} \cdot \mathbf{x}_1}{\mathbf{x}_2' \cdot \mathbf{x}_1} = \frac{\left(\mathbf{V}^{-1} \cdot \mathbf{u} \right)' \cdot \mathbf{P} \cdot \mathbf{V} \cdot \mathbf{u}}{\left(\mathbf{V}^{-1} \cdot \mathbf{u} \right)' \cdot \mathbf{V} \cdot \mathbf{u}} \equiv \mathcal{R}e(\mathbf{u}) = \frac{\mathbf{u}' \cdot \mathbf{D}_\mu \cdot \mathbf{u}}{\mathbf{u}' \cdot \mathbf{u}} \quad (\mathbf{u} \neq \mathbf{0}). \quad (382, 383)$$

Поскольку $D'_\mu = D_\mu$, то для диагональной матрицы D_μ в $\tilde{E}_0$ правые и левые собственные векторы тождественные. Отсюда векторы $\mathbf{x}_1$ и $\mathbf{x}_2$ для инвариантности бинарного отношения Релея должны претерпевать преобразование точно также, как правые и левые собственные векторы матрицы P , т. е. как $\mathbf{x}_1 = V\mathbf{u}$ при обратном переходе из $\tilde{E}_0$ в $\tilde{E} = V^{-1} \cdot \tilde{E}_0$ и как $\mathbf{x}_2 = V'^{-1}\mathbf{u}$ при обратном переходе из $\tilde{E}_0$ в $\tilde{E} = V' \cdot \tilde{E}_0$.

Ввиду полной идентичности в базисе $\tilde{E}_0$ вышеизученного монарного отношения Релея (362) и бинарного отношения Релея (383), а также инвариантности последнего к взаимным линейным преобразованиям базисов, все ранее полученные результаты решения экстремальной задачи для отношения (361) распространяются на отношение (381), — но пока только в базисе $\tilde{E}_0$. Геометрическая идентичность достигается при дополнительном требовании к ортонормированности базиса $\tilde{E}_0$ в обоих вариантах задачи. (Попросту в $\mathcal{E}^n$ он должен быть декартовым.) В исходных аффинных базисах правые и левые собственные векторы матрицы P расходятся как $\mathbf{x}_t = V\mathbf{u}_t$ и $\mathbf{y}_t = V'^{-1}\mathbf{u}_t$. Соответственно градиентное уравнение при трансляции в базис $\tilde{E}$ подвергается, как и отношение Релея, подобной бинаризации:

$$\left. \begin{aligned} D_\mu \cdot \mathbf{u} &= \mathcal{Re}(\mathbf{u}) \cdot \mathbf{u} = \mu \cdot \mathbf{u}, \\ \mathbf{u}' \cdot D_\mu &= \mathcal{Re}(\mathbf{u}) \cdot \mathbf{u}' = \mu \cdot \mathbf{u}'; \end{aligned} \right\} \leftrightarrow \left\{ \begin{aligned} P \cdot \mathbf{x} &= \mathcal{Re}(\mathbf{x}, \mathbf{y}) \cdot \mathbf{x} = \mu \cdot \mathbf{x}, \\ \mathbf{y}' \cdot P &= \mathcal{Re}(\mathbf{x}, \mathbf{y}) \cdot \mathbf{y}' = \mu \cdot \mathbf{y}'. \end{aligned} \right. \quad (384, 385)$$

При $\mu_i \neq \mu_j$ уравнения (384), (385) дают решения в виде n пар *как бы биортогональных* собственных векторов $\mathbf{x}_t$ и $\mathbf{y}_t$ с точностью до свободных ненулевых множителей: $\mathbf{x}_t = c_t \cdot \mathbf{r}_t$ и $\mathbf{y}_t = c_t \cdot \mathbf{v}_t$, где $\mathbf{r}_t$ и $\mathbf{v}_t$ — *как бы биортонормированные* собственные векторы, выраженные в $\tilde{E}$ и $\tilde{E}$. Причём $\mathbf{r}'_t \cdot \mathbf{v}_t = \mathbf{v}'_t \cdot \mathbf{r}_t = 1$, $\mathbf{r}'_j \cdot \mathbf{v}_i = \mathbf{v}'_j \cdot \mathbf{r}_i = 0$ ($\mathbf{r}_j \cdot \mathbf{v}'_i = \mathbf{v}_j \cdot \mathbf{r}'_i = Z$), но $\mathbf{r}_t \cdot \mathbf{v}'_t = P_t$, где $P_t = P - \mu_t \cdot I$ есть t -я собственная матрица ранга $n - 1$, P_t есть аффинный проектор на $\langle \ker P_t \rangle \equiv \langle \text{im } \mathbf{x}_t \rangle$ параллельно $\langle \text{im } P_t \rangle$.

Если некие μ_t имеют кратность $k_t > 1$, то их линейно независимые собственные векторы в прямых суммах образуют правые $A_{\mathbf{r}_t}$ и левые $A_{\mathbf{v}_t}$ собственные $n \times k$ -*линеоры*, задающие правые $\langle \text{im } A_{\mathbf{r}_t} \rangle$ и левые $\langle \text{im } A_{\mathbf{v}_t} \rangle$ собственные подпространства. На этих парных подпространствах в $\mathcal{E}^n$ бинарное отношение Релея стационарно и имеет значение μ_t .

Линеоры A_t и Al_t геометрически как бы биортгональные. Но если они составляются из k_t как бы биортонормированных правых $\mathbf{r}_t$ и левых $\mathbf{v}_t$ собственных векторов, то принимают структуру и имеют свойства квазибиортгональных матриц $E_{\mathbf{r}_t} = |\mathbf{r}_{t/1}, \dots, \mathbf{r}_{t/k}|$ и $E_{\mathbf{l}_t} = |\mathbf{v}_{t/1}, \dots, \mathbf{v}_{t/k}|$. Причём $E_{\mathbf{r}_t}' \cdot E_{\mathbf{l}_t} = E_{\mathbf{l}_t}' \cdot E_{\mathbf{r}_t} = I$, $E_{\mathbf{r}_j}' \cdot E_{\mathbf{l}_i} = E_{\mathbf{l}_j}' \cdot E_{\mathbf{r}_i} = Z$ ($E_{\mathbf{r}_j} \cdot E_{\mathbf{l}_i}' = E_{\mathbf{l}_j} \cdot E_{\mathbf{r}_i}' = Z$); $E_{\mathbf{r}_t} \cdot E_{\mathbf{l}_t}' = P_t$, где $P_t = P - \mu_t \cdot I$ есть t -я собственная матрица ранга $n - k$, P_t есть аффинный проектор на $\langle \ker P_t \rangle \equiv \langle \text{im } A_t \rangle$ параллельно $\langle \text{im } P_t \rangle$, применяемый в общем спектральном разложении простой матрицы типа

$$P = \sum_{t=1}^m \mu_t \cdot \vec{P}_t \left(\sum_{t=1}^m \vec{P}_t = I \right). \text{ При этом } E_{\mathbf{l}_t} \cdot E_{\mathbf{r}_t}' = \left(\vec{P}_t \right)' = \vec{P}_t' \text{ — аффинный}$$

проектор на $\langle \ker P_t' \rangle \equiv \langle \text{im } Al_t \rangle$ параллельно $\langle \text{im } P_t' \rangle$. Причём полные множества пар правых и левых собственных линеоров с значением μ_t производятся, например, из пары как бы биортонормированных линеоров умножением справа на свободную несингулярную $k_t \times k_t$ -матрицу C_t .

Для иллюстрации экстремальных свойств отношений Релея (361) и (381) выпишем достаточно важные неравенства, имеющие прикладное значение, в том числе в процедурах оптимизации целевых функций.

$$\mu_{\min} \cdot \mathbf{x}' \cdot \mathbf{x} \leq \mathbf{x}' \cdot S \cdot \mathbf{x} \leq \mu_{\max} \cdot \mathbf{x}' \cdot \mathbf{x}. \quad (386)$$

Согласно эрмитовой аналогии (§ 3.1), имеем:

$$\mu_{\min} \cdot \mathbf{x}^* \cdot \mathbf{x} \leq \mathbf{x}^* \cdot H \cdot \mathbf{x} \leq \mu_{\max} \cdot \mathbf{x}^* \cdot \mathbf{x}. \quad (387)$$

В частности, для евклидовых (эрмитовых) пространств имеем:

$$\mu_{\min} \cdot \|\mathbf{x}\|^2 \leq \mathbf{x}' \cdot S \cdot \mathbf{x} \leq \mu_{\max} \cdot \|\mathbf{x}\|^2. \quad (388)$$

Если $G = G'$ — матрица Гессе скалярной функции $y(\mathbf{x})$, то имеем:

$$\mu_{\min} \cdot d\mathbf{x}' d\mathbf{x} \leq d\mathbf{x}' \cdot G \cdot d\mathbf{x} \leq \mu_{\max} \cdot d\mathbf{x}' d\mathbf{x}. \quad (389)$$

Алгебраический смысл этих неравенств интерпретируется исходя из инвариантности квадратичных или эрмитовых форм при допустимых для них преобразованиях координат, в том числе в виде (378), (379).

Для вещественной простой матрицы P с вещественным спектром собственных значений и её взаимных собственных векторов $\mathbf{x}_1$ и $\mathbf{x}_2$ с билинейной формой во взаимных аффинных базисах также имеем:

$$\mu_{\min} \cdot \mathbf{x}_2' \cdot \mathbf{x}_1 \leq \mathbf{x}_2' \cdot P \cdot \mathbf{x}_1 \leq \mu_{\max} \cdot \mathbf{x}_2' \cdot \mathbf{x}_1. \quad (390)$$

§ 4.7. Метод наименьших квадратов Гаусса в одномерном и многомерном вариантах

В распространённых прикладных разделах и практических задачах, связанных с аппроксимацией и оптимизацией, часто принимается, что переменная $\mathbf{x} = (x_1, x_2, \dots, x_n)$ задаётся точно, а целевая функция $\tilde{y}(\mathbf{x})$ находится со случайным отклонением от истинного значения $y(\mathbf{x})$. Тогда *аппроксимационный анализ* поведения этой целевой функции трансформируется в *регрессионный анализ*. При этом обычно исходят из *нормальной* линейной регрессии, в которой принимается (возможно, с отдельной статистической проверкой этой гипотезы), что при каждом фиксированном точном значении $\mathbf{x}_k$ выборочное единичное значение случайной ошибки для целевой функции $(\tilde{y}_k - y_k)$ как статистика подчиняется **закону нормального распределения Гаусса** с нулевым математическим ожиданием $\mu = 0$ и постоянной дисперсией $\sigma^2 = \text{const}$. Этот фундаментальный математический закон впервые установил Гаусс исходя из созданного им классического *метода наименьших квадратов* (1821 — 1823 гг.) и анализа распределения случайной ошибки [29, 31].

Далее рассмотрим генезис метода наименьших квадратов Гаусса (вообще одного из наиболее употребительных в обработке результатов наблюдений и экспериментов) — сначала при $n = 1$, а затем при $n > 1$. При этом используется *аналитическая минимизация* целевой функции в форме суммы квадратов отклонений $\tilde{y}_q$ от *прямой линии регрессии функции на переменную* в зависимости от её коэффициентов.

Сразу же отметим, что как в методе наименьших квадратов, так и в корреляционном и регрессионном анализе на его основе применяются линейные (в том числе линеаризованные) функции $y(x)$. Тут x и/или y , в частности, зависимы от других величин, например, $x = t^{-1}$, $y = \ln z$. Тогда говорят о *линеаризации* типа $z(t) \rightarrow y(x)$, т. е. в координатах $\langle y, x \rangle$. В случае $n = 1$ целевая функция выражается линейным образом как

$$y = y(x) = a_0 + a \cdot x. \quad (391)$$

Пусть имеется N единичных опытных значений функции $\tilde{y}_q$ при N значениях $x = x_q$, где $q = 1, 2, \dots, N$ (выборка). Эти значения $\tilde{y}_q$ имеют случайные отклонения от истинных значений y_q . Задача метода состоит в том, чтобы по имеющейся выборке значений аргумента (точных) и функции (неточных) дальше вычислить наилучшие статистические оценки для коэффициентов a и a_0 в (391).

А именно, это суть оценки $\dot{a}$ и $\dot{a}_0$ для априори неизвестных, но предположительно точных коэффициентов в функции (391). Для этого применяется *линейная регрессия* типа $\tilde{y}|x$. Она выражается формулой:

$$\dot{y} = \dot{y}(x) = \dot{a}_0 + \dot{a} \cdot x. \quad (392)$$

Отсюда вычисляются статистические оценки значений функции $\dot{y}_q$ при конкретных значениях x_q . Идея метода наименьших квадратов Гаусса заключается в том, что для вычисления оценок коэффициентов $\dot{a}$ и $\dot{a}_0$ минимизируется сумма квадратов разностей $\dot{y}_q - \tilde{y}_q$. Далее имеем стандартную процедуру минимизации функции от 2-х переменных:

$$\Psi_y = \Psi_y(\dot{a}_0, \dot{a}) = \sum_{q=1}^N (\dot{y}_q - \tilde{y}_q)^2 = \sum_{q=1}^N (\dot{a} \cdot x_q + \dot{a}_0 - \tilde{y}_q)^2 = \min; \quad (393)$$

$$\left. \begin{aligned} \frac{1}{2} \cdot \frac{\partial \Psi_y}{\partial \dot{a}_0} &= \sum_{q=1}^N (\dot{a}_0 + \dot{a} \cdot x_q - \tilde{y}_q) = 0, \\ \frac{1}{2} \cdot \frac{\partial \Psi_y}{\partial \dot{a}} &= \sum_{q=1}^N (\dot{a}_0 + \dot{a} \cdot x_q - \tilde{y}_q) \cdot x_q = 0. \end{aligned} \right\} \quad (394)$$

Используя средние арифметические, из системы (394) получаем систему из 2-х усреднённых уравнений с двумя неизвестными $\dot{a}$ и $\dot{a}_0$:

$$\overline{\tilde{y}} = \dot{a}_0 + \dot{a} \cdot \overline{x}, \quad (395)$$

$$\overline{\tilde{y} \cdot x} = \dot{a}_0 \cdot \overline{x} + \dot{a} \cdot \overline{x^2}. \quad (396)$$

Отсюда имеем искомые оценки для коэффициентов:

$$\begin{aligned} \dot{a} &= (\overline{\tilde{y} \cdot x} - \overline{\tilde{y}} \cdot \overline{x}) / (\overline{x^2} - \overline{x}^2) = \\ &= \left(N \cdot \sum_{q=1}^N \tilde{y}_q x_q - \sum_{q=1}^N \tilde{y}_q \cdot \sum_{q=1}^N x_q \right) / \left[N \cdot \sum_{q=1}^N x_q^2 - \left(\sum_{q=1}^N x_q \right)^2 \right], \end{aligned} \quad (397)$$

$$\begin{aligned} \dot{a}_0 &= (\overline{\tilde{y}} \cdot \overline{x^2} - \overline{\tilde{y} \cdot x} \cdot \overline{x}) / (\overline{x^2} - \overline{x}^2) = \\ &= \left(\sum_{q=1}^N \tilde{y}_q \cdot \sum_{q=1}^N x_q^2 - \sum_{q=1}^N \tilde{y}_q x_q \cdot \sum_{q=1}^N x_q \right) / \left[N \cdot \sum_{q=1}^N x_q^2 - \left(\sum_{q=1}^N x_q \right)^2 \right], \end{aligned} \quad (398)$$

где знаменатель должен быть обязательно ненулевым (!).

Такие значения оптимальны в смысле их наилучшего приближения к точным значениям a и a_0 в (391) — при условии, что ошибки $(\tilde{y}_k - y_k)$ подвержены нормальному распределению с постоянной дисперсией.

Формулы (395)–(398) значительно упрощаются для специальных *центрированных* планов расположения точек x_q :

$$\sum_{q=1}^N x_q = 0 = \bar{x} \rightarrow \dot{a}_0 = \bar{\tilde{y}}, \quad \dot{a} = \overline{\tilde{y} \cdot x} / \overline{x^2}. \quad (399)$$

Ещё весьма важный и упрощённый вариант метода возникает, когда в функции (391) заведомо известно, что $a_0 = 0$. Тогда он преобразуется в *централизованный метод наименьших квадратов*. При этом (395) формально также трансформируется в однородное линейное уравнение. Оценка единственного коэффициента $\dot{a}$ существенно упрощается:

$$\begin{aligned} \bar{\tilde{y}} = \dot{a} \cdot \bar{x} &\leftrightarrow \sum_{q=1}^N \tilde{y}_q = \dot{a} \cdot \sum_{q=1}^N x_q, \\ \dot{a} = \sum_{q=1}^N \tilde{y}_q / \sum_{q=1}^N x_q &= \bar{\tilde{y}} / \bar{x} = \overline{\tilde{y} \cdot x} / \overline{x^2}, \end{aligned} \quad (400)$$

где знаменатель, естественно, ненулевой (!). Формально 2-ые формулы в (399) и (400) совпадают. Искусственным путём такой вариант можно получить преобразованием целевой функции вида:

$$y(x) \rightarrow \varphi(x) = y(x) - a_0 = a \cdot x \sim \dot{y}(x) \rightarrow \dot{\varphi}(x) = \dot{y}(x) - \tilde{y}(0) = \dot{a} \cdot x.$$

Из полученных формул (395)–(398) непосредственно видно, что *выбор масштабов по осям переменной и функции не имеет значения*.

С целью полноты изложения метода при $n = 1$ осталось установить: как проверить статистически, действительно ли y и x связаны линейной зависимостью? Для этого применим простой *корреляционный* подход. В нём те же y и x считают равноценными переменными, содержащими случайные ошибки, выясняя степень наличия между ними *линейной корреляции*. Но, используя данный подход, в итоге будем по-прежнему считать, что x — точная величина, а $\tilde{y}$ содержит случайную ошибку.

Теперь, если гипотеза о линейной связи между y и x верна, то тогда минимизация суммы квадратов отклонений точек $(x_k, \tilde{y}_k)$ от прямой линии регрессии, *взятых вдоль оси x* , должна дать практически те же оценки коэффициентов $\dot{a}_{(x)}$ и $\dot{a}_{0(x)}$, что и полученные выше по оси y , т. е. $\dot{a}_{(y)}$ и $\dot{a}_{0(y)}$. Отношение однородных коэффициентов $\dot{a}$ или $\dot{a}_0$, вычисленных по указанным взаимным вариантам метода наименьших квадратов, должно дать некоторый коэффициент k , характеризующий отклонение целевой функции от линейной зависимости типа (391).

Далее имеем процедуру минимизации, но во 2-м варианте метода:

$$\begin{aligned} \Psi_x = \Psi_x(\dot{a}_0, \dot{a}) &= \sum_{q=1}^N (x_q - \dot{x}_q)^2 = \sum_{q=1}^N [x_q - (\tilde{y}_q - \dot{a}_0) / \dot{a}]^2 = \min; \\ \left. \begin{aligned} \frac{1}{2} \cdot \frac{\partial \Psi_x}{\partial \dot{a}_0} &= \sum_{q=1}^N [x_q - (\tilde{y}_q - \dot{a}_0) / \dot{a}] / \dot{a} = 0, \\ \frac{1}{2} \cdot \frac{\partial \Psi_x}{\partial \dot{a}} &= \sum_{q=1}^N [x_q - (\tilde{y}_q - \dot{a}_0) / \dot{a}] \cdot (\tilde{y}_q - \dot{a}_0) / \dot{a}^2 = 0. \end{aligned} \right\} \quad (401) \end{aligned}$$

Используя средние арифметические, здесь также из первого уравнения системы получаем уравнение типа (395). Поэтапно выражаем оценки $\dot{a}_{(x)}$ и $\dot{a}_{0(x)}$, например, через средние арифметические значения:

$$\left. \begin{aligned} \dot{a}_{(x)} &= (\overline{\tilde{y}^2} - \overline{\tilde{y}}^2) / (\overline{\tilde{y} \cdot x} - \overline{\tilde{y}} \cdot \overline{x}), \\ \dot{a}_{0(x)} &= (\overline{\tilde{y} \cdot x} \cdot \overline{\tilde{y}} - \overline{\tilde{y}^2} \cdot \overline{x}) / (\overline{\tilde{y} \cdot x} - \overline{\tilde{y}} \cdot \overline{x}). \end{aligned} \right\} \quad (402)$$

Те же формулы можно сразу получить через формулы (397), (398), применяя метод наименьших квадратов в координатах $\langle \tilde{y}, x \rangle$, где $\tilde{y}$ является уже как бы аргументом. Тогда $1/\dot{a}_{(x)}$ и $\dot{a}_{0(x)}/\dot{a}_{(x)}$ выражаются формулами (397), (398) с взаимным обменом x и $\tilde{y}$.

Нетрудно проверить, что отношение коэффициентов $\dot{a}$ и отношение коэффициентов $\dot{a}_0$ в обоих взаимных вариантах метода одно и то же:

$$k_{(x, \tilde{y})}^2 = \dot{a}_{(y)} \cdot \frac{1}{\dot{a}_{(x)}} = \dot{a}_{0(y)} \cdot \frac{1}{\dot{a}_{0(x)}} = \frac{(\overline{\tilde{y} \cdot x} - \overline{\tilde{y}} \cdot \overline{x})^2}{(\overline{x^2} - \overline{x}^2) \cdot (\overline{\tilde{y}^2} - \overline{\tilde{y}}^2)}. \quad (403)$$

Это есть квадрат *выборочного коэффициента линейной корреляции* $k_{x\tilde{y}}$ для точной переменной x и неточной функции $\tilde{y}$ (в рассматриваемом аспекте). Алгебраически $k_{x, \tilde{y}}^2$ может находиться в интервале $0 \div 1$, $k_{x\tilde{y}}$ может находиться в интервале $-1 \div +1$ (в силу *корреляционного неравенства*). Чем ближе $k_{x, \tilde{y}}^2$ к единице, тем более вероятна линейная зависимость y от x , и обратно. Обычно *корреляция* устанавливается между случайными скалярными переменными, характеризуя степень наличия между ними линейной связи, а *регрессия* выполняется в форме статистического отображения неточной скалярной функции на точную переменную исходя из предполагаемой между ними линейной связи. (Именно в этом состоит принципиальное различие этих понятий!)

Выборочный коэффициент линейной корреляции в данном случае выражается тождественными друг другу формулами:

$$\begin{aligned}
 k_{(x, \tilde{y})} &= \frac{\overline{\tilde{y} \cdot x} - \bar{\tilde{y}} \cdot \bar{x}}{\left[\left(\overline{x^2} - \bar{x}^2 \right) \cdot \left(\overline{\tilde{y}^2} - \bar{\tilde{y}}^2 \right) \right]^{1/2}} = \\
 &= \frac{\frac{1}{N} \cdot \sum_{q=1}^N (x_q - \bar{x}) \cdot (\tilde{y}_q - \bar{\tilde{y}})}{\left[\frac{1}{N} \sum_{q=1}^N (x_q - \bar{x})^2 \cdot \frac{1}{N} \sum_{q=1}^N (\tilde{y}_q - \bar{\tilde{y}})^2 \right]^{1/2}} = \\
 &= \frac{N \cdot \sum_{q=1}^N x_q \tilde{y}_q - \sum_{q=1}^N x_q \cdot \sum_{q=1}^N \tilde{y}_q}{\left\{ \left[N \cdot \sum_{q=1}^N x_q^2 - \left(\sum_{q=1}^N x_q \right)^2 \right] \cdot \left[N \cdot \sum_{q=1}^N \tilde{y}_q^2 - \left(\sum_{q=1}^N \tilde{y}_q \right)^2 \right] \right\}^{1/2}}. \quad (404)
 \end{aligned}$$

Формулы (403), (404) лежат в основе *корреляционного анализа* для пары случайных величин, между которыми тестируется линейная связь.

Все данные формулы симметричны относительно x и $\tilde{y}$. Весьма интересно то, что линейные преобразования u и x , с точностью до знака, никак не влияют на $k_{x\tilde{y}}$ и тем более на $k_{x,\tilde{y}}^2$ [15, т. 1]. Объясняется этот факт дисперсионно-ковариационной природой величин числителя и знаменателя в (404), но именно при *линейной регрессии* $\tilde{y}|x$.

В числителе (404) фигурирует *выборочная ковариация* x и $\tilde{y}$:

$$v \langle x_q, \tilde{y}_q \rangle = \frac{1}{N} \cdot \sum_{q=1}^N (x_q - \bar{x}) \cdot (\tilde{y}_q - \bar{\tilde{y}}) = \overline{\tilde{y} \cdot x} - \bar{\tilde{y}} \cdot \bar{x}. \quad (405)$$

В знаменателе (404) фигурируют *выборочные дисперсии* для x и $\tilde{y}$:

$$\left. \begin{aligned}
 \Delta^2 \langle x_q \rangle &= \frac{1}{N} \cdot \sum_{q=1}^N (x_q - \bar{x})^2 = \overline{x^2} - \bar{x}^2, \\
 s^{*2} \langle \tilde{y}_q \rangle &= \frac{1}{N} \cdot \sum_{q=1}^N (\tilde{y}_q - \bar{\tilde{y}})^2 = \overline{\tilde{y}^2} - \bar{\tilde{y}}^2.
 \end{aligned} \right\} \quad (406)$$

В пределе, когда $\sigma^2(\tilde{y}) \rightarrow 0$, все случайные выборочные характеристики стремятся к собственным математическим ожиданиям и, естественно,

$$\mu \left\{ \Delta^2 \langle x_q \rangle \right\} = \Delta^2 \langle x_q \rangle. \quad \text{Далее имеем:}$$

$$\begin{aligned}
 \mu \{v \langle x_q, \tilde{y}_q \rangle\} &= \mu \left\{ \frac{1}{N} \cdot \sum_{q=1}^N (x_q - \bar{x}) \cdot (\tilde{y}_q - \bar{y}) \right\} = \\
 &= \frac{1}{N} \cdot \sum_{q=1}^N \left[(x_q - \bar{x}) \cdot \mu (\tilde{y}_q - \bar{y}) \right] = \\
 &= \frac{1}{N} \cdot \sum_{q=1}^N (x_q - \bar{x}) \cdot (y_q - \bar{y}) = v \langle x_q, y_q \rangle, \tag{407}
 \end{aligned}$$

$$\begin{aligned}
 \mu \{d^2 \langle \tilde{y}_q \rangle\} &= \mu \left\{ \frac{1}{N} \cdot \sum_{q=1}^N (\tilde{y}_q - \bar{y})^2 \right\} = \\
 &= \frac{1}{N} \cdot \sum_{q=1}^N \mu \left\{ \left[(y_q - \bar{y}) + (\tilde{y}_q - y_q) - (\bar{y} - \bar{y}) \right]^2 \right\} = \\
 &= \frac{1}{N} \cdot \sum_{q=1}^N \mu \left\{ (y_q - \bar{y})^2 + (\tilde{y}_q - y_q)^2 + (\bar{y} - \bar{y})^2 + 2(y_q - \bar{y}) \cdot (\tilde{y}_q - y_q) - \right. \\
 &\quad \left. - 2(y_q - \bar{y}_q) \cdot (\bar{y} - \bar{y}) - 2(\tilde{y}_q - y_q) \cdot (\bar{y} - \bar{y}) \right\} = \\
 &= \frac{1}{N} \cdot (N \cdot \Delta^2 \langle y_q \rangle + N \cdot \sigma^2(\tilde{y}) + \sigma^2(\tilde{y}) + 0 - 0 - 2\sigma^2(\tilde{y})) = \\
 &= \Delta^2 \langle y_q \rangle + \frac{N-1}{N} \cdot \sigma^2(\tilde{y}). \tag{408}
 \end{aligned}$$

Параметр $k_{(x, \tilde{y})}^2 = v^2 \langle x_q, \tilde{y}_q \rangle / \Delta^2 \langle x_q \rangle \cdot s^{*2} \langle \tilde{y}_q \rangle$ при $\sigma^2(\tilde{y}) \rightarrow 0$ и линейной зависимости типа $y = a_0 + a \cdot x$, очевидно, стремится к 1; $\Delta^2 \langle y_q \rangle$ и $v^2 \langle x_q, y_q \rangle$ суть *плановые дисперсия* и *ковариация* для $u(x)$.

Вообще же, при линейной зависимости $y = a_0 + a \cdot x$ безразлично, какую сумму квадратов тут минимизировать — по оси y или по оси x . С учётом теоремы Пифагора, можно минимизировать сумму квадратов расстояний от точек до прямой линии регрессии. Результат одинаков! Отметим, что для любых переменных x и y (точных и неточных) имеет место *алгебраическое псевдокорреляционное неравенство* (409):

$$-1 \leq \frac{\overline{y \cdot x} - \bar{y} \cdot \bar{x}}{\left[(\overline{x^2} - \bar{x}^2) \cdot (\overline{y^2} - \bar{y}^2) \right]^{1/2}} \leq +1 \rightarrow \frac{(\overline{y \cdot x} - \bar{y} \cdot \bar{x})^2}{(\overline{x^2} - \bar{x}^2) \cdot (\overline{y^2} - \bar{y}^2)} \leq 1.$$

* * *

Далее рассмотрим генезис метода наименьших квадратов Гаусса для случая $n > 1$. Интересно, что *инвариантность метода к общим линейным, в том числе к масштабным преобразованиям координат и в данном случае имеет место, что подчёркивает его универсальность!*

Линейная (линеаризованная) целевая функция имеет общий вид:

$$y = y(x_1, x_2, \dots, x_n) = a_0 + a_1 \cdot x_1 + a_2 \cdot x_2 + \dots + a_n \cdot x_n.$$

С векторной формой переменной эта зависимость имеет вид:

$$y = y(\mathbf{x}) = a_0 + \mathbf{x}' \cdot \mathbf{a}. \quad (410)$$

Здесь традиционно $\mathbf{a}$ есть $n \times 1$ -вектор-столбец из коэффициентов a_i . Для реализации метода при $n > 1$ применяется *линейная регрессия* $\tilde{y}|\mathbf{x}$:

$$\dot{y} = \dot{y}(x_1, x_2, \dots, x_n) = \dot{a}_0 + \dot{a}_1 \cdot x_1 + \dot{a}_2 \cdot x_2 + \dots + \dot{a}_n \cdot x_n. \quad (411)$$

Представим модель регрессии при $n > 1$ в векторно-скалярной форме:

$$\dot{y} = \dot{y}(\mathbf{x}) = \dot{a}_0 + \mathbf{x}' \cdot \dot{\mathbf{a}}, \quad (412)$$

где $\mathbf{x}$ — $n \times 1$ -векторная переменная, $\dot{\mathbf{a}}$ — $n \times 1$ -векторный коэффициент, $\dot{a}_0$ — скалярный коэффициент. Для их вычисления также аналитически решается задача минимизации суммы квадратов разностей ($\dot{y}_q - \tilde{y}_q$):

$$\psi = \psi(\dot{\mathbf{a}}, \dot{a}_0) = \sum_{q=1}^N (\mathbf{x}_q' \cdot \dot{\mathbf{a}} + \dot{a}_0 - \tilde{y}_q)^2 = \min; \quad (413)$$

$$\left. \begin{aligned} \frac{1}{2} \cdot \frac{\partial \psi}{\partial \dot{a}_0} &= \sum_{q=1}^N (\mathbf{x}_q' \cdot \dot{\mathbf{a}} + \dot{a}_0 - \tilde{y}_q) = 0, \\ \frac{1}{2} \cdot \frac{\partial \psi}{\partial \dot{\mathbf{a}}} &= \sum_{q=1}^N \mathbf{x}_q \cdot (\mathbf{x}_q' \cdot \dot{\mathbf{a}} + \dot{a}_0 - \tilde{y}_q) = \mathbf{0}. \end{aligned} \right\} \quad (414)$$

Используя средние арифметические, из (414) получаем систему из 2-х усреднённых уравнений с двумя неизвестными $\dot{\mathbf{a}}$ и $\dot{a}_0$:

$$\overline{\tilde{y}} = \overline{\mathbf{x}'} \cdot \dot{\mathbf{a}} + \dot{a}_0, \quad (415)$$

$$\overline{\tilde{y} \cdot \mathbf{x}} = \overline{\mathbf{x}} \cdot \dot{a}_0 + \overline{\mathbf{x} \cdot \mathbf{x}'} \cdot \dot{\mathbf{a}}. \quad (416)$$

Выразив из первого уравнения $\dot{a}_0$ и подставив это значение во второе уравнение, получаем уравнение относительно $\dot{\mathbf{a}}$:

$$(\overline{\mathbf{x} \cdot \mathbf{x}'} - \overline{\mathbf{x}} \cdot \overline{\mathbf{x}'}) \cdot \dot{\mathbf{a}} = \overline{\tilde{y} \cdot \mathbf{x}} - \overline{\tilde{y}} \cdot \overline{\mathbf{x}}.$$

В развёрнутой форме записи оно тождественно уравнению:

$$\left(N \cdot \sum_{q=1}^N \mathbf{x}_q \mathbf{x}_q' - \sum_{q=1}^N \mathbf{x}_q \cdot \sum_{q=1}^N \mathbf{x}_q' \right) \cdot \dot{\mathbf{a}} = N \cdot \sum_{q=1}^N \tilde{y}_q \mathbf{x}_q - \sum_{q=1}^N \tilde{y}_q \cdot \sum_{q=1}^N \mathbf{x}_q.$$

В кратчайшей матрично-векторной форме записи имеем:

$$\mathbf{S} \cdot \dot{\mathbf{a}} = \mathbf{w}, \quad (417)$$

где $\mathbf{S} = \mathbf{S}'$, причём элементы матрицы $\mathbf{S}$ и вектора $\mathbf{w}$ выражаются как

$$\begin{aligned} s_{ij} &= s_{ji} = N^2 \cdot \left(\overline{x_i \cdot x_j} - \bar{x}_i \cdot \bar{x}_j \right) = \\ &= \frac{N^2}{N} \cdot \sum_{q=1}^N (x_{i,q} - \bar{x}_i) \cdot (x_{j,q} - \bar{x}_j) = N \cdot \sum_{q=1}^N x_{i,q} x_{j,q} - \sum_{q=1}^N x_{i,q} \cdot \sum_{q=1}^N x_{j,q}, \\ w_i &= N^2 \cdot \left(\overline{\tilde{y} \cdot x_i} - \bar{\tilde{y}} \cdot \bar{x}_i \right) = \\ &= N^2 \cdot \left[\frac{1}{N} \sum_{q=1}^N (x_{i,q} - \bar{x}_i) \cdot (\tilde{y}_q - \bar{\tilde{y}}) \right] = N \cdot \sum_{q=1}^N x_{i,q} \tilde{y}_q - \sum_{q=1}^N x_{i,q} \cdot \sum_{q=1}^N \tilde{y}_q. \end{aligned}$$

Далее имеем:

$$\left. \begin{aligned} \dot{\mathbf{a}} &= \mathbf{S}^{-1} \cdot \mathbf{w}, \\ \dot{a}_0 &= \bar{\tilde{y}} - \bar{\mathbf{x}}' \cdot \dot{\mathbf{a}}. \end{aligned} \right\} \quad (418)$$

Для однозначности $\dot{\mathbf{a}}$ и $\dot{a}_0$ необходимо, чтобы выполнялось $\det \mathbf{S} \neq 0$, что достигается выбором плана расположения точек $\mathbf{x}_q$ в $\mathcal{A}^n$. Причём при $\mathbf{x} \rightarrow \mathbf{z} = \mathbf{V}\mathbf{x}$ имеем инвариантность $\mathbf{a}_{(x)} \rightarrow \mathbf{a}_{(z)} = \mathbf{V} \mathbf{a}_{(x)}$, $\mathbf{a}_{0(z)} = \mathbf{a}_{0(x)}$.

Формулы (415)–(418) значительно упрощаются для специальных *центрированных* планов расположения точек $\mathbf{x}_q$:

$$\sum_{q=1}^N \mathbf{x}_q = \mathbf{0} = \bar{\mathbf{x}} \quad \rightarrow \quad \dot{a}_0 = \bar{\tilde{y}}, \quad \dot{\mathbf{a}} = \left\{ \overline{\mathbf{x} \cdot \mathbf{x}'} \right\}^{-1} \cdot \bar{\tilde{y} \cdot \mathbf{x}}. \quad (419)$$

Ещё важный вариант метода возникает тогда, когда в функции (412) заведомо известно, что $\dot{a}_0 = 0$. Он преобразуется в *централизованный метод наименьших квадратов*. Тогда уравнение (415) принимает вид:

$$\bar{\tilde{y}} = \bar{\mathbf{x}}' \cdot \dot{\mathbf{a}} \leftrightarrow \sum_{q=1}^N \tilde{y}_q = \sum_{q=1}^N \mathbf{x}'_q \cdot \dot{\mathbf{a}}.$$

Это уравнение имеет однозначное решение только при $n \leq 2$. Вообще при $n > 1$ искомое однозначное решение получается из (416) при $\dot{a}_0 = 0$, т. е. из **системы линейных нормальных уравнений Гаусса**:

$$\left\{ \sum_{q=1}^N \mathbf{x}_q \mathbf{x}'_q \right\} \cdot \dot{\mathbf{a}} = \sum_{q=1}^N \tilde{y}_q \cdot \mathbf{x}_q \rightarrow \dot{\mathbf{a}} = \left\{ \sum_{q=1}^N \mathbf{x}_q \mathbf{x}'_q \right\}^{-1} \cdot \left(\sum_{q=1}^N \tilde{y}_q \mathbf{x}_q \right). \quad (420)$$

Причём, решения (419) и (420) отличаются только коэффициентом $\dot{a}_0$. *Искусственным путём* этот вариант можно получить преобразованием: $\mathbf{y}(\mathbf{x}) \rightarrow \psi(\mathbf{x}) = \mathbf{y}(\mathbf{x}) - \dot{a}_0 = \mathbf{x}' \cdot \dot{\mathbf{a}} \sim \dot{\mathbf{y}}(\mathbf{x}) \rightarrow \dot{\psi}(\mathbf{x}) = \dot{\mathbf{y}}(\mathbf{x}) - \dot{\tilde{y}}(0) = \mathbf{x}' \cdot \dot{\mathbf{a}}$.

Глава 5. Численные методы оптимизации

§ 5.1. Общие положения

Численная оптимизация целевой функции $y(\mathbf{x})$ — это определённый пошаговый вычислительный процесс, сводящий решение задачи на её экстремум к некоторому алгоритму, продуцирующему в итоге либо последовательность значений аргумента $\mathbf{x}_{(1)}, \mathbf{x}_{(2)}, \dots, \mathbf{x}_{(k)}, \dots$, неуклонно приближающуюся с ростом k к точке экстремума $\mathbf{s}^*$ с заданной степенью точности δ по $|\Delta \mathbf{x}|$, либо последовательность значений целевой функции $y(\mathbf{x}_{(1)}), y(\mathbf{x}_{(2)}), \dots, y(\mathbf{x}_{(k)}), \dots$, неуклонно приближающуюся с ростом k к экстремуму $y(\mathbf{s}^*)$ с заданной степенью точности $\pm \varepsilon$. В первом случае говорят о сходимости процедуры по аргументу, во втором случае говорят о сходимости процедуры по функции. Очевидно, что для *эволюционной* целевой функции оба вида сходимости тождественны. Тогда говорят о сходимости процесса оптимизации вообще [4, 8, 21]. В простейших случаях при некотором k возможно совпадение с искомым результатом, т. е. $\mathbf{x}_{(k)} = \mathbf{s}^*$ и $y(\mathbf{x}_{(k)}) = y(\mathbf{s}^*)$ с остановкой процесса. (Это теоретически реализуется при совпадении порядков функции и процедуры.)

Наиболее распространённые и действенные алгоритмы такого рода суть *итерации*. В качестве примера итерации можно указать *предельный метод* вычисления экстремальных корней алгебраического уравнения, в том числе векового, с вещественным положительным спектром (§ 4.2). В наиболее общем виде итерация понимается как последовательное применение одной и той же вычислительной процедуры к числу или некоему набору чисел в их некотором детерминированном ряду, начиная с исходного числа или набора чисел, и далее к последующим числам, производимым в результате выполнения процедуры. Для обеспечения однозначной сходимости итерационных процедур оптимизации $y(\mathbf{x})$ к точке её экстремума $\mathbf{s}^*$ необходимо такое свойство целевой функции, как *строгая унимодальность* (см. § 1.1), хотя бы на рассматриваемой области значений, — т. е. наличие только одного и причём строгого экстремума, а также отсутствие стационарных седловин (перегибов).

Ранее в §§ 1.1 и 1.9 рассматриваемые в этой монографии целевые скалярные функции были наделены как бы априори существенным в данном аспекте свойством *эволюционности*. Строго математически для $y(\mathbf{x})$ и $y(\mathbf{x})$ оно обеспечивается её непрерывностью и непрерывной дифференцируемостью.

Но в ряде численных процедур, например, для поиска экстремума непрерывность скалярной функции на компакте (отрезке на оси $\mathcal{R}$ или закрытой области в $\mathcal{A}^n$) задаётся как её **непрерывность по Липшицу**. Функция $y(\mathbf{x})$ (или $y(\mathbf{x})$) определяется как непрерывная по Липшицу на некоей области $\mathcal{D} \subset \mathcal{E}^n$ с коэффициентом $L > 0$, если выполняется следующее требование (при $n \geq 1$):

$$\frac{|y(\mathbf{x}_{(2)}) - y(\mathbf{x}_{(1)})|}{\|\mathbf{x}_{(2)} - \mathbf{x}_{(1)}\|} \leq L : \forall \mathbf{x}_{(1)}, \mathbf{x}_{(2)} \in \mathcal{D}. \quad (421)$$

Здесь $\mathcal{D}$ — область определения целевой функции $y(\mathbf{x})$; в частности, $\mathcal{D} \equiv \mathcal{E}^n$ (см. §§ 1.1 и 1.9). Ясно, что данное понятие требует метризации координатного пространства $\mathcal{A}^n$ обычно в $\mathcal{E}^n$. Геометрический смысл числа L заключается в том, что он равен максимуму модуля градиента, или максимуму тангенса угла в $\mathcal{E}^{n+1}$ наклона касательных к $y(\mathbf{x})$. Если скалярные переменные x_i суть какие-либо физические величины с размерностью, или масштабом по осям m_i , то изменяя эти масштабы, коэффициенту L можно придать любое *конечное* значение. Поэтому тут более важен сам факт конечности L , характеризующий *равномерную непрерывность* целевой функции $y(\mathbf{x})$. Однако понятие непрерывности по Липшицу с коэффициентом $L > 0$ даёт возможность делать самые общие оценки скорости сходимости процедуры.

Основным параметром для любого метода численной оптимизации является его порядок. Он определяется тем, что именно применяется в расчётах очередного приближения точки экстремума целевой функции: значения функции, значения её 1-производной (градиента), значения её 2-й производной (матрицы Гессе). Теоретически оптимален порядок процедуры, равный уровню экстремума функции (гл. 1).

В следующих параграфах рассматриваются основные классические итерационные методы оптимизации целевых функций — в одномерном и многомерном вариантах. Для большей конкретности итерационные методы одномерной оптимизации наглядно рассматриваются в варианте поиска *максимума*. Соответственно для функции $[-y(\mathbf{x})]$ те же самые процедуры обеспечивают поиск *минимума* $y(\mathbf{x})$.

§ 5.2. Итерационная одномерная оптимизация

Одномерная оптимизация имеет как самостоятельное значение, так и входит в качестве вспомогательного процесса в *директивные* процедуры для многомерной оптимизации. Численная оптимизация целевой функции применяется тогда, когда аналитическое решение её уравнения стационарности весьма сложно, либо вовсе невозможно. Кроме того, на практике часто аналитический вид целевой функции $u(x)$ вообще неизвестен; её значения определяются экспериментально при точно задаваемых значениях аргумента. Причём есть достаточно веские основания полагать наличие экстремальной зависимости $u(x)$. Вообще же, если нахождение значений функции при конкретных значениях аргумента требует довольно высоких затрат, то имеется естественное стремление свести к минимуму количество точек x_i , где вычисляются или экспериментально находятся значения $u(x_i)$.

В методах 0-го порядка непосредственно применяют только значения функции $u(x_i)$. В методах 1-го порядка применяют как значения функции $u(x_i)$, так и её 1-й производной $u'(x_i)$, либо её некую опосредованную через значения функции *разностную оценку*. В методах 2-го порядка применяют как значения функции $u(x_i)$, так и её 1-й производной $u'(x_i)$, 2-й производной $u''(x_i)$, либо их опосредованные через значения функции *разностные оценки*.

Вначале (по методам 0-го порядка) находят двухточечный интервал скалярного аргумента $[x_{(1)}, x_{(2)}]$, в котором заведомо находится точка экстремума s^* предположительно строго унимодальной и непрерывной целевой функции $u(x)$. В самом общем виде эта процедура сводится к поиску экстремума функции при дискретном изменении аргумента тем или иным способом. На первом этапе, разумеется, нужно выбрать исходную точку a и далее одно из двух направлений от a , в котором $u(x)$ изменяется должным образом. После этого выбирается конкретная процедура поиска в данном направлении — либо увеличивающимися, либо уменьшающимися последовательными шагами.

По первому варианту поиска довольно популярен способ с удвоением размера шагов. Делают шаги Δx , $2\Delta x$, $4\Delta x$, ... от a пока не пройдут искомый экстремум целевой функции от дискретной переменной x в трёхточечном интервале $[x_{(1)}, x_{(c)}, x_{(2)}]$, где $x_{(1)} = x_{(k-1)}$, $x_{(c)} = x_k$, $x_{(2)} = x_{(k+1)}$. Из дальнейшего будет ясно, что здесь вместо удвоения шага вполне логично применять *повышающий* коэффициент золотого сечения $\mu = (\sqrt{5} + 1)/2 \approx 1,618$.

По второму варианту поиска довольно популярен способ, связанный с *уполовиниванием* размера шагов. Выбирают довольно большой шаг аргумента от точки a , делят его пополам, образуя некий трёхточечный интервал аргумента $[a, x_{(c)}, b]$. Если на данном трёхточечном интервале функция $y(x)$ необходимым образом экстремальная, то его принимают за искомый. Если же $y(x)$ на данном интервале не экстремальная, то образуют в ту же сторону новый интервал от точки b . Его опять-таки можно оставить прежним или поделить в зависимости от характера дискретного изменения функции. Из дальнейшего будет ясно, что здесь вместо уполовинивания шага вполне логично применять *понижающий* коэффициент золотого сечения $\lambda = (\sqrt{5} - 1)/2 \approx 0,618$. Причём $\mu \cdot \lambda = 1$.

По обоим указанным способам на каком-то шаге процедуры поиска находят трёхточечный интервал $[x_{(1)}, x_{(c)}, x_{(2)}]$, или по крайним точкам двухточечный интервал $[x_{(1)}, x_{(2)}]$, содержащий точку экстремума $y(x)$.

Здесь, разумеется, должна работать больше интуиция вычислителя или экспериментатора. Но в итоге, более или менее удачно, будет найден искомый интервал аргумента для заведомо локализованного поиска на нём экстремума целевой функции. Обозначим его для дальнейшего как $[a_1, b_1]$. Это, по определению, есть исходный *экстримный интервал* для аргумента. Он имеет особое значение в методах 0-го и 1-го порядка.

В численных процедурах оптимизации по умолчанию принимается то, что все шаговые значения аргумента — точные числа, а получаемые для них значения целевой функции находятся с задаваемой абсолютной погрешностью ε . (Она, кстати, определяет допускаемое округление этих значений — с учётом правила Гаусса.) Поэтому сравнение значений $y(a_k)$ и $y(b_k)$ осуществляют также с заданной погрешностью ε . Причём условие

$$|y(b_k) - y(a_k)| \leq \varepsilon \quad (422)$$

эквивалентно равенству обоих значений целевой функции $y(x)$.

Отсюда оптимизация $y(x)$ на экстримном интервале $[a_1, b_1]$ должна строиться алгоритмически так, чтобы в результате каждой новой итерации последующий интервал $[a_k, b_k]$ неуклонно сужался и был бы также экстримным, т. е. при этом он продолжал бы включать s^* как свою внутреннюю точку. Критерий остановки общего процесса оптимизации есть условие типа (422) — при достаточно близких значениях аргумента. Данный факт свидетельствует именно о *строгом* экстремуме функции в точке s^* .

§ 5.3. Методы дихотомии и золотого сечения 0-го порядка

В данном параграфе последовательно рассматриваются два наиболее распространённых итерационных метода 0-го порядка для решения задачи одномерной оптимизации на экстримном интервале $[a_1, b_1]$:

$$y(x) \rightarrow \text{extr}, x \in [a_1, b_1]. \quad (423)$$

Причём $y(x)$ здесь строго унимодальная целевая функция, непрерывная по Липшицу, согласно (421).

Лемма 3 (об иерархии на экстримном интервале). Пусть $y(x)$ — строго унимодальная функция, s^+ — её максимум на интервале $[a_1, b_1]$. Тогда для любых точек $a_k, b_k \in [a_1, b_1]$ справедливы утверждения:

- 1) если $y(a_k) < y(b_k)$, то $s^+ \in [a_k, b_1]$;
- 2) если $y(a_k) > y(b_k)$, то $s^+ \in [a_1, b_k]$;
- 3) если $y(a_k) = y(b_k)$, то $s^+ \in [a_k, b_k] \begin{matrix} \subset [a_k, b_1] \\ \subset [a_1, b_k] \end{matrix}$.

Аналогичная по смыслу лемма действует на экстримном интервале, содержащем точку минимума s^- . Но в ней нужно только поменять знаки неравенств, что эквивалентно замене $y(x)$ на $[-y(x)]$. Смысл леммы состоит в том, что, сравнивая значения $y(x)$ в двух внутренних точках какого-либо экстримного интервала, можно далее перейти к новому экстримному интервалу, но уже с меньшей длиной и содержащемуся внутри предыдущего интервала. Что ещё важно, новый интервал — либо $[a_k, b_1]$, либо $[a_1, b_k]$ — будет содержать три точки с известными значениями в них целевой функции $y(x)$. Поэтому для последующего повторения подобной процедуры в рамках последовательных итераций потребуется нахождение значения функции всего лишь в одной точке нового интервала!

На этой главной исходной идее базируются излагаемые ниже два классических итерационных метода одномерной оптимизации строго унимодальной целевой функции на некотором исходном экстримном интервале $[a_1, b_1]$ — рис. 10. Это [например, 8, 19] *метод дихотомии* и *метод золотого сечения*. (Как ранее указывалось, для большей конкретности подобные методы рассматриваются в варианте поиска *максимума* целевой функции.)

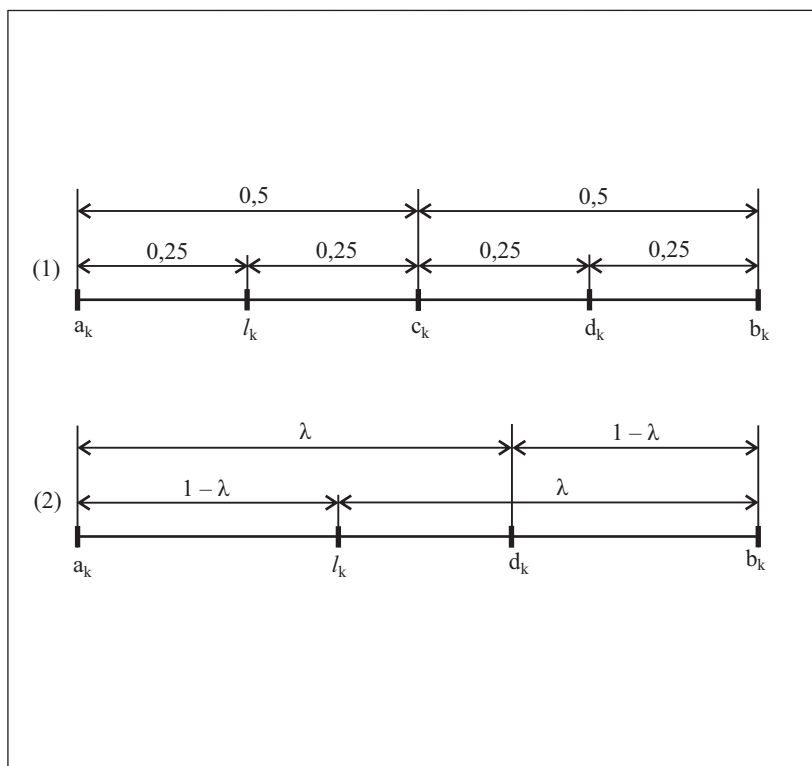


Рис. 10. Схемы деления экстримных интервалов в итерационных методах 0-го порядка:

(1) — метод дихотомии,

(2) — метод золотого сечения ($\lambda = \frac{\sqrt{5}-1}{2} \approx 0,618$).

Метод дихотомии

Название данного метода объясняется тем, что в нём каждая новая точка в экстримном интервале получается путём деления специально выбираемого отрезка пополам (см. рис. 10). Ниже излагается алгоритм, реализующий метод дихотомии, с учётом сравнения значений целевой функции в точках экстримного интервала в начале каждого шага.

Алгоритм 1. Исходные данные: $a_1, b_1, c_1 = 1/2 \cdot (a_1 + b_1); y(a_1), y(b_1), y(c_1)$.
 Полагаем $k = 1$ и переходим к п. 1.

1. Если $y(a_k) \geq y(b_k)$, то $l_k = 1/2 \cdot (a_k + c_k)$, находим $y(l_k)$ и переходим к п. 2.
 Если $y(a_k) < y(b_k)$, то $d_k = 1/2 \cdot (b_k + c_k)$, находим $y(d_k)$ и переходим к п. 3.

2. Если $y(l_k) \geq y(c_k)$, то полагаем $a_{k+1} = a_k, b_{k+1} = c_k, c_{k+1} = l_k$.
 Увеличиваем номер шага на 1 и переходим к п. 1. Если $y(l_k) < y(c_k)$, то $d_k = 1/2 \cdot (c_k + d_k)$, находим $y(d_k)$. Если $y(d_k) \leq y(c_k)$, то полагаем $a_{k+1} = l_k, b_{k+1} = d_k, c_{k+1} = c_k$. Увеличиваем номер шага на 1 и переходим к п. 1. Если $y(d_k) > y(c_k)$, то полагаем $a_{k+1} = c_k, b_{k+1} = b_k, c_{k+1} = d_k$. Увеличиваем номер шага на 1 и переходим к п. 1.

3. Если $y(d_k) \geq y(c_k)$, то полагаем $a_{k+1} = c_k, b_{k+1} = b_k, c_{k+1} = d_k$.
 Увеличиваем номер шага на 1 и переходим к п. 1. Если $y(d_k) < y(c_k)$, то $l_k = 1/2 \cdot (a_k + c_k)$, находим $y(l_k)$. Если $y(l_k) \leq y(c_k)$, то полагаем $a_{k+1} = l_k, b_{k+1} = d_k, c_{k+1} = c_k$. Если $y(l_k) > y(c_k)$, то полагаем $a_{k+1} = a_k, b_{k+1} = c_k, c_{k+1} = l_k$. Увеличиваем номер шага на 1 и переходим к п. 1.

Процедура этого алгоритма обрывается на каком-то его шаге при выполнении требования оценочного неравенства (422).

Согласно лемме 3, после k -го шага имеем:

$$s^* \in [a_{k+1}, b_{k+1}] \subset [a_k, b_k] \subset \dots \subset [a_1, b_1].$$

В результате деления экстримного интервала на каждом шаге пополам после k -го шага имеем:

$$b_{k+1} - a_{k+1} = (b_1 - a_1)/2^k. \quad (424)$$

Кроме того, из данного алгоритма видно, что на каждом шаге поиска равновероятно делается либо одно (левое для l_k или правое для d_k), либо два (левое и правое для l_k и d_k) вычисления целевой функции, т. е. в среднем по 1,5 раза. Для k -го шага количество вычислений (экспериментов) оценивается здесь как $N = 1,5k$. Тогда, с учётом (424), погрешность изложенного метода дихотомии с ростом N оценивается как

$$\Delta_k = (b_1 - a_1)/2^k \approx (b_1 - a_1)/2^{2N/3} \approx 0,63^N \cdot (b_1 - a_1). \quad (425)$$

Метод золотого сечения

Название данного метода объясняется тем, что деление экстримных интервалов выполняется по правилу золотого сечения в отношениях $\lambda : (1 - \lambda)$ и $(1 - \lambda) : \lambda$, где $\lambda = (\sqrt{5} - 1) / 2 \approx 0,618$ — положительное решение уравнения $1/x = x / (1 - x)$, см. рис. 10. Термин *золотое сечение* для применения к пропорциям в архитектуре ввёл Леонардо да Винчи. Теоретически метод при достаточно большом k (количестве шагов) требует минимум вычислений значений $y(x)$ при заданной погрешности. Однако при конечном числе k теоретически более оптимален только **метод чисел Фибоначчи**. (Но при $k \rightarrow \infty$ он попросту формально тождествен методу золотого сечения — см. например [8, 21].)

Ниже излагается алгоритм метода золотого сечения с левыми (l_k) и правыми (d_k) золотыми точками.

Алгоритм 2. Исходные данные: $a_1, b_1, \Delta_1 = b_1 - a_1, l_1 = b_1 - \lambda \cdot \Delta_1, d_1 = a_1 + \lambda \cdot \Delta_1; y(a_1), y(b_1), y(l_1), y(d_1)$. Полагаем $k = 1$ и переходим к п. 1.

1. Если $y(l_k) \geq y(d_k)$, то $\Delta_{k+1} = d_k - a_k, a_{k+1} = a_k, b_{k+1} = d_k, l_{k+1} = l_k, d_{k+1} = a_k + \lambda \cdot \Delta_{k+1}$. Увеличиваем номер шага на 1 и переходим к п. 1. Если $y(l_k) < y(d_k)$, то $\Delta_{k+1} = b_k - l_k, a_{k+1} = l_k, b_{k+1} = b_k, l_{k+1} = b_k - \lambda \cdot \Delta_{k+1}, d_{k+1} = d_k$. Увеличиваем номер шага на 1 и переходим к п. 1.

Процедура этого алгоритма обрывается на каком-то его шаге при выполнении требования оценочного неравенства (422).

Согласно лемме 3, после k -го шага имеем:

$$s^* \in [a_{k+1}, b_{k+1}] \subset [a_k, b_k] \subset \dots \subset [a_1, b_1].$$

В результате деления экстримного интервала после k -го шага имеем:

$$b_{k+1} - a_{k+1} = (b_1 - a_1) \cdot \lambda^k. \quad (426)$$

Из сравнения (424) и (426) видно, что $\lambda > 1/2$ и, казалось бы, скорость движения к экстремуму тут ниже. Но из алгоритма видно, что для каждого шага делается только одно вычисление, т. е. для k -го шага количество вычислений (экспериментов) оценивается как $N = k$. С учётом (426), погрешность метода золотого сечения с ростом N оценивается как

$$\Delta_k = (b_1 - a_1) \cdot \lambda^N \approx 0,618^N \cdot (b_1 - a_1). \quad (427)$$

Причём сравнение (425) и (427) показывает, что это вообще более эффективная процедура в сравнении с методом дихотомии, хотя и более сложная.

Из алгоритма метода также видно, что на каждом шаге соотношение предшествующего большего интервала Δ_{k-1} и последующего меньшего интервала Δ_k всегда одно и то же:

$$\frac{1}{\lambda} = \frac{\lambda}{1-\lambda} = \frac{1-\lambda}{2\lambda-1} = \frac{2\lambda-1}{2-3\lambda} = \dots = \frac{F_{k-1} - F_k \cdot \lambda}{F_{k+1} \cdot \lambda - F_k}. \quad (428)$$

Откуда следует, что

$$\lambda^2 + \frac{F_{k+1} - F_{k-1}}{F_k} \cdot \lambda - 1 = 0. \text{ (По формуле Бине: } F_k = [\lambda^k - (-\lambda)^{-k}] / \sqrt{5}.)$$

Поскольку в данном методе $\lambda^2 + \lambda - 1 = 0$, то $F_{k+1} = F_{k-1} + F_k$, $k = 2, 3, \dots$. Отсюда F_k — числа Фибоначчи. Они по известной легенде были открыты этим знаменитым математиком в результате его научного наблюдения за процессом увеличения популяции от пары кроликов в 1202 г. (!) Согласно соотношениям типа (428), числа Фибоначчи с увеличением числа k стремятся к членам геометрической прогрессии с основанием $1/\lambda = 1 + \lambda = \mu \approx 1,618$:

$$\left. \begin{aligned} k = 2m &\rightarrow F_k / F_{k-1} > 1/\lambda, \\ k = 2m+1 &\rightarrow F_k / F_{k-1} < 1/\lambda, \\ k \rightarrow \infty &\rightarrow F_k / F_{k-1} = 1/\lambda. \end{aligned} \right\} \quad (429)$$

§ 5.4. Метод парных касательных 1-го порядка

Название метода объясняется тем, что каждая следующая точка в экстримном интервале — рис. 11 получается в результате пересечения пары касательных (левой и правой) с противоположным по знаку наклоном к оси абсцисс. (Позднее в § 5.5 с целью численного решения скалярного уравнения стационарности будет рассмотрен классический *метод касательных* Ньютона и его разностный аналог — метод хорд.)

Реализация методов 1-го порядка на том же самом экстримном интервале, однако, требует для целевой функции $y(x)$, помимо строгой унимодальности и равномерной непрерывности, ещё и того, чтобы знак её 2-й производной строго отвечал характеру искомого экстремума.

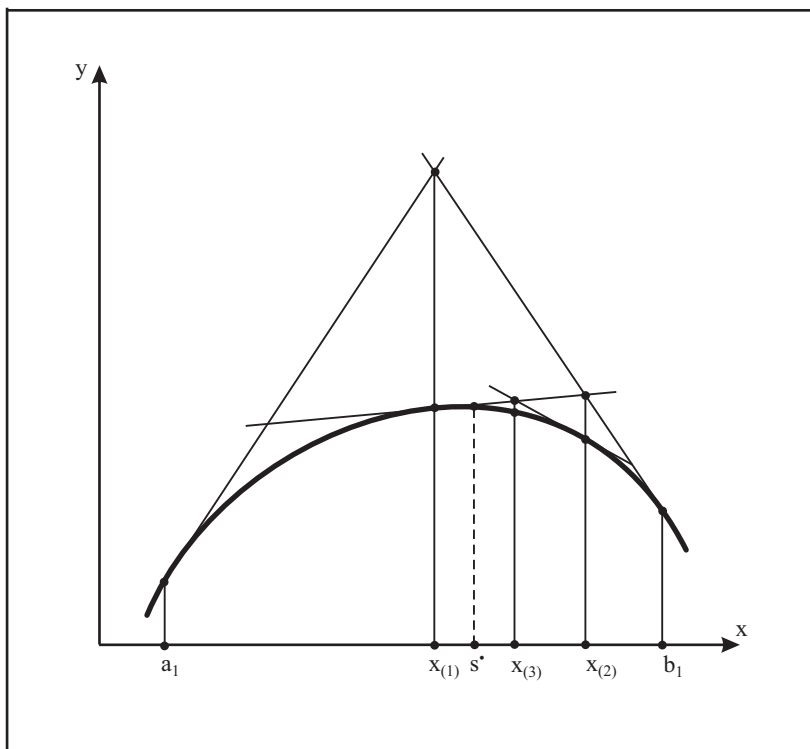


Рис. 11. Схема метода парных касательных.

А именно, при поиске максимума необходимое требование есть

$$y''(x) < 0: x \in [a_1, b_1],$$

при поиске минимума необходимое требование есть

$$y''(x) > 0: x \in [a_1, b_1].$$

Они более строгие, нежели те, что в части знака $y''(x)$ задают классические правила (2) и (3) — гл. 1. В численных процедурах их эквивалентом являются понятия *выпуклости* и *вогнутости* функции. Функция $y(x \in [a, b] \subset \mathcal{D})$ называется выпуклой на числовом отрезке $[a, b] \subset \mathcal{D}$, если при $0 \leq \alpha \leq 1$ выполняется неравенство:

$$y[\alpha \cdot x_{(i)} + (1 - \alpha) \cdot x_{(j)}] \leq \alpha \cdot y(x_{(i)}) + (1 - \alpha) \cdot y(x_{(j)}): \forall x_{(i)}, x_{(j)} \in [a, b]. \quad (430)$$

Аналогичным образом, но с изменением знака данного неравенства на противоположный определяется вогнутость функции $y(x)$.

Методы 1-го порядка заслуживают особого внимания тогда, когда 1-я производная целевой функции вычисляется достаточно просто. В частности, в рассматриваемом методе парных касательных (рис. 11) дополнительно к значениям функции в левой и правой точках в них же вычисляют значения 1-й производной $y'(x)$. После этого получают линейные функции $\varphi(x)$ для касательных — левой и правой. В варианте поиска максимума имеем:

$$[\varphi_l(x)]_k = y(a_k) + y'(a_k) \cdot (x - a_k), \quad (431)$$

$$[\varphi_d(x)]_k = y(b_k) - y'(b_k) \cdot (b_k - a) = y(b_k) + y'(b_k) \cdot (x - b_k), \quad (432)$$

где $y'(a_k) > 0$, $y'(b_k) < 0$. Точка пересечения указанной пары касательных, т. е. точечное решение уравнения $[\varphi_l(x)]_k = [\varphi_d(x)]_k$, даёт очередное $(k+1)$ -е приближение к точке максимума s^+ :

$$x_{k+1} = \frac{[y'(a_k) \cdot a_k - y(a_k)] - [y'(b_k) \cdot b_k - y(b_k)]}{y'(a_k) - y'(b_k)}. \quad (433)$$

Далее вычисляют $y(x_{k+1})$ и сравнивают со значениями $y(a_k)$ и $y(b_k)$.

Если разность с любым из них укладывается в неравенство (422), то процесс поиска максимума целевой функции прекращают в точке $x_{k+1} = s^+$ с заданной точностью ε (по функции). В s^+ с соответствующей ε точностью выполняется лемма Ферма о необходимом условии экстремума $y'(x_{k+1}) = 0$.

В противном случае вычисляют $y'(x_{k+1})$ и выбирают следующую пару точек a_{k+1} и b_{k+1} , т. е. левую и правую. Если $y'(x_{k+1}) > 0$, то приравнивают $x_{k+1} = a_{k+1}$, $b_k = b_{k+1}$; если $y'(x_{k+1}) < 0$, то приравнивают $a_k = a_{k+1}$, $x_{k+1} = b_{k+1}$. Далее повторяют ту же процедуру при $k + 1$. В конце процедуры числитель и знаменатель в (433) стремятся к нулю, как и обе производные y' . Поэтому особую важность имеет точность выполняемых вычислений.

В данном методе весьма существенно то, что на каждом шаге левая и правая точки частного интервала выбираются так, чтобы касательные в них имели соответственно положительный и отрицательный наклоны. Такая возможность обусловлена здесь именно выпуклостью целевой функции на всём интервале. Отсюда всегда $x_{k+1} \in [a_k, b_k]$. Но поскольку далее или $\{a_{k+1} = x_{k+1}, b_{k+1} = b_k\}$, или $\{a_{k+1} = a_k, b_{k+1} = x_{k+1}\}$, то и всегда $x_{k+1} \in [a_{k+1}, b_{k+1}]$, $[a_{k+1}, b_{k+1}] \subset [a_k, b_k]$.

Следовательно, имеем цепь вложенных экстримных интервалов, сужающихся после каждого очередного шага алгоритма:

$$[a_1, b_1] \supset [a_2, b_2] \supset \dots \supset [a_k, b_k] \supset [a_{k+1}, b_{k+1}].$$

Согласно лемме 3, после k -го шага $s^+ \in [a_{k+1}, b_{k+1}]$. Осталось оценить, насколько сужаются интервалы после каждого шага. Из (433) имеем в 2-х вариантах выбора нового интервала степень сужения:

$$\left. \begin{aligned} \frac{b_k - x_{k+1}}{b_k - a_k} &= \frac{y'(a_k) - \frac{y(b_k) - y(a_k)}{b_k - a_k}}{y'(a_k) + [-y'(b_k)]} = \frac{y'(a_k) + [-y'(c_k)]}{y'(a_k) + [-y'(b_k)]}, \\ \frac{x_{k+1} - a_k}{b_k - a_k} &= \frac{\frac{y(b_k) - y(a_k)}{b_k - a_k} + [-y'(b_k)]}{y'(a_k) + [-y'(b_k)]} = \frac{y'(c_k) + [-y'(b_k)]}{y'(a_k) + [-y'(b_k)]}. \end{aligned} \right\} \quad (434)$$

Нетрудно проверить, что для кривой $y(x)$ строго 2-го порядка, т. е. для параболы степени 2, обе величины равны точно 1/2. Поэтому для целевой функции, близкой именно ко 2-му порядку на экстримном интервале, метод касательных даёт сходимость со скоростью не менее скорости сходимости геометрической прогрессии со знаменателем 1/2, т. е. аналогично методу дихотомии (§ 5.3). Но, если кривая $y(x)$ более сильно выпуклая на $[a_1, b_1]$ (например, уровень экстремума менее 2-х), то эта скорость возрастает.

§ 5.5. Метод Ньютона 2-го порядка и его разностные аналоги

Для оптимизации таких целевых функций, которые в окрестности экстремума допускают параболическую аппроксимацию, целесообразно использовать метод 2-го порядка. Для сходимости процедуры нужно, чтобы функция в этой окрестности была равномерно непрерывной и чтобы начальное приближение аргумента c_1 находилось именно в зоне выпуклости (вогнутости) её производной. Пусть функция $y(x)$ в некоей окрестности точки экстремума s^* , по крайней мере, трижды непрерывно дифференцируемая. При этом, ограничиваясь частью 2-го порядка её разложения по формуле Тейлора в точке c , имеем аппроксимацию:

$$y(x) \approx q(x) = y(c) + y'(c) \cdot (x - c) + 1/2 \cdot y''(c) \cdot (x - c)^2. \quad (435)$$

Функция $q(x)$ отличается от целевой функции $y(x)$ в окрестности точки c только на остаточный член ряда Тейлора, например, в форме Лагранжа (11) порядка не менее 3-х. Парабола (435) в точке c имеет касательную, выражаемую производной:

$$q'(x) = y'(c) + y''(c) \cdot (x - c). \quad (436)$$

Согласно методу Ньютона 2-го порядка, точка экстремума параболы (435) при подстановке в неё $c = c_k$, а именно точка $x = c_{k+1}$, есть очередное приближение к точке экстремума s^* целевой функции $y(x)$. Формально по лемме Ферма c_{k+1} находится из уравнения стационарности $q'(x) = 0$. Геометрически она трактуется по методу Ньютона (рис. 12), как точка пересечения касательной (436) с осью абсцисс x [63]:

$$c_{k+1} = c_k + y'(c_k) / [-y''(c_k)], \quad y''(c_k) < 0. \quad (437)$$

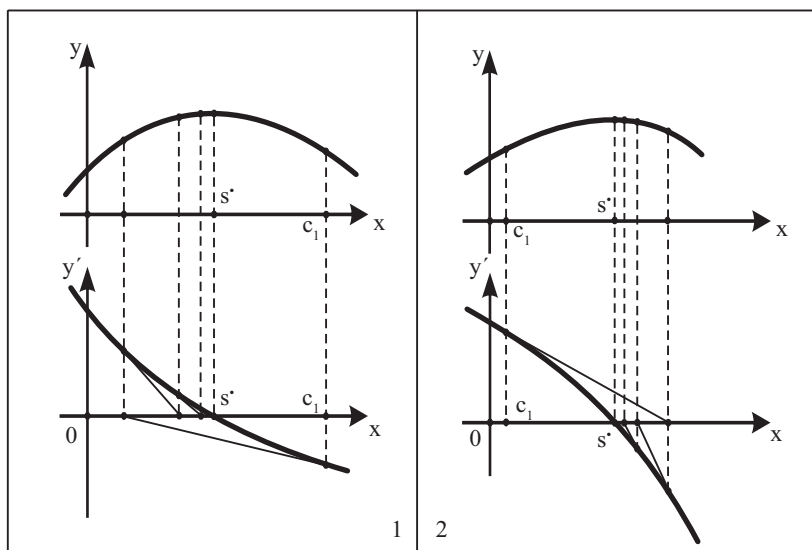


Рис. 12. Схемы пары эквивалентных численных процедур — метода 2-го порядка и метода касательных Ньютона (для решения уравнения стационарности) при поиске максимума 2-го уровня в двух основанных вариантах:

$$(1) y''(x) < 0, y'''(x) > 0; \quad (2) y''(x) < 0, y'''(x) < 0.$$

Отсюда следует двоякая интерпретация классического метода Ньютона. Наиболее общо он трактуется как метод 1-го порядка и тогда именуется как **метод касательных Ньютона**. В этой форме исходно он был предназначен для решения (скалярного) уравнения типа $f(x) = 0$ [63]. При этом уравнение стационарности $y'(x) = 0$ есть только его частный случай. Но в численной оптимизации этот же метод, по отношению к исходной целевой функции $y(x)$, является методом оптимизации Ньютона 2-го порядка и, вместе с тем, методом касательных при решении уравнения её стационарности! (Отсюда имеются опять-таки две возможности его *разностной модификации* — полная и частичная.)

По сути, итерация (437) есть некоторое гладкое *итерационное отображение*

$$x_{(k+1)} = \varphi(x_{(k)}). \quad (438)$$

Для сходимости в предельную точку, согласно **теореме Банаха**, оно должно быть именно *сужающим* [20]. В этом случае всегда имеется единственная точка s^* , к которой сходится последовательность точек $x_{(1)}, x_{(2)}, \dots$. Здесь это точка экстремума функции. Аналитически она является решением уравнения

$$x = \varphi(x). \quad (439)$$

В методе Ньютона *степень сжатия* отображения (438) на каждом шаге итерации оценивается сверху через коэффициент L_φ в требовании непрерывности по Липшицу к функции $\varphi(x)$. Понятно, что исходное приближение $x_{(1)} = c_{(1)}$ должно находиться настолько близко к s^* , чтобы выполнялось требование сужения отображения: $L_\varphi < 1$. Однако, для того чтобы скорость сходимости метода Ньютона превышала таковую для методов 0-го и 1-го порядка (см. §§ 5.2, 5.3), целесообразны значения $L_\varphi < 1/2$. На каждом шаге итерации сужение интервала в методе Ньютона оценивается сверху как

$$\left| s^* - c_{k+1} \right| \leq \frac{L_\varphi^k}{1 - L_\varphi} \cdot |c_2 - c_1|. \quad (440)$$

При этом следует иметь в виду, что и коэффициент L_φ на каждом шаге также уменьшается и часто квадратично. В силу этого факта, скорость сходимости метода Ньютона при определённых требованиях к значениям $y''(c_1)$ и $y'''(x)$ в окрестности экстремума имеет порядок 2, а не 1 [20, 21].

Ввиду того, что метод Ньютона базируется на вычислении 1-й и 2-й производной в точках c_k , он допускает две разностные модификации.

По *методу хорд* касательная (436) в точке c_k аппроксимируется хордой с небольшим шагом Δ_k по оси x . При приближении к экстремуму шаг Δ_k должен сужаться, например, пропорционально модулю $|c_k - c_{k-1}|$. По сути, этот метод применяется для численного решения уравнения $y'(x) = 0$, используя разностную аппроксимацию только для $y''(c_k)$, но через $y'(x)$. Тогда имеем *полуразностную* форму для итерации (437):

$$c_{k+1} = c_k + y'(c_k) / \left[-\frac{\Delta y'}{\Delta x}(c_k) \right], \quad (441)$$

$$\text{где } \frac{\Delta y'}{\Delta x}(c_k) = \frac{y'(c_k + \Delta_k) - y'(c_k)}{\Delta_k} \approx y''(c_k).$$

По *методу парабол* в (437) используют разностные аппроксимации и для $y'(c_k)$, и для $y''(c_k)$ через $y(x)$ с теми же шагами по оси x . Имеем:

$$c_{k+1} = c_k + \frac{\Delta y}{\Delta x}(c_k) / \left[-\frac{\Delta^2 y}{\Delta x^2}(c_k) \right], \quad (442)$$

где

$$\frac{\Delta y}{\Delta x}(c_k) = \frac{y(c_k + \Delta_k) - y(c_k - \Delta_k)}{2 \cdot \Delta_k}, \quad \frac{\Delta^2 y}{\Delta x^2}(c_k) = \frac{y(c_k + \Delta_k) + y(c_k - \Delta_k) - 2 \cdot y(c_k)}{\Delta_k^2}.$$

В итоге (442) приводится к форме:

$$c_{k+1} = c_k - \frac{\Delta_k}{2} \cdot \frac{y(c_k + \Delta_k) - y(c_k - \Delta_k)}{y(c_k + \Delta_k) + y(c_k - \Delta_k) - 2 \cdot y(c_k)}. \quad (443)$$

Ту же самую формулу получаем, если $y(x)$ в точке c аппроксимируется параболой, используя интерполяционный многочлен Лагранжа степени 2:

$$q(x) = y(c) \cdot \frac{x - (c - \Delta)}{c - (c - \Delta)} \cdot \frac{x - (c + \Delta)}{c - (c + \Delta)} + \\ + y(c - \Delta) \cdot \frac{x - c}{(c - \Delta) - c} \cdot \frac{x - (c + \Delta)}{(c - \Delta) - (c + \Delta)} + y(c + \Delta) \cdot \frac{x - c}{(c + \Delta) - c} \cdot \frac{x - (c - \Delta)}{(c + \Delta) - (c - \Delta)}. \quad (444)$$

Разностные аналоги метода Ньютона эффективно применять тогда, когда затруднено или даже невозможно находить производные целевой функции. Иногда метод Ньютона для обеспечения быстрой сходимости применяют на второй стадии оптимизации после методов 0-го или 1-го порядка.

§ 5.6. Итерационная многомерная оптимизация

В координатном пространстве, задаваемом некими независимыми переменными x_i , весьма существенными являются такие их свойства, как однородность или неоднородность в качестве *физических* величин. Обычно для скалярных функций типа $y = (x_1, x_2, \dots, x_n) = y(\mathbf{x})$ все её частные переменные x_i молчаливо подразумеваются однородными величинами (§ 1.9). Однако на практике эти частные переменные x_i имеют конкретные физические размерности — одинаковые или нет. Именно этот факт отвечает тут их однородности или неоднородности. Поэтому в процедурах многомерной оптимизации целевых функций такие дополнительные свойства переменных x_i обязательно нужно принимать в расчёт. Например, совершенно естественным является отображение физических пространственных координат x_1, x_2 и x_3 в евклидовом пространстве $\mathcal{E}^3$ и пространственно-временных координат x_1, x_2, x_3 и ct в псевдоевклидовом пространстве $\mathcal{P}^{3+1}$. Однако такие физически неоднородные переменные x_i , как давление, температура, концентрация, фигурирующие в целевых функциях, могут корректно отображаться лишь в неких физических аффинных пространствах $\mathcal{A}^n$ с допустимыми в них *масштабными* преобразованиями координат.

С учётом этого, важно изначально в целевых функциях применять именно такие естественные коэффициенты m_i к переменным x_i или/и именно такие сочетания переменных, чтобы в итоге их размерность или взаимно уничтожалась, или была хотя бы однородной. Переменные в чисто математических функциях обладают свойством безразмерности. Однако в прикладных разделах, а также в описательных исследованиях зачастую исходно фигурируют переменные разнообразной размерности. Поэтому, например, в физико-химических теориях подобия закономерно вводятся новые критериальные переменные. Причём они формируются из исходных переменных так, чтобы все их размерности уничтожались вследствие компенсации. Далее они применяются в целевых функциях.

Как один из самых ярких примеров в истории фундаментальных наук можно тут привести ввод коэффициента «с» к стреле времени t , впервые осуществлённый Пуанкаре в 1904 — 1905 гг. для придания физической однородности 4-м переменным, задающим базовое пространство-время в релятивистской теории движения материи. На основе этого Пуанкаре совершенно естественным *геометрическим* путём пришёл к 2-м важным следствиям, что коэффициент «с» есть скорость света в пустоте и что эта скорость максимальная для всех материальных явлений!

Весьма важно здесь то, что при линейных (в том числе масштабных) преобразованиях координат такие существенные тут свойства целевой функции (если они исходно имеются), как эволюционность или хотя бы равномерная непрерывность как 1-го, так и 2-го порядка, а также знакоопределённость квадратичной формы не нарушаются! Отсюда в численном анализе для $y(\mathbf{x})$ инвариантами линейных преобразований являются их важные свойства — аналоги: непрерывность по Липшицу и выпуклость (вогнутость). Но связанные с этими свойствами конкретные числовые коэффициенты, характеризующие максимальный наклон или минимальную кривизну поверхности $y(\mathbf{x})$, могут изменяться в любых конечных пределах, сохраняя лишь свои знаки. Конкретные числовые значения этих коэффициентов, однако, влияют на скорость сходимости.

Движение к экстремуму $y(\mathbf{x})$ в n -мерном координатном пространстве может осуществляться в форме итераций (§ 5.1). Причём общий процесс содержит либо только одностипные — многомерные *главные* итерации, либо как те же главные итерации, так и одномерные *побочные* итерации по *директивному вектору* (которые уже рассматривались в § 5.2). Второй из этих вариантов по отношению к первому, по сути, есть его *директивная* модификация. Возможные директивные процедуры далее, в общем, те же самые, что используют для одномерной оптимизации, т. е. либо пошаговые, либо поинтервальные.

§ 5.7. Покоординатные методы 0-го порядка

Простейшие процедуры для многомерной оптимизации целевой функции $y(\mathbf{x})$ основаны на циклическом *покоординатном* изменении аргумента $\mathbf{x}$ обычно в естественном порядке координат: $x_1, x_2, \dots, x_n$, — начиная с некоторого его начального значения $\mathbf{x} = \mathbf{c}$. Применяются две основные разновидности процедуры покоординатной оптимизации — с постоянным и с переменным (по ортам $\mathbf{e}_i$) шагом в пределах данного k -го цикла изменения n координат. Рассмотрим эти процедуры для определённости на примере поиска максимума целевой функции $y(\mathbf{x})$. (При поиске минимума попросту знак $y(\mathbf{x})$ в нижеуказанных процедурах изменяют на противоположный.)

Простой циклический процесс с постоянным k -м шагом в пределах k -го цикла сопровождается поочерёдным альтернативным изменением частных координат x_i с шагами $\pm \alpha_k \cdot \mathbf{e}_i$, где $\mathbf{e}_i$ — единичные орты. Конкретно, при очередной $(k + i + 1)$ -й итерации k -го цикла имеем далее нижеуказанный алгоритм.

- а. Если $y(\mathbf{c}_{k+i} + \alpha_k \cdot \mathbf{e}_{i+1}) > y(\mathbf{c}_{k+i})$, то принимают 1-ю альтернативу:
 $\mathbf{c}_{k+i+1} = \mathbf{c}_{k+i} + \alpha_k \cdot \mathbf{e}_{i+1}$ и переходят к $(k + i + 2)$ -й итерации.
- б. Если $y(\mathbf{c}_{k+i} + \alpha_k \cdot \mathbf{e}_{i+1}) \leq y(\mathbf{c}_{k+i})$, то переходят к 2-й альтернативе:
 $\mathbf{c}_{k+i+1} = \mathbf{c}_{k+i} - \alpha_k \cdot \mathbf{e}_{i+1}$.
- с. Если $y(\mathbf{c}_{k+i} - \alpha_k \cdot \mathbf{e}_{i+1}) > y(\mathbf{c}_{k+i})$, то принимают 2-ю альтернативу:
 $\mathbf{c}_{k+i+1} = \mathbf{c}_{k+i} - \alpha_k \cdot \mathbf{e}_{i+1}$ и переходят к $(k + i + 2)$ -й итерации.
- д. Если $y(\mathbf{c}_{k+i} - \alpha_k \cdot \mathbf{e}_{i+1}) \leq y(\mathbf{c}_{k+i})$, то полагают $\mathbf{c}_{k+i+1} = \mathbf{c}_{k+i}$ и переходят к $(k + i + 2)$ -й итерации k -го цикла.

Если неравенства (b) и (d) здесь имеют место для всех $1 \leq i \leq n$, то уменьшают, например, вдвое величину α_k и соответственно $(k + 1)$ -го шага. В противном случае α_k и соответственно шаг по ортам оставляют прежним. Далее переходят к $(k + 1)$ -му циклу по той же схеме. Процедура в целом заканчивается тогда, когда выполняется требование:

$$y(\mathbf{c}_{2k}) - y(\mathbf{c}_k) \leq \varepsilon, \quad (445)$$

где ε — заданная погрешность для целевой функции (§ 5.2).

Составной циклический процесс с переменным i -м шагом в k -ом цикле включает в себя поэтапную одномерную оптимизацию целевой функции в направлении возможных изменений каждой из частных координат x_i :

$$f(\alpha_{i+1}) = y(\mathbf{c}_{k+i} + \alpha_{i+1} \cdot \mathbf{e}_{i+1}) = \max: \alpha_{i+1} \in (-\infty, +\infty). \quad (446)$$

Одномерная оптимизация может осуществляться, например, любым из способов, описанных в §§ 5.3 — 5.5. В принципе, иногда возможны и даже более целесообразны совершенно иные подходы, например, аналитический, интерполяционный, приближение степенным рядом Тейлора.

Интересно также отметить, что *аналитическое* применение данной процедуры к задаче минимизации функции *невязки* $\|\delta(\mathbf{x})\|$ линейного уравнения (333) при $\det A \neq 0$ даёт классический метод Зейделя [8, 40] для численного решения невырожденной системы линейных уравнений. Поэтому описанная процедура покоординатного поиска экстремума целевой функции так же, как и менее общая процедура численного решения системы линейных алгебраических уравнений, в литературе известны как **покоординатный метод Зейделя**.

§ 5.8. Градиентный метод Коши 1-го порядка

Градиентный метод в своём базовом варианте был предложен Коши в 1847 г. [56]. Как и покоординатный метод Зейделя (см. выше), он исходно был предназначен для решения совместной системы уравнений на основе минимизации модуля её невязки $\|\delta(\mathbf{x})\|$. Как известно, Коши первым начал рассматривать линейную часть приращения функции dy как главную часть её общего приращения Δy , обосновав таким путём всё дифференциальное исчисление. Для функции $y(\mathbf{x})$ её 1-й дифференциал представляется в виде скалярного произведения в $\mathcal{E}^n$ в тригонометрической форме:

$$dy = \frac{dy}{d\mathbf{x}} d\mathbf{x} = \left| \frac{dy}{d\mathbf{x}} \right| \cdot |d\mathbf{x}| \cdot \cos \varphi, \quad (447)$$

$1 \times n \quad n \times 1$

где φ — угол в $\mathcal{E}^n$ между градиентом и вектором приращения аргумента. Отсюда следует, что при $|d\mathbf{x}| = \text{const}$ имеем 3 характерных варианта:

$$\left. \begin{aligned} dy &= \max \langle dy \rangle \leftrightarrow (\varphi = 0), \\ dy &= \min \langle dy \rangle \leftrightarrow (\varphi = \pm \pi), \\ dy &= 0 \leftrightarrow (\varphi = \pm \pi/2). \end{aligned} \right\} \quad (448)$$

Поэтому при заданном шаге аргумента $|d\mathbf{x}|$ максимальное возрастание функции $y(\mathbf{x})$ именно в линейной части dy происходит по направлению градиента $\mathbf{g}(\mathbf{x}) = dy/d\mathbf{x}$, а максимальное её убывание происходит в направлении антиградиента $-\mathbf{g}(\mathbf{x}) = -dy/d\mathbf{x}$. Промежуточный вариант при $\varphi = \pm \pi/2$ отвечает направлениям вдоль линий уровня на поверхности $y(\mathbf{x})$ в $\mathcal{E}^{n+1}$. Полное множество касательных к линиям уровня в данной точке $\mathbf{x}$ составляет линейное подпространство размерности $n - 1$, дополняющее ортогонально градиент в евклидовом координатном пространстве $\mathcal{E}^n$.

По методу оптимизации Коши 1-го порядка движение к экстремуму функции $y(\mathbf{x})$ из начальной и из промежуточных точек $\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_k$ в исходном аффинном координатном пространстве $\mathcal{A}^n$ осуществляют по *директивному вектору* — градиенту или антиградиенту функции с заданными шагами вдоль этих направлений.

Если шаги $|dx|$ для каждой из этих точек уменьшаются и в пределе стремятся к «+0», то для некоторого класса целевых функций (например, выпуклых на заданной области аргумента $\mathbf{x}$) для каждой начальной точки $\mathbf{c}_1$ однозначно производится определённая непрерывная линия движения, или *предельная кривая* $\mathbf{x} = \mathbf{x}(\alpha)$, приводящая в итоге из $\mathbf{c}_1$ в точку экстремума $\mathbf{s}^*$. Эта кривая получается точным образом решением **дифференциального уравнения Коши**:

$$d\mathbf{x}/d\alpha = \pm \beta(\alpha) \cdot \mathbf{g}'[\mathbf{x}(\alpha)], \alpha \geq 0, \mathbf{x}(0) = \mathbf{c}_1.$$

Приближённое (разностное) решение уравнения **методом ломаных Эйлера** даёт вышеуказанный градиентный метод Коши для численного поиска экстремума эволюционной скалярной функции $y(\mathbf{x})$. Анализ устойчивости решения данного дифференциального уравнения Коши, а, следовательно, и анализ сходимости градиентного метода Коши для конкретной целевой функции $y(\mathbf{x})$ вообще осуществляют, используя **метод функций Ляпунова** [8].

Поскольку слишком частое (т. е. пошаговое) вычисление градиента целевой функции увеличивает трудоёмкость градиентного метода, то гораздо большее распространение получили его же директивные модификации: *метод крутого восхождения* (при поиске максимума) и *метод скорейшего спуска* (при поиске минимума). Направление движения из точки $\mathbf{c}_k$ по градиенту или антиградиенту сохраняется прямолинейным в плоскости $\langle \mathbf{x} \rangle$ вплоть до достижения промежуточного экстремального значения функции $y(\mathbf{c}_{k+1})$. В каждой новой точке $\mathbf{c}_{k+1}$ опять вычисляют градиент или антиградиент. Затем повторяют ту же самую процедуру одномерной оптимизации, но уже вдоль нового прямолинейного направления в $\langle \mathbf{x} \rangle$. Этим прямолинейным отрезкам в плоскости $\langle \mathbf{x} \rangle$, в свою очередь, на поверхности $y(\mathbf{x})$ отвечают некие ломаные траектории, изменяющие свои направления для каждой новой промежуточной точки. Рассмотрим: как это происходит.

Ввиду того, что в промежуточной точке $\mathbf{c} \in \langle \mathbf{x} \rangle$ целевая функция $y(\mathbf{x})$ по данному направлению стационарная (т. е. $dy = 0$), то в ней и в (447) $\varphi = 0$. Отсюда это направление в точке $\mathbf{c}$ совпадает с касательной к линии уровня. Согласно (448), новый градиент в точке $\mathbf{c}$ ортогонален данной касательной. Поэтому и новое направление движения в ней изменяется перпендикулярно предыдущему направлению. В итоге поэтапные движения как по методу крутого восхождения, так и по методу скорейшего спуска происходят *перпендикулярными зигзагами*.

Такова геометрическая картина этих двух главных модификаций градиентного метода Коши. В зигзагообразности общего движения заключается главный недостаток обеих модификаций, несмотря на их геометрическую наглядность и очевидную простоту реализации. В каждой новой промежуточной точке $\mathbf{c}_k$ вычисляют новый градиент, практически ортогональный предыдущему. Процесс, например, при поиске максимума заканчивается тогда, когда

$$y(\mathbf{c}_{k+1}) - y(\mathbf{c}_k) \leq \varepsilon. \quad (449)$$

По директивным модификациям метода Коши точки промежуточных экстремумов $\mathbf{c}_{k+1}$ вычисляются одномерной оптимизацией:

$$f(\alpha_{k+1}) = y[\mathbf{c}_k \pm \alpha_{k+1} \cdot \mathbf{g}'(\mathbf{c}_k)] = \text{extr}, \quad \mathbf{c}_{k+1} = \mathbf{c}_k \pm (\alpha_{k+1})^\bullet \cdot \mathbf{g}'(\mathbf{c}_k). \quad (450)$$

Причём одномерная оптимизация может осуществляться как численным методом, так и аналитическим или интерполяционным способом.

Скорость сходимости и метода крутого восхождения, и метода скорейшего спуска в большой степени зависит от соотношения между максимальным и минимальным собственными значениями матрицы Гессе целевой функции в заданной области. Это совершенно очевидно чисто геометрически. Когда отношение близко к 1, то линии уровня поверхности $y(\mathbf{x})$ близки по форме к окружностям (сферам), что способствует направлению градиента именно к истинному максимуму, и антиградиента — именно к истинному минимуму функции, если таковые имеются и реализуются в точечном виде. При такой картине движения скорость сходимости методов весьма высокая (конкретно, отвечает скорости сходимости для методов одномерной оптимизации). Однако, если это отношение заметно более 1, то линии уровня имеют *овражную* форму, т. е. вытянуты вдоль некоторых направлений в $\langle \mathbf{x} \rangle$. При этом зигзагообразная траектория общего движения к экстремуму петляет перпендикулярно поперёк, а не вдоль *оврага*, что резко замедляет скорость сходимости этих методов.

В связи с указанным обстоятельством особенно важно учитывать такой возможный фактор воздействия на скорость сходимости обоих методов, как изменение масштабов по осям m_i — для *неоднородных* переменных. Ведь исходно выбор этих масштабов для них совершенно произволен. Масштабы по осям существенно влияют на формирование главной диагонали матрицы Гессе — её наиболее действенной части в определении собственных значений и отсюда формы линий уровня скалярной функции $y(\mathbf{x})$.

В итоге, суммируя всё вышесказанное, можно сделать главный вывод, что градиентный метод даёт идеальную сходимость, отвечающую методам *одномерной* оптимизации, именно в том случае, когда целевая функция $y(\mathbf{x})$ с точностью до коэффициентов модели неполного 2-го порядка имеет вид:

$$y(\mathbf{x}) \approx k(\mathbf{x}) = y(\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n) + \sum_{i=1}^n g_i (x_i - c_i) + 1/2 \cdot g \cdot \sum_{i=1}^n (x_i - c_i)^2,$$

где $g = \text{const}$, а g_i (компоненты градиента), которые могут как угодно различаться между собой. Разумеется, такой требуемый вид целевой функции, даже в модели 2-го порядка, маловероятен. Поэтому отсюда имеется слишком большой возможный разброс в скоростях сходимости процедур по градиентному методу — от максимальной до нулевой. Это же значительно ограничивает применение градиентного метода в его первоизданном виде. Данный недостаток преодолевается в излагаемом далее масштабно-градиентном методе.

§ 5.9. Масштабно-градиентный метод неполного 2-го порядка и его директивная модификация

Рассмотренный выше градиентный метод, в сущности, базируется на аппроксимации целевой функции в окрестности $\mathbf{c}$ моделью 1-го порядка:

$$y(\mathbf{x}) \approx l(\mathbf{x}) = y(\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n) + \sum_{i=1}^n g_i (x_i - c_i) = y(\mathbf{c}) + \mathbf{g}(\mathbf{c}) \cdot (\mathbf{x} - \mathbf{c}), \quad (451)$$

где g_i — компоненты $1 \times n$ -вектора градиента в точке $\mathbf{c}$. Но заметно точнее исходную целевую функцию $y(\mathbf{x})$ в окрестности $\mathbf{c}$ аппроксимирует модель неполного 2-го порядка вида:

$$\begin{aligned} y(\mathbf{x}) \approx k(\mathbf{x}) &= y(\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n) + \sum_{i=1}^n g_i (x_i - c_i) + 1/2 \cdot \sum_{i=1}^n g_{ii} (x_i - c_i)^2 = \\ &= y(\mathbf{c}) + \mathbf{g}(\mathbf{c}) \cdot (\mathbf{x} - \mathbf{c}) + 1/2 \cdot (\mathbf{x} - \mathbf{c})' \cdot \mathbf{D}(\mathbf{c}) \cdot (\mathbf{x} - \mathbf{c}), \end{aligned} \quad (452)$$

где $g_{ii} \neq 0$ — компоненты диагональной $n \times n$ -матрицы Гессе $\mathbf{D}$ в точке $\mathbf{c}$. Кроме того, важно то, что при знакоопределённых коэффициентах $g_{ii} \neq 0$ эта модель позволяет легко получать очередное $(k + 1)$ -е приближение к точке экстремума.

В каждой новой точке $\mathbf{c}_k$ сразу же имеем очередное приближение $\mathbf{c}_{k+1}$ к точке экстремума $\mathbf{s}^*$ целевой функции $y(\mathbf{x})$ в результате решения градиентного уравнения стационарности типа:

$$\begin{aligned} d\mathbf{k}(\mathbf{x})/d\mathbf{x}' &= \mathbf{g}'(\mathbf{c}_k) + D(\mathbf{c}_k) \cdot (\mathbf{x} - \mathbf{c}_k) = \mathbf{0} \rightarrow \\ \rightarrow \mathbf{c}_{k+1} &= \mathbf{c}_k + [-D(\mathbf{c}_k)]^{-1} \cdot \mathbf{g}'(\mathbf{c}_k) = \mathbf{c}_k + \begin{vmatrix} g_1 / -g_{11} \\ \dots \\ g_n / -g_{nn} \end{vmatrix}. \end{aligned} \quad (453)$$

Обратим внимание на простоту обращения матрицы Гессе в данном методе. Как и все методы 2-го порядка, он инвариантен по отношению к выбору масштабов m_i по осям аффинных координат x_i . Так, выполним их произвольное масштабное преобразование $\mathbf{u} = D_m \mathbf{x}$, где D_m — диагональная положительная масштабная матрица, и преобразуем в новых масштабах вторую часть формулы (453). В итоге имеем то же самое инвариантное решение:

$$\begin{aligned} D_m \cdot \mathbf{x}_{(k+1)} &= D_m \cdot \mathbf{x}_{(k)} + [-D_m^{-1} D(\mathbf{x}_{(k)}) D_m^{-1}]^{-1} \cdot D_m^{-1} \cdot \mathbf{g}'(\mathbf{x}_{(k)}) \rightarrow \\ \rightarrow \mathbf{u}_{(k+1)} &= \mathbf{u}_{(k)} + [-D(\mathbf{u}_{(k)})]^{-1} \cdot \mathbf{g}'(\mathbf{u}_{(k)}) = \mathbf{u}_{(k)} + D_m \begin{vmatrix} g_1 / -g_{11} \\ \dots \\ g_n / -g_{nn} \end{vmatrix}. \end{aligned} \quad (454)$$

Следовательно, (453) и (454) отображают один и тот же процесс, но в разных масштабно преобразованных системах аффинных координат в $\mathcal{A}^n$. В частности, масштабная матрица типа $D_m = \sqrt{g^{-1} \cdot [\pm D(\mathbf{c}_k)]}$ преобразует матрицу Гессе $[-D(\mathbf{c}_k)]$ к виду $\{\pm g \cdot I\}$, а сам процесс — к форме градиентного метода Коши при $g_{ii} = g = \text{const}$ (см. выше в § 5.8, в том числе при $g = 1$).

$$\mathbf{v}_{(k+1)} = \mathbf{v}_{(k)} \pm g^{-1} \cdot \mathbf{g}'(\mathbf{v}_{(k)}). \quad (455)$$

Поэтому данный метод можно трактовать как градиентный метод, но в котором все масштабы по осям выбраны так, чтобы нивелировать различие в собственных значениях g_{ii} неполной (т. е. диагональной) матрицы Гессе $D(\mathbf{c}_k)$ — для повышения скорости сходимости процесса. Но, что самое главное, в данном методе именно выбору масштабов по осям x_i (как в градиентном методе) явно не придаётся значения. В частности, борьба с *овражной* сходимостью (см. § 5.8) выполняется автоматически. Отсюда и принятое тут название метода, вынесенное в заголовок параграфа.

Масштабно-градиентный метод допускает также директивную модификацию с побочным движением, как и в методе с одномерной оптимизацией типа (450), в направлении на $\mathbf{c}_{k+1}$, т. е. в данном варианте — по *директивному вектору* неполного 2-го порядка:

$$\mathbf{j}(\mathbf{c}_k) = [-D(\mathbf{c}_k)]^{-1} \cdot \mathbf{g}'(\mathbf{c}_k). \quad (456)$$

Этот вектор по компонентам имеет те же самые размерности, что и вектор $\mathbf{x}$. В его директивной модификации при побочной одномерной оптимизации весьма логично применять однотипные процедуры 2-го порядка из § 5.5.

Масштабно-градиентный метод базируется на модели (452) неполного 2-го порядка. Геометрически это трактуется в $\mathcal{A}^n$ как аппроксимация поверхности $y(\mathbf{x})$ на окрестности точки $\mathbf{x}_k$ эллиптическим параболоидом с эллиптическими n осями, направленными как оси координат базиса $\mathbf{x}_i$, и с параболической осью, параллельной оси y . Это, разумеется, есть значительно более точная аппроксимация поверхности $y(\mathbf{x})$, нежели плоскостью в градиентном методе.

§ 5.10. Общий метод Ньютона 2-го порядка и его директивная модификация

Объяснение названию изучаемого метода было дано ранее в § 5.5. В рассматриваемом его общем — многомерном варианте применяется аппроксимация целевой функции моделью полного 2-го порядка:

$$\begin{aligned} y(\mathbf{x}) \approx q(\mathbf{x}) &= y(c_1, c_2, \dots, c_n) + \sum_{i=1}^n g_i (x_i - c_i) + 1/2 \sum_{i \neq j}^n \sum_{j=1}^n g_{ij} (x_i - c_i)(x_j - c_j) = \\ &= y(\mathbf{c}) + \mathbf{g}(\mathbf{c}) \cdot (\mathbf{x} - \mathbf{c}) + 1/2 \cdot (\mathbf{x} - \mathbf{c})' \cdot \mathbf{G}(\mathbf{c}) \cdot (\mathbf{x} - \mathbf{c}), \end{aligned} \quad (457)$$

где $g_{ij} = g_{ji}$ — компоненты симметричной $n \times n$ -матрицы Гессе $\mathbf{G}$ в точке $\mathbf{c}$. Геометрически это трактуется в $\mathcal{A}^n$ как аппроксимация поверхности $y(\mathbf{x})$ в окрестности точки $\mathbf{x}_k$ эллиптическим параболоидом с эллиптическими n осями, направленными как собственные векторы матрицы Гессе, и с параболической осью, параллельной оси y . Это, разумеется, есть гораздо более точная аппроксимация $y(\mathbf{x})$, нежели плоскостью в градиентном методе и параболоидом в масштабнo-градиентном методе. Однако необходимое тут обращение матрицы Гессе — вычислительно трудоёмкая операция. Оно применяется именно для поиска очередного приближения к точке экстремума.

Общий метод Ньютона формально следует из градиентного метода Коши, но с общей *линейной* оптимизацией векторной переменной.

Так, если $\det G \neq 0$, то данная модель позволяет выполнять очередную итерацию в результате решения градиентного уравнения стационарности:

$$dq(\mathbf{x})/d\mathbf{x}' = \mathbf{g}'(\mathbf{c}_k) + G(\mathbf{c}_k) \cdot (\mathbf{x} - \mathbf{c}_k) = \mathbf{0} \rightarrow \mathbf{c}_{k+1} = \mathbf{c}_k + [-G(\mathbf{c}_k)]^{-1} \cdot \mathbf{g}'(\mathbf{c}_k). \quad (458)$$

Это соответствует поиску строгого экстремума уровня $p = 2$ (§ 1.1) любого характера. Как все методы 2-го порядка, общий метод Ньютона инвариантен по отношению к выбору масштабов m_i по осям аффинных координат x_i . Хотя, конечно, этот выбор как-то влияет на скорость его сходимости, если рассматривать влияние членов в разложении $y(\mathbf{x})$ по формуле Тейлора порядка более 2-х. Для функции порядка 2, как $q(\mathbf{x})$ в (457), этот метод даёт из (458) сразу же искомый экстремум за одну итерацию — при условии, что он имеется и что он строгий (точечный); при этом $G = \text{Const}$ и $\det G \neq 0$.

Если не учитывать все вычислительные затраты на обращения матрицы Гессе, необходимые после каждой итерации, то в общем методе Ньютона поэтапное сужение интервала оценивается аналогично и по тем же самым соображениям, что было показано в § 5.5. Степень сжатия отображения $\mathbf{x} = \varphi(\mathbf{x})$ на каждой итерации оценивается сверху аналогично формуле (440) — через коэффициент L_φ в требовании непрерывности по Липшицу для функции $\varphi(\mathbf{x})$:

$$\|\mathbf{s}^* - \mathbf{c}_{k+1}\| \leq \frac{L_\varphi^k}{1 - L_\varphi} \cdot \|\mathbf{c}_2 - \mathbf{c}_1\|.$$

Здесь для начальной точки $\mathbf{c}_1$ целевая функция в её окрестности должна быть равномерно непрерывной и строго выпуклой (вогнутой), причём так, чтобы выполнялось хотя бы требование $L_\varphi < 1$, а лучше $L_\varphi < 1/2$. При каждой новой итерации коэффициент L_φ продолжает уменьшаться и часто квадратично. Это и обеспечивает скорость сходимости общего метода Ньютона, согласно Л. В. Канторовичу, также порядка 2, а не 1 [20, 21]. Но последнее имеет место при выполнении определённых требований к матрице Гессе в точке $\mathbf{c}_1$ и к матрице третьих производных на области $\langle \mathbf{x} \rangle$.

Общий метод Ньютона тоже допускает *директивную* модификацию при побочном движении с одномерной оптимизацией в направлении на $\mathbf{c}_{k+1}$, т. е. по *директивному вектору* 2-го порядка типа:

$$\mathbf{p}(\mathbf{c}_k) = [-G(\mathbf{c}_k)]^{-1} \cdot \mathbf{g}'(\mathbf{c}_k). \quad (459)$$

Этот вектор по компонентам имеет те же размерности, что и вектор $\mathbf{x}$. При произвольном масштабном преобразовании $\mathbf{u} = \mathbf{D}_m \mathbf{x}$ имеем:

$$\mathbf{p}(\mathbf{u}_{(k)}) = [-\mathbf{D}_m^{-1} \cdot \mathbf{G}(\mathbf{x}_{(k)}) \cdot \mathbf{D}_m^{-1}]^{-1} \cdot \mathbf{D}_m^{-1} \cdot \mathbf{g}'(\mathbf{x}_{(k)}) = \mathbf{D}_m \cdot \mathbf{p}(\mathbf{x}_{(k)}). \quad (460)$$

Т. е. и здесь директивный вектор преобразуется масштабно всегда так, как и $\mathbf{x}$! В общем методе Ньютона, как в предыдущем, для одномерной (побочной) оптимизации $y(\mathbf{x})$ весьма логично применять однотипные с ним процедуры 2-го порядка из § 5.5, в том числе их разностные модификации. Директивная модификация, разумеется, значительно повышает эффективность общего метода Ньютона в целом. Но, что особо важно, снижается трудоёмкость вычислений, затрачиваемых на обращения матриц Гессе (их количество сокращается). Метод Ньютона для повышения эффективности оптимизации в целом целесообразно применять после градиентного и масштабно-градиентного метода.

Если матрица Гессе $\mathbf{G}(\mathbf{x})$ в окрестности строгого экстремума $y(\mathbf{x})$ плохо обусловлена (т. е. $\det \mathbf{G} \approx 0$) или в окрестности нестрогого экстремума $y(\mathbf{x})$ вырождена (т. е. $\det \mathbf{G} = 0$ или более общо $\text{rang } \mathbf{G} < n$), то общий метод Ньютона малоэффективен или даже неэффективен. В первом случае помогает укрупнение масштабов m_i по осям базиса, но это снижает точность оптимизации. Универсальный способ разрешения проблемы состоит в применении метода квадратичной регуляризации по Тихонову (см. § 5.10).

Отметим также, что рассмотренные выше основные *именные* методы многомерной оптимизации изначально были предложены для иных вычислительных целей, т. е. методы Коши и Зейделя — для численного решения систем совместных линейных уравнений, метод Ньютона — для численного нахождения корней уравнения $f(\mathbf{x}) = 0$. Они оказались настолько всеобщими, что с успехом в изложенных вариантах нашли применение и для оптимизации целевых функций. Если вспомнить, что и решение системы линейных уравнений $\mathbf{B}\mathbf{x} = \mathbf{a}$ достигается минимизацией модуля невязки — также целевой функции от векторного аргумента, то все рассмотренные методы многомерной оптимизации 0-го, 1-го и 2-го порядков можно поставить в соответствие релаксационным методам решения систем линейных уравнений.

Кроме того, для ряда многомерных методов 1-го и 2-го порядков возможность их разностных модификаций оговаривалась пока только для директивных модификаций — на стадии побочной (одномерной) оптимизации. Однако для главных (многомерных) итераций разностная модификация тоже возможна, но она потребует вычисления значений целевой функции по неким определённым планам расположения точек в координатном пространстве (гл. 6 и 7).

§ 5.11. Регуляризация по Тихонову в методах 2-го порядка

Метод регуляризации, предназначенный для численного решения некорректных задач линейной алгебры и анализа (см. ранее в § 2.4), изначально был предложен для этого А. Н. Тихоновым в 1965 г. Далее этот метод всесторонне развивался им и его школой в ряду многочисленных последующих публикаций [36, 37]. Главная цель данного весьма общего метода заключается в том, чтобы при решении неких особых задач, связанных с обращением матричных операторов, в условиях производимой обращением явной неопределённости или множественности решения вычислять однозначную, или корректную по Тихонову *точечную* оценку решения исходной задачи.

Например, с такими подобными проблемами часто сталкиваются при решении численными методами 2-го порядка задач на экстремум целевых функций $y(\mathbf{x})$ хотя с определённой, но плохо обусловленной, или даже с полуопределённой матрицей Гессе $d^2y/d\mathbf{x}'d\mathbf{x}$. В широкой трактовке метода регуляризации Тихонова дополнительно применяется некоторая специально выбираемая функция $\alpha \cdot \theta(\mathbf{x})$, или *стабилизатор*. Соответственно при этом оптимизируют композиционную функцию Тихонова: $T_\alpha(\mathbf{x}) = y(\mathbf{x}) + \alpha \cdot \theta(\mathbf{x})$. Скалярный множитель α выбирают, с учётом необходимой размерности и достаточно малым по величине. Стабилизатор $\alpha \cdot \theta(\mathbf{x})$ должен быть того же порядка малости, что и общая погрешность для $y(\mathbf{x})$. Точечный экстремум скалярной функции $\theta(\mathbf{x})$ на области плохой обусловленности или вырожденности общего решения исходной задачи $\langle \mathbf{x} \rangle$ определяет её корректное (т. е. точечное) решение. Оно, вообще говоря, не является каким-то наилучшим решением этой задачи. Главное здесь то, что метод регуляризации в итоге обеспечивает результативность общего процесса оптимизации.

Квадратичная регуляризация, применённая в первой публикации [36], исходит из требования минимума квадрата евклидовой нормы для радиус-вектора решения задачи ($\min \theta(\mathbf{x}) = \mathbf{x}'\mathbf{x}$ или $\frac{1}{2} \mathbf{x}'\mathbf{x}$). Она даёт однозначное нормальное (квази)решение, например, *плохо обусловленного линейного уравнения* типа $A\mathbf{x} = \mathbf{a}$ при некоем значении множителя корректности α . (При этом лишь при операциях с *точными* переменными и их целевыми функциями оправдано применение значения $\alpha \rightarrow 0$.) Причём нормальные решения не инвариантны к масштабным преобразованиям — см. § 4.5.

Такой подход, например, может применяться в вышерассмотренных методах численной оптимизации с моделями целевых функций $y(\mathbf{x})$ неполного и полного 2-го порядка.

§ 5.12. Условная численная оптимизация

Ранее в гл. 2 (§§ 2.1—2.4) была рассмотрена *аналитическая* условная оптимизация целевой функции $y = y(\tilde{x})$ на некоторой гладкой регулярной q -поверхности $\langle \tilde{x} \rangle \subset \mathcal{A}^n$, заданной или параметрическим способом, или ограничительным способом. Оба способа, в сущности, являются функциональными. Принципиальное же различие между ними состоит в том, что в первом случае множество $\langle \tilde{x} \rangle$ является *образом* некоего преобразования, а во втором случае множество $\langle \tilde{x} \rangle$ является *ядром* некоего преобразования. Это позволяет при выполнении определённых требований к $y(\tilde{x})$ и $\langle \tilde{x} \rangle$ далее вполне естественным путём переходить к *численным* методам решения подобных задач условной оптимизации. Иные же возможные способы задания допустимого множества $\langle \tilde{x} \rangle$ для векторной переменной $\tilde{x}$ не рассматриваются ввиду того, что выходят за рамки содержания данной монографии. А именно, такие способы относятся уже к сфере интересов математического программирования [см., например, 21].

Если переменная $\tilde{x}$ принадлежит некоторой q -плоскости $\langle \tilde{x} \rangle \subset \mathcal{A}^n$, то для условной оптимизации целевой функции $y(\tilde{x})$ логично использовать тоже условные (т. е. проективные) аналоги директивных векторов 1-го и 2-го порядков с применением симметричных характеристических проекторов $\overleftarrow{AA'}$ или $\overrightarrow{A'A}$ (см. §§ 2.1—2.2).

Что очень важно, эти проекторы для обоих вышеуказанных способов задания плоскости $\langle \tilde{x} \rangle$ — параметрического и ограничительного тут постоянны и вычисляются из данной матрицы A всего только один раз. В координатном пространстве $\mathcal{A}^n$ они имеют аффинный характер, но в $\mathcal{E}^n$ они проецируют ортогонально.

Но если переменная $\tilde{x}$ принадлежит некоторой гладкой регулярной криволинейной q -поверхности $\langle \tilde{x} \rangle$, то для условной оптимизации целевой функции $y(\tilde{x})$ необходимо применять более общие процедуры. Таковыми, например, являются *методы нормальных проекций* 1-го или 2-го порядков с использованием промежуточной линеаризации q -поверхности $\langle \tilde{x} \rangle$ перед каждой главной итерацией.

В свою очередь, рассмотренные ранее (см. § 2.4) функциональные предельные методы условной аналитической оптимизации целевых функций с большим и малым параметром аналогично трансформируются в соответствующие численные процедуры — с решением предельного уравнения стационарности и её анализом.

§ 5.12.1. Методы 1-го и 2-го порядка для функций от линейно зависимой или ограниченной векторной переменной

В этом параграфе рассматриваются численные методы 1-го и 2-го порядков условной оптимизации функции $y = y(\tilde{\mathbf{x}})$, где $\tilde{\mathbf{x}} \in \langle \tilde{\mathbf{x}} \rangle \subset \mathcal{A}^n$ — или линейно зависимая, или линейно ограниченная вектор-переменная. С целью упрощения общей задачи примем, что $\langle \tilde{\mathbf{x}} \rangle \subset \mathcal{D} \subset \mathcal{A}^n$, где $\mathcal{D}$ — область определения целевой функции $y(\mathbf{x})$. Кроме того, при *условном* характере её экстремума имеем повсюду $dy/d\mathbf{x}(\tilde{\mathbf{x}}) \neq \mathbf{0}$.

В первом варианте плоскость $\langle \tilde{\mathbf{x}} \rangle$ задаётся параметрически (§ 2.1):

$$\langle \tilde{\mathbf{x}} \rangle: \tilde{\mathbf{x}} = l_1(\mathbf{u}) = \tilde{\mathbf{c}} + A_1(\mathbf{u} - \mathbf{c}) = A_1\mathbf{u}, \quad A_1 \cdot \mathbf{c} = \tilde{\mathbf{c}} \in \langle \tilde{\mathbf{x}} \rangle; \quad (461)$$

где $\mathbf{u} \in \mathcal{A}^q$ — независимая вектор-переменная размерности $q < n$, $\mathbf{c}$ — точка в $\mathcal{A}^q$, $A_1 = \text{Const}$ — $n \times q$ -матрица трансляции из $\mathcal{A}^q$ в $\mathcal{A}^n$, $\text{rang } A_1 = q$.

Во втором варианте плоскость $\langle \tilde{\mathbf{x}} \rangle$ задаётся ограничительно (§ 2.2):

$$\langle \tilde{\mathbf{x}} \rangle: l_2(\tilde{\mathbf{x}}) = A_2\tilde{\mathbf{x}} - \mathbf{a} = \mathbf{0}, \quad \mathbf{a} \in \langle \text{im } A_2 \rangle, \quad (462)$$

где $l_2 \in \mathcal{A}^m$ — функция ограничения размерности $m < n$, $\mathbf{a}$ — точка в $\mathcal{A}^m$, $A_2 = \text{Const}$ — $m \times n$ -матрица преобразования из $\mathcal{A}^n$ в $\mathcal{A}^m$, $\text{rang } A_2 = m$. Важно, в частности, если $\tilde{\mathbf{c}} \in \langle \tilde{\mathbf{x}} \rangle$, то $A_2\tilde{\mathbf{c}} - \mathbf{a} = \mathbf{0}$ и $l_2(\tilde{\mathbf{x}}) = A_2(\tilde{\mathbf{x}} - \tilde{\mathbf{c}}) = \mathbf{0}$ — упрощённая линейная функция ограничения.

При $q = n - m$ имеется однозначная взаимосвязь обоих способов задания q -плоскости $\langle \tilde{\mathbf{x}} \rangle$, что определяется, во-первых, одним и тем же начальным значением переменной $\tilde{\mathbf{x}} = \tilde{\mathbf{c}}$ и, во-вторых, сингулярными соотношениями типа:

$$\langle \text{im } A_1 \rangle \equiv \langle \ker A_2 \rangle \leftrightarrow A_1 \overset{\leftarrow}{A_1}' = A_2 \overset{\rightarrow}{A_2}' = \Pi = \Pi', \quad (463)$$

где $\overset{\leftarrow}{A_1}'$ — проектор в $\mathcal{A}^n$ на $\langle \text{im } A_1 \rangle$ параллельно $\langle \ker A_1' \rangle$,

$\overset{\rightarrow}{A_2}'$ — проектор в $\mathcal{A}^n$ на $\langle \ker A_2 \rangle$ параллельно $\langle \text{im } A_2' \rangle$.

(В евклидовом координатном пространстве они попросту ортогопроекторы.)

Соответственно изложенные в §§ 5.8, 5.9 и 5.10 методы оптимизации трансформируются в условные аналоги. Для реализации этого подхода директивные векторы проецируются своими проекторами на допустимые направления, т. е. из начальной точки $\tilde{\mathbf{c}} \in \langle \tilde{\mathbf{x}} \rangle$ на q -плоскость $\langle \tilde{\mathbf{x}} \rangle$.

По методу *условного* градиента или антиградиента направление движения из начальной точки $\tilde{\mathbf{c}} \in \langle \tilde{\mathbf{x}} \rangle$ осуществляется по директивным векторам условного градиента или антиградиента

$$\pm \left[\frac{dy}{dx}(\tilde{\mathbf{c}}) \right]' = \pm \Pi \cdot \left(\frac{dy}{dx}(\tilde{\mathbf{c}}) \right)' \quad (464)$$

вдоль некоей прямолинейной траектории в плоскости $\langle \tilde{\mathbf{x}} \rangle$ — вплоть до достижения условно экстремального значения функции $y(\tilde{\mathbf{c}}^*)$ по этому направлению. После данной операции в точке $\tilde{\mathbf{c}}^*$ вычисляют условный градиент или антиградиент и затем повторяют процедуру одномерной оптимизации $y(\tilde{\mathbf{x}})$ вдоль нового прямолинейного направления в пределах плоскости $\langle \tilde{\mathbf{x}} \rangle$. Направление всегда выдерживается автоматически в $\langle \tilde{\mathbf{x}} \rangle$, в силу *линейного* характера переменной $\tilde{\mathbf{x}}$ в $\mathcal{A}^n$.

В *условном* аналоге масштабно-градиентного метода к итерации (453) применяется то же проективное преобразование:

$$\begin{aligned} \left[\frac{dk(\mathbf{x})}{dx'} \right] &= \Pi \cdot \mathbf{g}'(\tilde{\mathbf{c}}_k) + \Pi \cdot D(\tilde{\mathbf{c}}_k) \cdot \Pi \cdot (\tilde{\mathbf{x}} - \tilde{\mathbf{c}}_k) = \mathbf{0} \rightarrow \\ \rightarrow \tilde{\mathbf{c}}_{k+1} &= \tilde{\mathbf{c}}_k + \Pi \cdot [-D(\tilde{\mathbf{c}}_k)]^{-1} \cdot \Pi \cdot \mathbf{g}'(\tilde{\mathbf{c}}_k). \end{aligned} \quad (465)$$

Наконец, в *условном* аналоге общего метода Ньютона к итерации (458) применяется то же самое проективное преобразование:

$$\begin{aligned} \left[\frac{dq(\mathbf{x})}{dx'} \right] &= \Pi \cdot \mathbf{g}'(\tilde{\mathbf{c}}_k) + \Pi \cdot G(\tilde{\mathbf{c}}_k) \cdot \Pi \cdot (\tilde{\mathbf{x}} - \tilde{\mathbf{c}}_k) = \mathbf{0} \rightarrow \\ \rightarrow \tilde{\mathbf{c}}_{k+1} &= \tilde{\mathbf{c}}_k + \Pi \cdot [-G(\tilde{\mathbf{c}}_k)]^{-1} \cdot \Pi \cdot \mathbf{g}'(\tilde{\mathbf{c}}_k). \end{aligned} \quad (466)$$

Последние два метода условной численной оптимизации допускают тоже свои директивные модификации — побочные движения в пределах плоскости $\langle \tilde{\mathbf{x}} \rangle$ из точки $\tilde{\mathbf{c}}_k$ в направлении на $\tilde{\mathbf{c}}_{k+1}$, т. е. по своим условным директивным векторам 2-го порядка с использованием процедуры одномерной оптимизации также 2-го порядка.

§ 5.12.2. Методы нормальных проекций 1 и 2-го порядка

Эти методы базируются на идее промежуточных линеаризаций гладкой и регулярной криволинейной q -поверхности $\langle \tilde{\mathbf{x}} \rangle$ в начале каждой главной итерации, т. е. последовательно в точках $\tilde{\mathbf{c}}_1, \tilde{\mathbf{c}}_2, \dots, \tilde{\mathbf{c}}_k \in \langle \tilde{\mathbf{x}} \rangle$. Линеаризация осуществляется, согласно требованиям леммы 1 (§ 2.1) и леммы 2 (§ 2.2), для аппроксимации в них условных (проективных) директивных векторов.

При параметрическом способе задания q -поверхности $\langle \tilde{\mathbf{x}} \rangle$ в точке $\tilde{\mathbf{c}}$ имеем линеаризацию: $\langle I \rangle \equiv \tilde{\mathbf{c}} + \langle \text{im } d\tilde{\mathbf{x}}/d\mathbf{u}(\mathbf{c}) \rangle$ — см. формулу (63).

При ограничительном способе задания q -поверхности $\langle \tilde{\mathbf{x}} \rangle$ в точке $\tilde{\mathbf{c}}$ имеем линеаризацию: $\langle I \rangle \equiv \tilde{\mathbf{c}} + \langle \ker d\mathbf{h}/d\tilde{\mathbf{x}}(\tilde{\mathbf{c}}) \rangle$ — см. формулу (73).

Здесь применяются значения функциональных первых тензор-производных в исходной точке:

$d\tilde{\mathbf{x}}/d\mathbf{u}$ — $m \times q$ -матрица Якоби постоянного ранга $q < n$,

$d\mathbf{h}/d\tilde{\mathbf{x}}$ — $m \times n$ -матрица Якоби постоянного ранга $m < n$.

Далее с целью наглядной геометрической интерпретации общей процедуры перейдём в евклидово координатное пространство $\mathcal{E}^n$, соответственно при этом введя в нём понятие евклидова расстояния между его отдельными точками. Хотя, разумеется, при неоднородных частных переменных x_i (см. § 5.6) такое понятие имеет неоднозначный характер и находится всегда в зависимости от выбираемых масштабов по осям координат, которые, по сути, произвольные. Это в подобных случаях надо учитывать.

Главные (итерации) или побочные движения из каждой точки $\tilde{\mathbf{c}} \in \langle \tilde{\mathbf{x}} \rangle$ выполняют, согласно директивным векторам или (464), или (465), или (466) в зависимости от используемого основного метода для условной оптимизации. В результате вначале делают аппроксимацию $\tilde{\mathbf{c}}^\circ \notin \langle \tilde{\mathbf{x}} \rangle$ для точки промежуточного экстремума $\tilde{\mathbf{c}}^\bullet \in \langle \tilde{\mathbf{x}} \rangle$ целевой функции $y(\tilde{\mathbf{x}})$ по заданному направлению. Ведь вследствие искривления q -поверхности $\langle \tilde{\mathbf{x}} \rangle$ траектория движения по ней из $\tilde{\mathbf{c}}$ в заданном направлении в $\mathcal{E}^n$ криволинейная. Следовательно, точка $\tilde{\mathbf{c}}^\circ$ находится в $\mathcal{E}^n$ на некотором расстоянии d от $\langle \tilde{\mathbf{x}} \rangle$. Если q -поверхность $\langle \tilde{\mathbf{x}} \rangle$ имеет положительную или отрицательную кривизну (без изменения знака), то при её гладкости и регулярности однозначно решается задача минимизации в $\mathcal{E}^n$ квадрата расстояния d от точки $\tilde{\mathbf{c}}^\circ$ до поверхности $\langle \tilde{\mathbf{x}} \rangle$ с вычислением новой точки промежуточного условного экстремума $y(\tilde{\mathbf{x}})$ как ортопроекции $\tilde{\mathbf{c}}^\bullet \in \langle \tilde{\mathbf{x}} \rangle$. Минимальное расстояние в $\mathcal{E}^n$ геометрически здесь отображает вектор нормальной проекции из $\tilde{\mathbf{c}}^\circ$ на $\langle \tilde{\mathbf{x}} \rangle$.

По тем же самым геометрическим соображениям и при данных требованиях к q -поверхности этот однозначный вектор ортогонален в точке пересечения с поверхностью по отношению к новой касательной к ней в этой же точке аппроксимирующей q -плоскости $\langle l \rangle$ — см. выше. Теоретически он вычисляется из уравнения относительно $\tilde{\mathbf{x}}$

$$\Pi(\tilde{\mathbf{x}}) \cdot (\tilde{\mathbf{x}} - \tilde{\mathbf{c}}^{\circ}) = \mathbf{0}, \quad (467)$$

где функциональный характеристический ортопроектор $\Pi(\tilde{\mathbf{x}})$ в $\mathcal{E}^n$ выражается в 2-х вариантах задания q -поверхности (см. §§ 2.1, 2.2) как:

$$\Pi(\tilde{\mathbf{x}}) = \left[\frac{d\tilde{\mathbf{x}}}{d\mathbf{u}} \cdot \left(\frac{d\tilde{\mathbf{x}}}{d\mathbf{u}} \right)^{\leftarrow} \right], \quad \Pi(\tilde{\mathbf{x}}) = \left[\left(\frac{d\mathbf{h}}{d\tilde{\mathbf{x}}} \right)^{\rightarrow} \cdot \frac{d\mathbf{h}}{d\tilde{\mathbf{x}}} \right]. \quad (468, 469)$$

Вычисленная в результате решения именно вспомогательной задачи квадратичной минимизации точка $\tilde{\mathbf{c}}^{\circ} \in \langle \tilde{\mathbf{x}} \rangle$ является исходной точкой для повторения вышеописанной процедуры. Таким образом, в итоге приходят к точке $\tilde{\mathbf{s}}^{\circ}$ условного экстремума целевой функции $y(\tilde{\mathbf{x}})$.

§ 5.12.3. Методы с большим и малым параметром

Ранее в § 2.4 было дано теоретическое обоснование предельных методов условной оптимизации функций $y(\tilde{\mathbf{x}})$ с большим и малым параметром, а также показана их полная тождественность (парное соответствие) и общая сфера применимости. Поскольку аналитическое решение предельных уравнений стационарности ($N \rightarrow \infty$ или $\alpha \rightarrow 0$) возможно только в некоторых простейших случаях, то для полной реализации всех возможностей предельных методов целесообразно прибегать к численным процедурам с заданной точностью.

В сущности обоих методов заложено следующее. Если условный экстремум в задаче с ограниченной векторной переменной $\tilde{\mathbf{x}}$ (§ 2.2) существует, то точка экстремума $\tilde{\mathbf{s}}^{\circ}$ обязательно является решением предельных уравнений типа (109) и (111), т. е. в широкой трактовке этих методов, или при симметрии матрицы Якоби $d\mathbf{h}/d\mathbf{x}$ является решением предельных уравнений (100) и (103), т. е. в специальной трактовке этих методов. Обратное утверждение также верно.

Предельный метод с большим параметром в широкой трактовке известен как *метод штрафных функций*. Он возник и начал развиваться в середине XX века. Идея метода принадлежит Куранту [см., например, 11]. Главным достоинством этого метода является то, что он сводит исходную задачу условной оптимизации к задаче на безусловную оптимизацию!

Численная процедура метода штрафных функций базируется на том, что с увеличением параметра N решение предельного уравнения (109) неуклонно приближается к точке $\tilde{\mathbf{S}}$ условной стационарности (экстремума) целевой функции. Причём в численной интерпретации метода применяется экстраполяция решения уравнения (109) к $N \rightarrow \infty$. Решение разбивается на последовательность безусловных оптимизаций $K_N(\mathbf{x})$ при $N_1, N_2, \dots, N_k$. Конечный этап есть экстраполяция решения к $N \rightarrow \infty$, когда она явно прослеживается. Для повышения точности и реализации сходимости с 2-х сторон от $\tilde{\mathbf{S}}^*$ (по линии экстраполяции) возможна модификация численного метода со значениями большого параметра $\pm N_1, \pm N_2, \dots, \pm N_k, \dots, \pm \infty$. Экстремум аппроксимируется как среднее арифметическое из левых и правых значений.

Этот метод при симметричности матрицы Якоби: $d\mathbf{h}/d\mathbf{x} = (d\mathbf{h}/d\mathbf{x})'$, но при этом $\text{rang } d\mathbf{h}/d\mathbf{x} = m = \text{const}$, в качестве предельного уравнения стационарности имеет более простое прародительское уравнение (103). Тогда $d\mathbf{h}/d\mathbf{x}$, вместе с тем, является и матрицей Гессе для некоторой иной скалярной функции типа $f(\mathbf{x}) = \int \mathbf{h}'(\mathbf{x}) d\mathbf{x}$. В этом случае условный экстремум $y(\tilde{\mathbf{x}})$ находится на области вырожденной стационарности интегральной функции ограничения $f(\mathbf{x})$.

Формально, умножая (103) на $\alpha = 1/N$, продуцируем эквивалентный ему метод с малым числовым параметром α , который иногда более удобен для численной реализации, в том числе при оценке скорости сходимости самой процедуры условной оптимизации [21]. Из линейной аппроксимации типа (105) видно, что при достаточно малом α решение предельного уравнения (100) или (111) линейно и опять с обеих сторон приближается к точке условного экстремума $\tilde{\mathbf{S}}^*$. Это является основой процесса его численной экстраполяции к $\alpha \rightarrow \pm 0$:

$$\mathbf{x}_S(\alpha) \approx \tilde{\mathbf{S}} + \alpha \cdot \mathbf{u}(\alpha). \quad (470)$$

* * *

Пример. Пусть некоторое вещество получается из n ингредиентов с относительными долями $x_1, x_2, \dots, x_n$ ($x_1 + x_2 + \dots + x_n = 1, x_i \geq 0$). Данная область ограничения есть правильный $(n - 1)$ -симплекс, или гиперсимплекс с рёбрами, равными $\sqrt{2}$, опирающимися на вершины единичных орт. При варьировании долей ингредиентов некое полезное свойство вещества y приобретает максимальное значение при $\mathbf{x} = \mathbf{s}$. Применяя проективный подход (§§ 2.2, 5.12.1), осуществляем движение к точке максимума $\mathbf{s}$ целевой функции $y(\mathbf{x})$ по директивному вектору условного градиента:

$$[dy/d\mathbf{x}] = dy/d\mathbf{x} - [(dy/dx_1 + dy/dx_2 + \dots + dy/dx_n)/n] \cdot (1, 1, \dots, 1).$$

(При $n = 2$ имеем одномерную оптимизацию вдоль гипотенузы.)

Глава 6. Планово-вычислительная n -мерная оптимизация по минимальным планам

§ 6.1. Общие положения

Планово-вычислительные методы n -мерной оптимизации целевой функции целесообразно применять тогда, когда вычисление её тензор-производных (градиента и матрицы Гессе) весьма сложно и трудоёмко. Значительно легче, используя некий план, найти их разностные оценки, применяя степенные *разностные модели целевой функции 1-го или 2-го квази порядка*, т. е. в виде суммы членов её степенного ряда Тейлора до 1-го или 2-го порядка включительно. Но в нём все частные производные, по сути, оцениваются как отношения методом конечных разностей!

В изучаемых далее методах оптимизации исходно используют лишь значения целевой функции в точках по некоему *минимальному плану* в $\mathcal{A}^n$ (где количество точек равно количеству коэффициентов модели). Затем на их основе аппроксимируют k -й директивный вектор в исходной точке $\mathbf{c}_k$. По директивному вектору из точки $\mathbf{c}_k$ осуществляют движение к промежуточному экстремуму в точке $\mathbf{c}_{k+1}$. Далее она становится новой исходной точкой. Процедуру повторяют до выполнения неравенства (449), что отвечает оценке точки стационарности $\mathbf{s}$ (где градиент обнуляется). Если необходимо, в ней вычисляют разностную оценку матрицы Гессе по новому плану для описания функции в окрестности точки экстремума. Эта картина на моделях 1-го и 2-го порядка приведена в §§ 5.8 — 5.10.

Разностные оценки, наряду с ошибкой округления, дополнительно содержат и *систематическую* ошибку аппроксимации. В зависимости от значений тензор-производных $y(\mathbf{x})$ порядка 2, 3, ... и плана её вклад может быть весьма существенным. Поэтому для снижения зависимости от такой малопредсказуемой ошибки движение к экстремуму функции в плановых методах 2-го порядка целесообразно осуществлять только по траекториям директивных векторов. На каждом его этапе реализуется одномерная численная оптимизация, но в определённом направлении. Её возможные процедуры — те же самые, что и вообще в одномерной численной оптимизации: или пошаговые, или поинтервальные (§ 5.2).

§ 6.2. Критерий адекватности плано-разностных моделей

Методы оптимизации, использующие разностные полиномиальные модели целевых функций, — это часть *аппроксимационного анализа*. В данном анализе по аналогии конструируются численные процедуры, базирующиеся на соответствующих функциональных процедурах, но теперь с использованием исходно лишь значений целевой функции. В n -мерном варианте эффективность подобных процедур значительно зависит от выбора *плана вычислений*, т. е. положения его точек $\mathbf{x}_q$ в $\mathcal{A}^n$, а также от размеров интервалов варьирования частных переменных x_i , которые должны быть достаточно малы (а точки близки к центру плана). Отсюда неизбежно встаёт вопрос о наиболее рациональном критерии для оценки адекватности целевой функции и её *разностной модели*.

Пусть в координатном пространстве $\mathcal{A}^n$ на окрестности точки $\mathbf{c}_k$ составлен некий *план вычислений*, состоящий из множества точек $\langle \mathbf{x}_q \rangle$, где $q = 0, \dots, N - 1$; N — количество точек в плане. Соответственно $\langle y_q \rangle$ есть полное множество значений целевой функции в точках плана. Необходимо тут каким-то корректным образом оценивать адекватность целевой функции $y(\mathbf{x})$ и получаемой на основе значений $\langle \mathbf{x}_q \rangle$ и $\langle y_q \rangle$ её разностной полиномиальной модели $\hat{y}(\mathbf{x})$. Критерием адекватности может являться *относительная ошибка модели на области плана*:

$$\xi = \frac{\sqrt{\frac{1}{N} \cdot \sum_{q=0}^{N-1} (y_q - \hat{y}_q)^2}}{y_{\max} - y_{\min}} \cdot 100\% \quad (471)$$

где y_q и $\hat{y}_q$ — значения функции и разностной модели в точках плана (они могут совпадать в некоторых точках, например, иногда в нулевой), $y_{\max}$ и $y_{\min}$ — максимальное и минимальное значения целевой функции на точках плана, оценивающие размах выборки $\langle y_q \rangle$. Принимаем, что при $\xi \leq 5\%$ модель адекватная и её можно применять для оптимизации целевой функции. Если $\xi > 5\%$, то, напротив, принимаем, что модель неадекватная. Тогда необходимо уменьшить интервалы варьирования, хотя бы для некоторых частных переменных x_i . Причём сделать отбор частных переменных x_i можно через анализ вкладов частных слагаемых в сумму квадратов под корнем в (471).

Заметим, что аналогичный критерий с порогом 5% применим для оценки адекватности целевой функции и её *регрессионной модели* в планировании эксперимента — см. § 7.4. Обычно уровень 5% для допустимой ошибки или обратно — уровень для надёжности 95% считаются приемлемыми в процессах плановой оптимизации [15].

§ 6.3. Вычислительная оптимизация по плану квадратный n -симплекс с разностной моделью 1-го порядка

Планово-вычислительная градиентная оптимизация в координатном пространстве $\mathcal{A}^n$ базируется на аппроксимации целевой функции $y(\mathbf{x})$ разностной моделью 1-го порядка с её коэффициентами, полученными в виде разностных оценок в точке $\mathbf{c}$, причём $g_0 = y(\mathbf{c})$:

$$y(\mathbf{x}) \approx \hat{y}(\mathbf{x}) = g_0 + \sum_{i=1}^n \hat{g}_i \cdot (x_i - c_i) = y(\mathbf{c}) + \hat{\mathbf{g}} \cdot (\mathbf{x} - \mathbf{c}), \quad (472)$$

где $\hat{g}_i = \frac{\Delta y}{\Delta x_i}$ — элементы $1 \times n$ -вектора разностного градиента $\hat{\mathbf{g}}$.

Самый простой и при этом *минимальный* план вычисления значений функции $y(\mathbf{x})$ для реализации метода — это *квадратный n -симплекс*. Все точки данного плана расположены в пределах одного квадранта координатного пространства, — как правило, 1-го. В принципе, когда уже имеется информация о направлении движения к оптимуму, может выбираться и более подходящий квадрант для (анти)градиента. Схема метода условно отображена на рис. 13 при $n = 2$ в декартовом базисе:

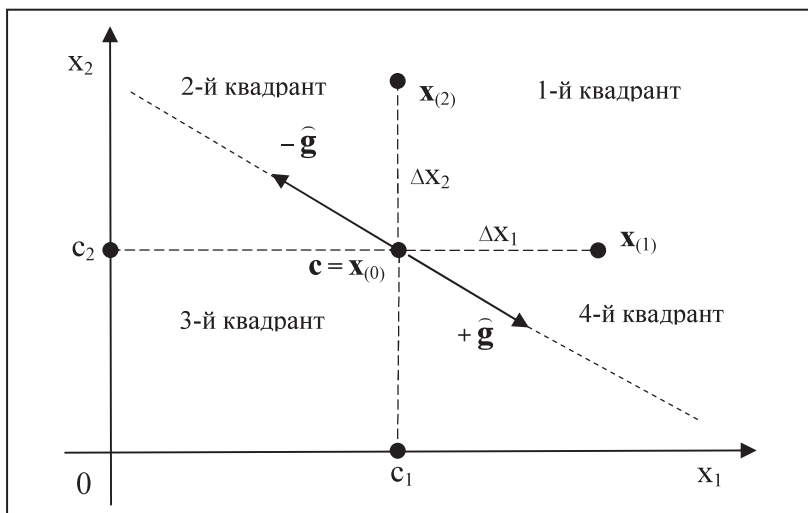


Рис. 13. Поиск максимума или минимума целевой функции $y(\mathbf{x})$ при $n = 2$ по плану квадратный n -симплекс в 1-м квадранте.

Из примера на рис. 13 видно следующее. Несмотря на то, что точки исходного плана находятся формально в 1-м квадранте, градиент $\hat{\mathbf{g}}$ при $n \geq 2$ может оказаться в любом из 2^n квадрантов. При этом вектор антиградиента $-\hat{\mathbf{g}}$ окажется ему зеркально симметричен относительно точки $\mathbf{c}$. Конкретно в примере, данном на рис. 13, векторы градиента и антиградиента расположены в 4-м и 2-м квадрантах. Далее имеем:

$\mathbf{x}_{(0)} = \mathbf{c} = (c_1, c_2)$, $\mathbf{x}_{(1)} = (c_1 + \Delta x_1, c_2)$, $\mathbf{x}_{(2)} = (c_1, c_2 + \Delta x_2)$ — план;
 Δx_1 и Δx_2 — частные вариации переменных x_i ;

$y(\mathbf{x}_{(0)}) = y_0$, $y(\mathbf{x}_{(1)}) = y_1$, $y(\mathbf{x}_{(2)}) = y_2$ — значения функции в точках плана.

Компоненты вектора градиента в примере выражаются так:

$$\hat{g}_1 = \frac{y_1 - y_0}{\Delta x_1}, \quad \hat{g}_2 = \frac{y_2 - y_0}{\Delta x_2}, \quad \hat{\mathbf{g}} = (\hat{g}_1, \hat{g}_2).$$

План содержит 3 точки. В общем случае при $n \geq 2$ количество точек в плане квадратный n -симплекс составляет $N_{\min} = n + 1 = K + 1$, т. е. равно количеству коэффициентов. Поэтому план является *минимальным*.

Компоненты директивного вектора градиента оцениваются так:

$$\hat{g}_i = \frac{y_i - y_0}{\Delta x_i}, \quad \hat{\mathbf{g}} = (\hat{g}_i). \quad (473)$$

Движению по вектору разностного градиента $\hat{\mathbf{g}}$ из начальной точки $\mathbf{c}$ соответствует *метод крутого восхождения*, а движению по вектору разностного антиградиента $-\hat{\mathbf{g}}$ из начальной точки $\mathbf{c}$ соответствует *метод скорейшего спуска* (§ 5.8), причём в их планово-разностных модификациях. Все ранее указанные недостатки градиентного метода остаются в силе и в данных модификациях, усугубляясь дополнительно систематическими ошибками, заложенными в разностной модели (472). Главный недостаток — это в той или иной степени *овражный* характер сходимости процесса к экстремуму, проявляемый из-за произвольности выбираемых масштабов по осям координат (например, для физических переменных — их размерностей). Кроме того, градиентный метод не даёт какой-то оценки искривления *поверхности отклика* $y(\mathbf{x})$ в $\mathcal{A}^{n+1}$. Эти недостатки преодолеваются в масштабно-градиентном методе.

Отметим, что выбранный выше в $\mathcal{A}^n$ в качестве плана вычислений функции отклика и её градиента квадрантный n -симплекс на множестве всевозможных n -симплексов является наиболее простым — с точки зрения как структуры этого плана вычислений, так и его реализации. Разумеется, он только чисто иллюстративно отображается в декартовом базисе, хотя, по сути, является аффинным геометрическим объектом.

§ 6.4. Вычислительная оптимизация по плану осевой n-крест с разностной моделью неполного 2-го порядка

Планово-вычислительная масштабнo-градиентная оптимизация в координатном пространстве $\mathcal{A}^n$ базируется на аппроксимации целевой функции отклика $y(\mathbf{x})$ разностной моделью неполного 2-го порядка с коэффициентами, полученными в виде разностных оценок в точке $\mathbf{c}$, причём $g_0 = y(\mathbf{c})$:

$$\begin{aligned} y(\mathbf{x}) &\approx \hat{y}(\mathbf{x}) = g_0 + \sum_{i=1}^n \hat{g}_i \cdot (x_i - c_i) + \frac{1}{2} \cdot \sum_{i=1}^n \hat{g}_{ii} \cdot (x_i - c_i)^2 = \\ &= y(\mathbf{c}) + \hat{\mathbf{g}} \cdot (\mathbf{x} - \mathbf{c}) + \frac{1}{2} \cdot (\mathbf{x} - \mathbf{c})' \cdot \hat{\mathbf{D}} \cdot (\mathbf{x} - \mathbf{c}), \end{aligned} \quad (474)$$

где $\hat{g}_i = \frac{\Delta y}{\Delta x_i}$ — элементы разностного $1 \times n$ -вектора градиента $\hat{\mathbf{g}}$,

$\hat{g}_{ii} = \frac{\Delta^2 y}{\Delta x_i^2}$ — элементы разностной $n \times n$ -матрицы Гессе $\hat{\mathbf{D}}$.

Наиболее простой, причём *минимальный* план вычислений значений целевой функции $y(\mathbf{x})$ для реализации метода — это *осевой n-крест*. Все точки данного плана расположены по осям координат в парных направлениях, одна точка находится в центре. Схема метода с планом условно отображена на рис. 14 при $n = 2$ в декартовом базисе.

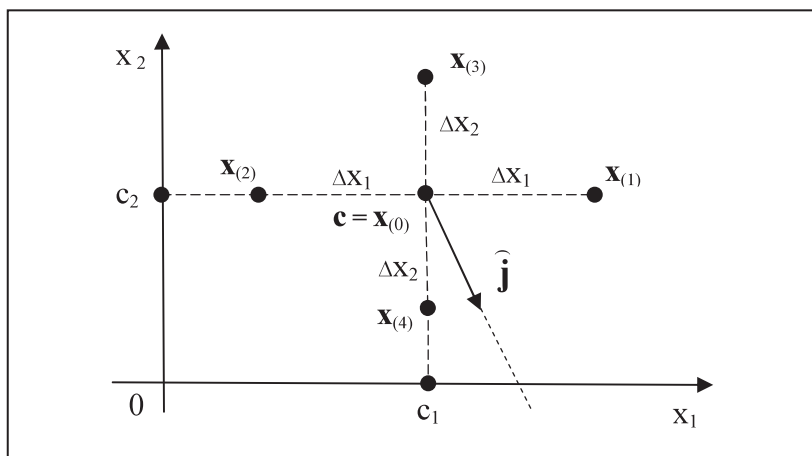


Рис. 14. Поиск максимума или минимума целевой функции $y(\mathbf{x})$ при $n = 2$ по плану осевой n -крест.

Обратим внимание на то, что директивный вектор неполного 2-го порядка $\hat{\mathbf{j}}$ находится всегда в тех квадрантах, что и векторы градиента $\hat{\mathbf{g}}$ или антиградиента $-\hat{\mathbf{g}}$, что обусловлено их чисто масштабным характером взаимосвязи. (На рис. 14 это условно 4-й квадрант.)

Конкретно в примере на рис. 14 имеем:

$$\mathbf{x}_{(0)} = \mathbf{c} = (c_1, c_2), \mathbf{x}_{(1)} = (c_1 + \Delta x_1, c_2), \mathbf{x}_{(2)} = (c_1 - \Delta x_1, c_2),$$

$$\mathbf{x}_{(3)} = (c_1, c_2 + \Delta x_2), \mathbf{x}_{(4)} = (c_1, c_2 - \Delta x_2) \text{ — план;}$$

Δx_1 и Δx_2 — частные вариации переменных x_i ;

$y(\mathbf{x}_{(0)}) = y_0, y(\mathbf{x}_{(1)}) = y_1, y(\mathbf{x}_{(2)}) = y_2, y(\mathbf{x}_{(3)}) = y_3, y(\mathbf{x}_{(4)}) = y_4$ — значения целевой функции в точках плана.

Компоненты вектора градиента и матрицы Гессе оцениваются так:

$$\left. \begin{aligned} \hat{g}_1 &= \frac{y_1 - y_2}{2 \cdot \Delta x_1}, \quad \hat{g}_2 = \frac{y_3 - y_4}{2 \cdot \Delta x_2}, \quad \hat{\mathbf{g}} = (\hat{g}_1, \hat{g}_2); \\ \hat{g}_{11} &= \frac{y_1 + y_2 - 2 \cdot y_0}{\Delta x_1^2}, \quad \hat{g}_{22} = \frac{y_3 + y_4 - 2 \cdot y_0}{\Delta x_2^2}, \quad \hat{\mathbf{D}} = \begin{Bmatrix} \hat{g}_{11} & 0 \\ 0 & \hat{g}_{22} \end{Bmatrix}. \end{aligned} \right\}$$

План на рис. 14 содержит 5 точек. В общем случае при $n \geq 1$ количество точек в плане осевой n-крест составляет $N = 2n + 1 = K + 1 = N_{\min}$. Следовательно, этот план является минимальным для модели (474). При $n \geq 1$ по первым слева формулам в (475) вычисляют последовательно $\hat{g}_i$ и $\hat{g}_{ii}$ вдоль каждой i-й оси координат:

$$\left. \begin{aligned} \hat{g}_i &= \frac{y_i - y_{i+1}}{2 \cdot \Delta x_i}, \quad \hat{\mathbf{g}} = (\hat{g}_i); \\ \hat{g}_{ii} &= \frac{y_i + y_{i+1} - 2 \cdot y_0}{\Delta x_i^2}, \quad \hat{\mathbf{D}} = \{\hat{g}_{ii}\}. \end{aligned} \right\} \quad (475)$$

В частности, при $n = 1$ процедура поиска сводится к *одномерному планированию* 2-го порядка, т. е. к разностному методу ньютоновского типа (§ 5.5).

Пошаговое движение к промежуточному или конечному экстремуму из точки $\mathbf{c}$ осуществляют при $n \geq 2$ по директивному вектору:

$$\hat{\mathbf{j}} = (-\hat{\mathbf{D}})^{-1} \cdot \hat{\mathbf{g}}' = - \begin{vmatrix} \hat{g}_1 / \hat{g}_{11} \\ \dots\dots\dots \\ \hat{g}_n / \hat{g}_{nn} \end{vmatrix}. \quad (476)$$

Этот вектор и радиус-вектор $\mathbf{x}$ по компонентам имеют одни и те же размерности. При изменениях масштабов по осям они преобразуются ковариантно. Формально масштабы не влияют на скорость сходимости.

Особо тут отметим то обстоятельство, что укрупнением масштабов по осям x_i можно при необходимости добиваться пропорционального уменьшения отношений $\hat{g}_i / \hat{g}_{ii}$ в (476), что важно, если знаменатели $\hat{g}_{ii}$ слишком малы и не точны. Тогда знаменатели $\hat{g}_{ii}$ увеличиваются на порядок больше, нежели числители $\hat{g}_i$.

Весьма наглядным признаком движения к экстремуму функции $y(\mathbf{x})$ является один знак для всех $\hat{g}_{ii} \neq 0$. (Но разные знаки хотя бы для пары $\hat{g}_{ii}$ свидетельствуют о локальной седловине целевой функции.) Причём с приближением к искомому экстремуму $y(\mathbf{x})$ числители $\hat{g}_i$ неизбежно с замедлением уменьшаются вплоть до выполнения результирующего неравенства типа (449). Хорошим признаком эффективности метода является то, что приближение к экстремуму сопровождается только уменьшением значений всех $\hat{g}_i$, а все $\hat{g}_{ii}$ при этом весьма значимые.

Если матрица $\hat{D}(\mathbf{x})$ в окрестности строгого экстремума $y(\mathbf{x})$ плохо обусловлена, т. е. при $\hat{\mathbf{g}} \rightarrow 0$ некоторые $\hat{g}_{ii} \approx 0$ (несмотря на укрупнение масштабов), или в окрестности нестрогого экстремума $y(\mathbf{x})$ вырождена, т. е. при $\hat{\mathbf{g}} \rightarrow 0$ некоторые $\hat{g}_{ii} = 0$, то данный метод в базовом варианте неэффективен, так как приводит к слишком большому разбросу результата из-за неизбежных ошибок округления и аппроксимации в значениях целевой функции. Однако действенным способом разрешения этой проблемы может быть применение метода квадратичной регуляризации по Тихонову (см. § 5.10).

В качестве параметра квадратичной регуляризации, например, здесь можно выбрать $\alpha = \max \{ \hat{g}_i / \Delta x_i \} \rightarrow 0$. Тогда в итоге имеем довольно устойчивую оценку:

$$\hat{\mathbf{j}}_\alpha = -(\hat{D} \pm \alpha I)^{-1} \cdot \hat{\mathbf{g}}', \quad (477)$$

согласованную с элементами $\hat{g}_{ii}$ по размерности и убывающую по величине по мере приближения к экстремуму функции. В таком случае благодаря приобретаемой устойчивости пошаговой вычислительной оптимизации соответственно резко увеличивается скорость сходимости процедуры к экстремуму целевой функции. Хотя сам её экстремум может быть иногда недостаточно строгим, а то и вообще нестрогим. Главное, тут находится такое сочетание значений частных переменных x_i , при котором целевая функция имеет экстремальное значение. (Масштабы по осям координат x_i при регуляризации могут оказывать влияние лишь на находимые значения переменных, но не целевой функции.)

§ 6.5. Вычислительная оптимизация по плану n-АКП с разностной моделью 2-го порядка

Плано-вычислительная оптимизация общеньютоновского типа в координатном пространстве $\mathcal{A}^n$ базируется на аппроксимации целевой функции разностной моделью полного 2-го порядка с коэффициентами, полученными в виде разностных оценок в точке $\mathbf{c}$, причём $g_0 = y(\mathbf{c})$:

$$\begin{aligned} y(\mathbf{x}) &\approx \hat{y}(\mathbf{x}) = g_0 + \sum_{i=1}^n \hat{g}_i \cdot (x_i - c_i) + \frac{1}{2} \cdot \sum_{j=1}^n \sum_{i=1}^n \hat{g}_{ij} \cdot (x_i - c_i)(x_j - c_j) = \\ &= y(\mathbf{c}) + \hat{\mathbf{g}} \cdot (\mathbf{x} - \mathbf{c}) + \frac{1}{2} \cdot (\mathbf{x} - \mathbf{c})' \cdot \hat{\mathbf{G}} \cdot (\mathbf{x} - \mathbf{c}), \end{aligned} \quad (478)$$

где $\hat{g}_i = \frac{\Delta y}{\Delta X_i}$ — элементы разностного $1 \times n$ -вектора градиента $\hat{\mathbf{g}}$,

$\hat{g}_{ii} = \frac{\Delta^2 y}{\Delta X_i^2}$, $\hat{g}_{ij} = \frac{\Delta^2 y}{\Delta X_i \Delta X_j}$ — элементы разностной $n \times n$ -матрицы Гессе $\hat{\mathbf{G}}$.

Наиболее простой, причём *минимальный* план вычислений значений целевой функции $y(\mathbf{x})$ для реализации метода с моделью типа (478) — это n -АКП (*асимметричный композиционный план*). Данный план есть композиция из осевого плана n -крест и его угловых точек в каком-то одном квадранте каждой координатной плоскости $\langle x_i, x_j \rangle$. Схема метода с планом условно показана на рис. 15 при $n = 2$ в декартовом базисе.

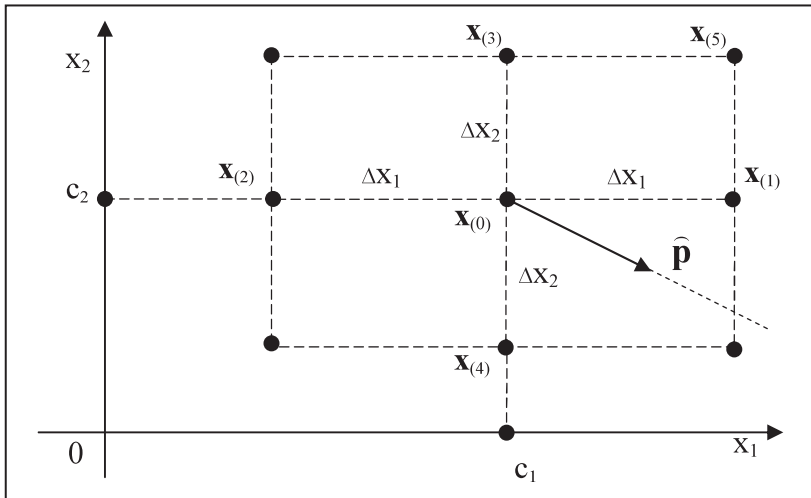


Рис. 15. Поиск максимума или минимума целевой функции $y(\mathbf{x})$ при $n = 2$ по асимметричному композиционному плану.

Директивный вектор 2-го порядка $\hat{\mathbf{p}}$, хотя и не обязательно должен находиться в том же квадранте, что и векторы $\hat{\mathbf{g}}$ или $-\hat{\mathbf{g}}$, но вероятнее находится в нём (или, по крайней мере, в смежном), что обусловлено их знакоопределённым симметрично-линейным характером взаимосвязи. Отсюда целесообразен выбор *градиентного* при поиске максимума или *антиградиентного* при поиске минимума квадранта для оценки $\hat{\mathbf{g}}_{ij}$!

Конкретно в примере при $n = 2$ на рис. 15 имеем:

$\mathbf{x}_{(0)} = \mathbf{c} = (c_1, c_2)$, $\mathbf{x}_{(1)} = (c_1 + \Delta x_1, c_2)$, $\mathbf{x}_{(2)} = (c_1 - \Delta x_1, c_2)$,
 $\mathbf{x}_{(3)} = (c_1, c_2 + \Delta x_2)$, $\mathbf{x}_{(4)} = (c_1, c_2 - \Delta x_2)$, $\mathbf{x}_{(5)} = (c_1 + \Delta x_1, c_2 + \Delta x_2)$ — план;
 Δx_1 и Δx_2 — частные вариации переменных x_i ;
 $y(\mathbf{x}_{(0)}) = y_0$, $y(\mathbf{x}_{(1)}) = y_1$, $y(\mathbf{x}_{(2)}) = y_2$, $y(\mathbf{x}_{(3)}) = y_3$, $y(\mathbf{x}_{(4)}) = y_4$, $y(\mathbf{x}_{(5)}) = y_5$ — значения целевой функции в точках плана.

Компоненты вектора градиента и матрицы Гессе выражаются так:

$$\left. \begin{aligned} \hat{g}_1 &= \frac{y_1 - y_2}{2 \cdot \Delta x_1}, \quad \hat{g}_2 = \frac{y_3 - y_4}{2 \cdot \Delta x_2}, \quad \hat{\mathbf{g}} = (\hat{g}_1, \hat{g}_2); \\ \hat{g}_{11} &= \frac{y_1 + y_2 - 2 \cdot y_0}{\Delta x_1^2}, \quad \hat{g}_{22} = \frac{y_3 + y_4 - 2 \cdot y_0}{\Delta x_2^2}, \\ \hat{g}_{12} &= \hat{g}_{21} = \frac{y_0 - y_1 + y_5 - y_3}{\Delta x_1 \cdot \Delta x_2}, \quad \hat{\mathbf{G}} = \begin{vmatrix} \hat{g}_{11} & \hat{g}_{12} \\ \hat{g}_{21} & \hat{g}_{22} \end{vmatrix}. \end{aligned} \right\} \quad (479)$$

При $n > 2$ для заданных сочетаний координат x_i и x_j применяются те же самые формулы для чистых и смешанных коэффициентов. Количество точек в n -АКП составляет:

$$N = 1 + 2n + C_n^2 = 1 + 2n + \frac{n(n-1)}{2} = N_{\min},$$

следовательно, он *минимален* для модели (478).

Движение к экстремуму осуществляется по директивному вектору:

$$\hat{\mathbf{p}} = (-\hat{\mathbf{G}})^{-1} \cdot \hat{\mathbf{g}}'. \quad (480)$$

Его компоненты имеют те же самые размерности, что и компоненты $\mathbf{x}$. Поэтому при изменении масштабов x_i они преобразуются ковариантно.

Если матрица Гессе $\hat{\mathbf{G}}$ на области данного плана плохо обусловлена (т. е. $\det \hat{\mathbf{G}} \approx 0$), то для однозначности и устойчивости результатов поиска целесообразно использовать метод регуляризации Тихонова (§ 5.10). Как параметр квадратичной регуляризации, например, можно выбрать $\alpha = \max\{\hat{g}_i / \Delta x_i\} \rightarrow 0$. Тогда в итоге имеем устойчивую оценку:

$$\hat{\mathbf{p}}_\alpha = -(\hat{\mathbf{G}} \pm \alpha \mathbf{I})^{-1} \cdot \hat{\mathbf{g}}'. \quad (481)$$

Глава 7. Планово-экспериментальная n-мерная оптимизация по ортогональным планам

§ 7.1. Общие положения

В этой главе рассматриваются основные аспекты и методы планово-экспериментальной n-мерной оптимизации целевой функции $y(\mathbf{x})$, чей аналитический вид априори неизвестен. В литературе она называется *функцией отклика* [46, 47]. Фактически это кибернетическое понятие:



Вся процедура оптимизации традиционно называется *планированием эксперимента*. В ней по неким изначальным логическим соображениям предполагается, что целевая функция y зависит от этих переменных, но степень функционального влияния последних, в том числе совместно, на фоне общей погрешности для y нужно ещё выявить. Что особо важно: отдельные оценки значения целевой функции отклика $\tilde{y}(\mathbf{x})$ при каждом задаваемом наборе *точных* значений переменных $\mathbf{x}_q = (x_1, x_2, \dots, x_n)$ всегда возможно найти *экспериментально*, но с некоторой случайной ошибкой $(\tilde{y}_q - y_q)$. В идеале её математическое ожидание нулевое: $\mu\{\tilde{y}_q - y_q\} = 0$, или $\mu\{\tilde{y}_q\} = y_q$. Поэтому $\mathbf{x} \rightarrow y(\mathbf{x})$ есть однозначное отображение, но $\mathbf{x} \rightarrow \tilde{y}(\mathbf{x})$ есть не вполне однозначное отображение.

В планово-экспериментальной оптимизации особое значение имеет *экономическая* составляющая общей процедуры, так как выполнение опытов или наблюдений является финансово затратным мероприятием. Отсюда имеется стремление свести их общее количество к минимуму! Фактически это означает стремление минимизировать как число точек в *планах эксперимента*, так и общее количество побочных движений к экстремуму по директивным векторам (см. §§ 5.8—5.10). Эти приёмы минимизации всех затрат в условиях производства или весьма дорогих научных исследований относятся к сфере *математической экономики*.

Для дальнейших необходимых статистических оценок принимают, что случайная величина $[\tilde{y}(\mathbf{x}) - y(\mathbf{x})]$ имеет *центральное однородное нормальное* распределение Гаусса; т. е. $\mu\{\tilde{y} - y\} = 0$, $\sigma\{\tilde{y}\} = \text{const}$. (Параметры однородного распределения не зависят от $\mathbf{x}_q$.) График для функции плотности вероятности при нормальном распределении $(\tilde{y} - y)$ известен как *кривая ошибок*. График кривой имеет характерную форму колокола [15, т.1]. Это центральное распределение широко применяется в статистических методах, в частности, в планировании эксперимента.

Исходя из априорной неизвестности аналитического вида целевой функции типа $y(\mathbf{x})$ в планово-экспериментальных методах оптимизации используют ряд её аппроксимирующих *разностных* моделей (см. гл. 6), как правило, 2-го порядка (неполного и полного). Эту модель находят на области в $\mathcal{A}^n$ по очередному плану эксперимента с центром в точке $\mathbf{c}_k$, причём с привлечением *множественной линейной регрессии* ($n > 1$). Координатное аффинное пространство при какой-либо геометрической необходимости вводом евклидовой метрики трансформируется в $\mathcal{E}^n$.

Ввиду аппроксимации целевой функции разностно-регрессионным полиномом 2-го порядка, в её статистической модели содержится как *систематическая*, так и *случайная ошибка*. В свою очередь, её *ошибка округления* как бы входит незначительной составляющей в случайную ошибку (когда само это округление выполняют, с учётом принятого правила Гаусса). Из модели вычисляют соответствующий директивный вектор. По приближённому k -му *директивному вектору* из центральной точки плана $\mathbf{c}_k$ осуществляют движение к промежуточному экстремуму той же функции отклика в точке $\mathbf{c}_{k+1}$. Затем всю процедуру повторяют вплоть до соблюдения в какой-то точке $\mathbf{s}^*$ неравенства типа (449). Её принимают за точку стационарности $y(\mathbf{x})$. В точке стационарности $\mathbf{s}^*$ вычисляют полную регрессионную оценку матрицы Гессе, по которой и судят о характере экстремума или в итоге подтверждают его. Заодно с применением этой оценки матрицы Гессе выполняется аппроксимация 2-го порядка целевой функции $y(\mathbf{x})$ в окрестности точки экстремума $\mathbf{s}^*$. В общих чертах это было продемонстрировано на моделях 1-го и 2-го порядка в §§ 6.2 — 6.4. Поскольку используемые модели содержат неустранимые ошибки 2-х видов в сравнении с априори точной целевой функцией, то они должны проверяться каждый раз на *адекватность* по отношению к функции отклика на области плана эксперимента.

Все допускаемые статистические гипотезы, в принципе, проверяемы с принимаемым уровнем *доверительной вероятности* с использованием стандартных статистических критериев и/или с ограничением сверху уровня допустимых *относительных ошибок*.

Ввиду того что в этой главе для разнообразных оценок довольно широко применяются методы теории вероятностей и математической статистики [29, 44], то необходимо кратко рассмотреть ряд исходных базовых понятий, но применительно к оценкам функции отклика $\tilde{y}(\mathbf{x})$. Эти понятия определяются или в какой-то конкретной точке $\mathbf{x}_q$ или на множестве точек $\langle \mathbf{x}_q \rangle$, например, на области плана эксперимента.

1. Математическое ожидание. (Момент 1-го порядка.)

$$\mu\{\tilde{y}_q\} = y_q, \quad \mu\{\tilde{y}_q - y_q\} = 0, \quad \text{где } y_q = y(\mathbf{x}_q), \quad \tilde{y}_q = \tilde{y}(\mathbf{x}_q).$$

Здесь и далее принимается, что все частные значения $\mathbf{x}_q$ задаются точно. (Нужная точность обеспечивается при выполнении экспериментов.)

Полезные свойства:

- а) математическое ожидание алгебраической суммы случайных величин равно той же алгебраической сумме их математических ожиданий;
- б) математическое ожидание произведения неслучайной величины на случайную величину равно произведению неслучайной величины на математическое ожидание случайной величины.

2. Выборочное среднее.

$$\bar{\tilde{y}}_q = \frac{1}{m} \cdot \sum_{j=1}^m \tilde{y}_{q,j}, \quad \text{где } \tilde{y}_{q,j} \text{ — статистически независимые значения}$$

функции $\tilde{y}$ в точке $\mathbf{x}_q$. По сути, эта характеристика есть среднее арифметическое для элементов независимой выборки $\langle \tilde{y}_{q,j} \rangle$ порядка m , где $j = \overline{1, m}$. Очевидно: $\mu\{\bar{\tilde{y}}_q\} = \mu\{\tilde{y}_q\} = y_q$.

3. Дисперсия случайной величины. (Момент 2-го порядка.)

$$\begin{aligned} \sigma^2\{\tilde{y}_q\} &= \sigma^2\{\tilde{y}_q - y_q\} = \mu\left\{\left[\tilde{y}_q - \mu\{\tilde{y}_q\}\right]^2\right\} = \mu\left\{\left(\tilde{y}_q - y_q\right)^2\right\}, \\ \sigma^2\{\bar{\tilde{y}}_q\} &= \sigma^2\{\bar{\tilde{y}}_q - y_q\} = \mu\left\{\left[\bar{\tilde{y}}_q - \mu\{\bar{\tilde{y}}_q\}\right]^2\right\} = \mu\left\{\left(\bar{\tilde{y}}_q - y_q\right)^2\right\} = \\ &= \mu\left\{\left[\frac{1}{m} \cdot \sum_{j=1}^m \tilde{y}_{q,j} - y_q\right]^2\right\} = \frac{1}{m^2} \cdot \mu\left\{\left[\sum_{j=1}^m (\tilde{y}_{q,j} - y_q)\right]^2\right\} = \\ &= \frac{1}{m^2} \cdot \mu\left\{\sum_{j=1}^m (\tilde{y}_{q,j} - y_q)^2 + \sum_{i=1}^m \sum_{j=1, j \neq i}^m (\tilde{y}_{q,j} - y_q) \cdot (\tilde{y}_{q,i} - y_q)\right\} = \\ &= \frac{1}{m^2} \cdot \left[m \cdot \sigma^2\{\tilde{y}_q\} + 0\right] = \sigma^2\{\tilde{y}_q\} / m = \sigma^2\{\tilde{y}_q - y_q\} / m. \end{aligned} \quad (482)$$

4. Выборочная смещённая дисперсия.

$$s^{*2} \{ \tilde{y} \} = \frac{1}{m} \cdot \sum_{j=1}^m \left(\tilde{y}_{q,j} - \bar{\tilde{y}}_q \right)^2 = \overline{\left(\tilde{y}_q - \bar{\tilde{y}}_q \right)^2} = \overline{\tilde{y}_q^2} - \bar{\tilde{y}}_q^2. \quad (483)$$

Её математическое ожидание составляет

$$\begin{aligned} \mu \{ s^{*2} \{ \tilde{y} \} \} &= \frac{1}{m} \cdot \sum_{j=1}^m \mu \left\{ \left(\tilde{y}_{q,j} - \frac{1}{m} \cdot \sum_{i=1}^m \tilde{y}_{q,i} \right)^2 \right\} = \\ &= \frac{1}{m} \cdot \sum_{j=1}^m \mu \left\{ \left[\left(\tilde{y}_{q,j} - y_q \right) - \left(\frac{1}{m} \cdot \sum_{i=1}^m \tilde{y}_{q,i} - y_q \right) \right]^2 \right\} = \\ &= \frac{1}{m^3} \cdot \sum_{j=1}^m \mu \left\{ \left[m \cdot \left(\tilde{y}_{q,j} - y_q \right) - \sum_{i=1}^m \left(\tilde{y}_{q,i} - y_q \right) \right]^2 \right\} = \\ &= \frac{1}{m^3} \cdot \sum_{j=1}^m \mu \left\{ m^2 \cdot \left(\tilde{y}_{q,j} - y_q \right)^2 - 2m \cdot \left(\tilde{y}_{q,j} - y_q \right) \cdot \sum_{i=1}^m \left(\tilde{y}_{q,i} - y_q \right) + \right. \\ &\quad \left. + \sum_{i=1}^m \left(\tilde{y}_{q,i} - y_q \right)^2 - \sum_{t=1}^m \sum_{i=1, i \neq t}^m \left(\tilde{y}_{q,t} - y_q \right) \cdot \left(\tilde{y}_{q,i} - y_q \right) \right\} = \\ &= \frac{1}{m^3} \cdot \left[m^3 \cdot \sigma^2 \{ \tilde{y} \} - 2m^2 \cdot \sigma^2 \{ \tilde{y} \} + m^2 \cdot \sigma^2 \{ \tilde{y} \} - 0 \right] = \\ &= \frac{m-1}{m} \cdot \sigma^2 \{ \tilde{y} \} = \mu \left\{ \left(\tilde{y}_{q,j} - \bar{\tilde{y}}_q \right)^2 \right\} = \sigma^2 \{ \tilde{y}_{q,j} - \bar{\tilde{y}}_q \} < \sigma^2 \{ \tilde{y} \}. \quad (484) \end{aligned}$$

5. Выборочная несмещённая дисперсия.

$$s^2 \{ \tilde{y} \} = \frac{1}{m-1} \cdot \sum_{j=1}^m \left(\tilde{y}_{q,j} - \bar{\tilde{y}}_q \right)^2, \quad (485)$$

$$\mu \{ s^2 \{ \tilde{y} \} \} = \frac{m}{m-1} \cdot s^{*2} \{ \tilde{y} \} = \sigma^2 \{ \tilde{y} \}. \quad (486)$$

В свою очередь, характеристики $\sigma \{ \tilde{y} \}$ или $s \{ \tilde{y} \}$ обычно определяются как *среднее квадратичное отклонение*, т. е. истинное или выборочное. Первое из них, в сущности, абстрактное. Второе из них весьма широко применяется для оценки вероятных интервалов случайной величины.

Выше в пунктах 3 и 4 встречается ряд математических ожиданий для произведений двух центрированных случайных величин. Подобная характеристика определяется как *ковариация* этих случайных величин. Поскольку центрированные величины *функционально независимы*, а их ошибки тут *статистически независимы*, то их ковариация нулевая.

6. Ковариация случайной величины и её выборочного среднего.

$$\begin{aligned}
v\{\tilde{y}_{q,i} - \bar{\tilde{y}}_q\} &= \mu\left\{\left(\tilde{y}_{q,i} - y_q\right) \cdot \left(\bar{\tilde{y}}_q - y_q\right)\right\} = \\
&= \mu\left\{\left(\tilde{y}_{q,i} - y_q\right) \cdot \left(\frac{1}{m} \cdot \sum_{j=1}^m \tilde{y}_{q,j} - y_q\right)\right\} = \frac{1}{m} \cdot \mu\left\{\left(\tilde{y}_{q,i} - y_q\right) \cdot \sum_{j=1}^m \left(\tilde{y}_{q,j} - y_q\right)\right\} = \\
&= \frac{1}{m} \cdot \mu\left\{\left(\tilde{y}_{q,i} - y_q\right)^2\right\} = \sigma^2\{\tilde{y}\} / m = \sigma^2\{\bar{\tilde{y}}_q - y_q\} = \sigma^2\{\bar{\tilde{y}}_q\}. \quad (487)
\end{aligned}$$

7. Математическое ожидание суммы выборочных ковариаций.

$$\begin{aligned}
\mu\left\{\sum_{j=1}^m \sum_{i=1}^m \substack{i \neq j} (\tilde{y}_{q,j} - \bar{\tilde{y}}_q) \cdot (\tilde{y}_{q,i} - \bar{\tilde{y}}_q)\right\} = \\
= \sum_{j=1}^m \sum_{i=1}^m \substack{i \neq j} \mu\left\{(\tilde{y}_{q,j} - \bar{\tilde{y}}_q) \cdot (\tilde{y}_{q,i} - \bar{\tilde{y}}_q)\right\} = 0. \quad (488)
\end{aligned}$$

8. Абсолютная ошибка. (Абсолютный момент 1-го порядка.)

$$\delta\{\tilde{y}\} = \mu\left\{\left|\tilde{y}_{q,j} - y_q\right|\right\} \text{— см. также далее § 7.3.}$$

Выборочная несмещённая абсолютная ошибка.

$$d\{\tilde{y}\} = \frac{1}{\sqrt{m(m-1)}} \cdot \sum_{j=1}^m \left|\tilde{y}_{q,j} - \bar{\tilde{y}}_q\right| = \frac{1}{\sqrt{m(m-1)}} \cdot \overline{\left|\tilde{y}_{q,j} - \bar{\tilde{y}}_q\right|} \quad (489)$$

Для нормального закона распределения плотности вероятности Гаусса её математическое ожидание есть величина, независящая от m ,

$$\mu\{d\{\tilde{y}\}\} = \delta\{\tilde{y}\} = \sqrt{2/\pi} \cdot \sigma\{\tilde{y}\} < \sigma\{\tilde{y}\}, \quad (490)$$

математическое ожидание выборочной смещённой абсолютной ошибки

$$\mu\left\{\left|\tilde{y}_{q,j} - \bar{\tilde{y}}_q\right|\right\} = \mu\left\{\frac{1}{m} \cdot \sum_{j=1}^m \left|\tilde{y}_{q,j} - \bar{\tilde{y}}_q\right|\right\} = \sqrt{\frac{m-1}{m}} \cdot \delta\{\tilde{y}\} < \delta\{\tilde{y}\}.$$

9. Математическое ожидание разности и дисперсия алгебраической суммы двух тождественно распределённых случайных величин.

$$\mu\left\{\left(\tilde{y}_{q,j} - \tilde{y}_{q,i}\right)\right\} = \mu\left\{\tilde{y}_{q,j}\right\} - \mu\left\{\tilde{y}_{q,i}\right\} = y_q - y_q = 0, \quad (491)$$

$$\begin{aligned}
\sigma^2\left\{\left(\tilde{y}_{q,j} \pm \tilde{y}_{q,i}\right)\right\} &= \mu\left\{\left[\left(\tilde{y}_{q,j} - y_q\right) \pm \left(\tilde{y}_{q,i} - y_q\right)\right]^2\right\} = \mu\left\{\left(\tilde{y}_{q,j} - y_q\right)^2\right\} + \\
&+ \mu\left\{\left(\tilde{y}_{q,i} - y_q\right)^2\right\} \pm 2\mu\left\{\left[\left(\tilde{y}_{q,j} - y_q\right) \cdot \left(\tilde{y}_{q,i} - y_q\right)\right]\right\} = 2\sigma^2\{\tilde{y}_q\}, \quad (492)
\end{aligned}$$

$$\sigma\left\{\left(\tilde{y}_{q,j} \pm \tilde{y}_{q,i}\right)\right\} = \sqrt{2} \cdot \sigma\{\tilde{y}_q\}. \quad (493)$$

§ 7.2. Полифакторная линейная регрессия

В *регрессионном анализе*, как правило, аналитический вид целевых функций неизвестен, — как и в планировании эксперимента вообще. Поэтому прибегают к их полиномиальным разностным моделям (гл. 6).

Выполнить же статистическую оценку точности полиномиальной модели функции отклика в случае нескольких исходных переменных x_i можно весьма корректно, используя *полифакторный* (множественный) регрессионный анализ. Этот анализ базируется на нормальном законе распределения случайной ошибки опыта ($\tilde{y} - y$) при экспериментальном нахождении отдельных значений целевой функции путём измерений или наблюдений. Это существенно расширило те возможности, которые ранее были заложены в базовом методе наименьших квадратов [15].

Основоположителем регрессионного анализа является К. Пирсон. Он внёс значительный вклад в становление, развитие и практическое применение этого важного раздела математической статистики. Ему же принадлежит и термин *нормальное распределение*. В частности, именно благодаря полифакторному регрессионному анализу стало возможным осуществлять достаточно надёжную аппроксимацию целевых функций отклика линейными регрессионными моделями, проверять с заданной доверительной вероятностью ряд попутных статистических гипотез, а также выявлять значимые и незначимые коэффициенты регрессионной модели. В полифакторном регрессионном анализе используют сложные *линейные разностные модели* априори неизвестной целевой функции $\hat{y} = \hat{y}(\mathbf{x})$ и аналогичные им *линейные регрессионные модели* функции отклика $\hat{y} = \hat{y}(\mathbf{x})$ 2-х тождественных друг другу видов $f_1(\mathbf{u}) \equiv f_2(\mathbf{v})$:

$$\hat{y}(\mathbf{x}) = f_1(\mathbf{u}) = b_0 + \mathbf{u}' \cdot \mathbf{b} = b_0 + \mathbf{b}' \cdot \mathbf{u} = b_0 + b_1 \cdot u_1 + b_2 \cdot u_2 + \dots + b_K \cdot u_K, \quad (494)$$

$$\hat{y}(\mathbf{x}) = f_2(\mathbf{v}) = \mathbf{h}' \cdot \mathbf{v} = \mathbf{v}' \cdot \mathbf{h} = b_0 \cdot 1 + b_1 \cdot u_1 + b_2 \cdot u_2 + \dots + b_K \cdot u_K. \quad (495)$$

1-я модель — *традиционная* и аналогична по структуре простой модели, которая применяется в методе наименьших квадратов непосредственно к линейной или линеаризованной целевой функции $y = y(\mathbf{x})$ (см. § 4.7); 2-я модель — *полнолинейная*. Последняя пригодна для оценок сразу всех её коэффициентов и их дисперсионных характеристик.

В принципе, обе данные модели применяются для оценок методом регрессионного анализа всех своих коэффициентов — либо b_0 и $\mathbf{b}$, либо в целом $\mathbf{h}$. Но при этом размерность самого матричного уравнения, из которого в итоге вычисляются векторные коэффициенты, в первом случае значительно меньше, чем во втором, что существенно (см. далее). Поэтому в таком аспекте линейная модель (494) предпочтительней.

Так, для разностной модели целевой функции $y(\mathbf{x})$ 2-го порядка при $n=3$ (см. § 6.5) имеем количество частных переменных факторов $K=9$, а количество коэффициентов регрессии $K+1=10$ (т. е. по максимуму):

$$\begin{aligned} y(\mathbf{x}) &\approx \hat{y} = \hat{y}(\mathbf{x}) = \hat{g}_0 + \hat{g}_1 \cdot x_1 + \hat{g}_2 \cdot x_2 + \hat{g}_3 \cdot x_3 + \frac{1}{2} \cdot \hat{g}_{11} \cdot x_1^2 + \frac{1}{2} \cdot \hat{g}_{22} \cdot x_2^2 + \\ &+ \frac{1}{2} \cdot \hat{g}_{33} \cdot x_3^2 + \hat{g}_{12} \cdot x_1 \cdot x_2 + \hat{g}_{13} \cdot x_1 \cdot x_3 + \hat{g}_{23} \cdot x_2 \cdot x_3 = \hat{f}_1(\mathbf{u}) = \hat{f}_2(\mathbf{v}) = \\ &= b_0 + b_1 \cdot u_1 + b_2 \cdot u_2 + b_3 \cdot u_3 + b_4 \cdot u_4 + b_5 \cdot u_5 + b_6 \cdot u_6 + b_7 \cdot u_7 + b_8 \cdot u_8 + b_9 \cdot u_9 = \\ &= b_0 \cdot u_0 + b_1 \cdot u_1 + b_2 \cdot u_2 + b_3 \cdot u_3 + b_4 \cdot u_4 + b_5 \cdot u_5 + b_6 \cdot u_6 + b_7 \cdot u_7 + b_8 \cdot u_8 + b_9 \cdot u_9, \\ b_0 &= \hat{g}_0, u_0 = 1; b_1 = \hat{g}_1, u_1 = x_1; \dots; b_4 = \frac{1}{2} \cdot \hat{g}_{11}, u_4 = x_1^2; \dots; b_9 = \hat{g}_{23}, u_9 = x_2 \cdot x_3. \end{aligned}$$

Причём традиционно при линейной регрессии $\mathbf{u}$ и $\mathbf{b}$ суть $K \times 1$ -вектор-столбцы, а $\mathbf{v}$ и $\mathbf{h}$ суть $(K+1) \times 1$ -вектор-столбцы. Например, имеем:

$$\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ \dots \\ u_n \\ u_{n+1} \\ \dots \\ u_K \end{pmatrix} \equiv \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \\ x_1 x_2 \\ \dots \\ x_n^2 \end{pmatrix}, \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_n \\ b_{n+1} \\ \dots \\ b_K \end{pmatrix}; \mathbf{v} = \begin{pmatrix} u_0 \\ u_1 \\ u_2 \\ \dots \\ u_n \\ u_{n+1} \\ \dots \\ u_K \end{pmatrix} \equiv \begin{pmatrix} 1 \\ x_1 \\ x_2 \\ \dots \\ x_n \\ x_1 x_2 \\ \dots \\ x_n^2 \end{pmatrix}, \mathbf{h} = \begin{pmatrix} b_0 \\ b_1 \\ b_2 \\ \dots \\ b_n \\ b_{n+1} \\ \dots \\ b_K \end{pmatrix}.$$

В случае исходной разностной модели 1-го порядка отсюда исчезают все смешанные произведения и квадраты переменных ($K = n = 3$); в случае исходной разностной модели неполного 2-го порядка отсюда исчезают все смешанные произведения переменных ($K = 2n = 6$).

Тогда реальной целевой функции $y = y(\mathbf{x})$ отвечают соответственно модели линейной регрессии $\tilde{y}|\mathbf{u}$ и $\tilde{y}|\mathbf{v}$ также 2-х видов:

$$\dot{y}(\mathbf{x}) = \dot{f}_1(\mathbf{u}) = \dot{b}_0 + \mathbf{u}' \cdot \dot{\mathbf{b}} = \dot{b}_0 + \dot{b}_1 \cdot u_1 + \dot{b}_2 \cdot u_2 + \dots + \dot{b}_K \cdot u_K, \quad (496)$$

$$\dot{y}(\mathbf{x}) = \dot{f}_2(\mathbf{v}) = \dot{\mathbf{h}}' \cdot \mathbf{v} = \mathbf{v}' \cdot \dot{\mathbf{h}} = \dot{b}_0 \cdot 1 + \dot{b}_1 \cdot u_1 + \dot{b}_2 \cdot u_2 + \dots + \dot{b}_K \cdot u_K. \quad (497)$$

Тут $\dot{b}_0$, $\dot{\mathbf{b}}$ и $\dot{\mathbf{h}}$ — коэффициенты 1-й и 2-й модели линейной регрессии.

Обе модели тождественны при централизованной регрессии, т. е. $b_0 = 0$.

$\mathbf{u} = (u_1, u_2, \dots, u_K)'$ — $K \times 1$ -радиус-вектор в факторном пространстве Φ^K ;

$\mathbf{v} = \begin{pmatrix} 1 \\ \mathbf{u} \end{pmatrix}$ — $(K+1) \times 1$ -радиус-вектор в факторном пространстве Φ^{K+1} .

* * *

Рассмотрим вначале линейную регрессию $\tilde{y}|\mathbf{u}$, базирующуюся на традиционных моделях типа (494), (496), используемых в классическом методе наименьших квадратов Гаусса ($n > 1$). Поэтому здесь возможно применять результаты, полученные ранее по этому методу в § 4.7 для линейной целевой функции y от векторного переменного $\mathbf{x}$ (410)–(412). Например, аналогом (415), (416) является система из 2-х усреднённых уравнений с двумя неизвестными $\dot{\mathbf{b}}$ и $\dot{b}_0$:

$$\tilde{y} = \overline{\mathbf{u}'} \cdot \dot{\mathbf{b}} + \dot{b}_0, \quad (498)$$

$$\overline{\tilde{y} \cdot \mathbf{u}} = \overline{\mathbf{u}} \cdot \dot{\mathbf{b}}_0 + \overline{\mathbf{u} \cdot \mathbf{u}'} \cdot \dot{\mathbf{b}}. \quad (499)$$

Выразив из первого уравнения $\dot{b}_0$ и подставив это значение во второе уравнение, получаем уравнение относительно $\dot{\mathbf{b}}$:

$$\left(\overline{\mathbf{u} \cdot \mathbf{u}'} - \overline{\mathbf{u}} \cdot \overline{\mathbf{u}'} \right) \cdot \dot{\mathbf{b}} = \overline{\tilde{y} \cdot \mathbf{u}} - \tilde{y} \cdot \overline{\mathbf{u}}.$$

В развёрнутой форме записи оно тождественно уравнению:

$$\left(N \cdot \sum_{q=0}^{N-1} \mathbf{u}_q \mathbf{u}_q' - \sum_{q=0}^{N-1} \mathbf{u}_q \cdot \sum_{q=0}^{N-1} \mathbf{u}_q' \right) \cdot \dot{\mathbf{b}} = N \cdot \sum_{q=0}^{N-1} \tilde{y}_q \mathbf{u}_q - \sum_{q=0}^{N-1} \tilde{y}_q \cdot \sum_{q=0}^{N-1} \mathbf{u}_q.$$

В кратчайшей матрично-векторной форме записи оно имеет вид:

$$\mathbf{S} \cdot \dot{\mathbf{b}} = \mathbf{w}.$$

Отсюда вычисляется коэффициент $\dot{\mathbf{b}}$:

$$\dot{\mathbf{b}} = \left(N \cdot \sum_{q=0}^{N-1} \mathbf{u}_q \mathbf{u}_q' - \sum_{q=0}^{N-1} \mathbf{u}_q \cdot \sum_{q=0}^{N-1} \mathbf{u}_q' \right)^{-1} \cdot \left(N \cdot \sum_{q=0}^{N-1} \tilde{y}_q \mathbf{u}_q - \sum_{q=0}^{N-1} \tilde{y}_q \cdot \sum_{q=0}^{N-1} \mathbf{u}_q \right). \quad (500)$$

После этого из (498) вычисляется коэффициент $\dot{b}_0$:

$$\dot{b}_0 = \overline{\tilde{y}} - \overline{\mathbf{u}'} \cdot \dot{\mathbf{b}}. \quad (501)$$

Для однозначности $\dot{\mathbf{b}}$ и $\dot{b}_0$ необходимо, чтобы выполнялось $\det \mathbf{S} \neq 0$, что достигается выбором плана расположения точек $\mathbf{u}_q$ в $\mathcal{P}^K$.

Формулы (500) и (501) значительно упрощаются для специальных *центрированных планов* расположения точек $\mathbf{u}_q$ с значением $\mathbf{u}_0 = \mathbf{0}$:

$$\sum_{q=0}^{N-1} \mathbf{u}_q = \mathbf{0} = \bar{\mathbf{u}} \rightarrow$$

$$\rightarrow \dot{\mathbf{b}}_0 = \bar{\tilde{y}}, \quad \dot{\mathbf{b}} = \left\{ \sum_{q=0}^{N-1} \mathbf{u}_q \mathbf{u}_q' \right\}^{-1} \cdot \left(\sum_{q=0}^{N-1} \tilde{y}_q \mathbf{u}_q \right) = \left\{ \overline{\mathbf{u} \cdot \mathbf{u}'} \right\}^{-1} \cdot \overline{\tilde{y} \cdot \mathbf{u}}. \quad (502)$$

В планировании эксперимента, в силу нормированности планов, имеем: либо $\mathbf{u}_0 = \mathbf{0}$, либо центральная точка в основном расчёте не принимается во внимание. (Общее количество точек в плане по-прежнему равно N .) Выбор именно центрированных планов эксперимента, согласно (502), существенным образом упрощает расчёты. Центрированность плана обеспечивается симметричностью расположения точек относительно начала координат.

Ещё один важный вариант, а именно *централизованная линейная регрессия*, возникает тогда, когда в функциях и в моделях (494), (495) заведомо точно известно, что $\mathbf{b}_0 = 0$. Тогда здесь возможно применить аналогию с централизованным методом наименьших квадратов (§ 4.7), с учётом того, что $\mathbf{u} = (u_1, u_2, \dots, u_K)'$, $\dot{\mathbf{b}} = (\dot{b}_1, \dot{b}_2, \dots, \dot{b}_K)'$. При этом уравнение (498) принимает вид:

$$\bar{\tilde{y}} = \bar{\mathbf{u}}' \cdot \dot{\mathbf{b}} \leftrightarrow \sum_{q=0}^{N-1} \tilde{y}_q = \sum_{q=0}^{N-1} \mathbf{u}_q' \cdot \dot{\mathbf{b}}. \quad (503)$$

Очевидно, это уравнение даёт однозначное решение только при $K \leq 2$. Но поскольку при множественной регрессии $K \gg 1$, то однозначное решение получается из (499) при $\mathbf{b}_0 = 0$. В данном случае решается **система линейных нормальных уравнений Гаусса**, но теперь уже в варианте полифакторной линейной регрессии:

$$\left\{ \sum_{q=0}^{N-1} \mathbf{u}_q \mathbf{u}_q' \right\} \cdot \dot{\mathbf{b}} = \sum_{q=0}^{N-1} \tilde{y}_q \mathbf{u}_q \rightarrow \dot{\mathbf{b}} = \left\{ \sum_{q=0}^{N-1} \mathbf{u}_q \mathbf{u}_q' \right\}^{-1} \cdot \left(\sum_{q=0}^{N-1} \tilde{y}_q \mathbf{u}_q \right). \quad (504)$$

Причём решение (502) отличается от (504) лишь коэффициент $\dot{\mathbf{b}}_0$.

Разумеется, матрица системы S (в фигурных скобках) при любых центрированных планах и при централизованной регрессии не может быть плохо обусловленной или вырожденной. *Искусственным путём* последний вариант можно получить преобразованием вида:

$$y = f(\mathbf{u}) \rightarrow \psi(\mathbf{u}) = f(\mathbf{u}) - b_0 = \mathbf{u}' \cdot \mathbf{b} \sim \dot{y} = \dot{f}(\mathbf{u}) \rightarrow \dot{\psi}(\mathbf{u}) = \dot{y}(\mathbf{u}) - \tilde{y}(0) = \mathbf{u}' \cdot \dot{\mathbf{b}}.$$

* * *

Далее рассмотрим наиболее общий подход к линейной регрессии, а именно линейную регрессию $\tilde{y}|\mathbf{v}$, базирующуюся на *полнолинейных* моделях типа (495), (497). Все скалярные переменные u_p по-прежнему определяются как p -ые частные факторы (но при этом $u_0 = 1 = \text{const}$). Поэтому $(K+1) \times 1$ -векторные переменные $\mathbf{v}$ суть точки *расширенного факторного пространства* Φ^{K+1} . Конкретная q -я точка обозначается как $\mathbf{v}_q$. Количество точек в плане равно N . Причём имеем:

$$\mathbf{v}_0 = \begin{pmatrix} 1 \\ \mathbf{u}_0 \end{pmatrix} = \begin{pmatrix} 1 \\ \mathbf{0} \end{pmatrix} \quad (q = 0). \quad (505)$$

Это обусловлено тем, что планы эксперимента всегда централизованы. Обратим здесь особое внимание на то, что любые планы размещения точек $\mathbf{x}_q$ как аргумента целевой функции $y(\mathbf{x})$ реализуются *конкретно* в исходном аффинном координатном пространстве $\mathcal{A}^n$, но однозначно ими определяемые точки $\mathbf{v}_q = \mathbf{v}(\mathbf{x}_q)$ реализуются *абстрактно* в более высокоразмерном факторном аффинном пространстве Φ^{K+1} . Именно над пространством Φ^{K+1} при данной линейной регрессии применяется, как базовый, метод наименьших квадратов (§ 4.7), а затем выполняются основанные здесь на законе центрального нормального распределения ошибки опыта $(\tilde{y} - y)$ необходимые статистические оценки. Разумеется, при этом полное отображение аффинного пространства $\mathcal{A}^n$ в более высокоразмерное аффинное пространство Φ^{K+1} заполняет его только частично. Поэтому метод наименьших квадратов работает над Φ^{K+1} шире, нежели над $\mathcal{A}^n$. Но, с другой стороны, это оправдано, поскольку факторы u_p в своей совокупности обладают большей степенью свободы, нежели исходные аргументы x_i функции $y(\mathbf{x})$. Подобное явление имело место и выше, так как при аппроксимации коэффициентов b_0 и $\mathbf{b}$, по сути, применялся метод наименьших квадратов над тем же Φ^{K+1} .

Пусть по некоторому плану эксперимента в $\mathcal{A}^n$ получена выборка из парных значений аргумента и целевой функции отклика $\langle \mathbf{x}_q, \tilde{y}_q \rangle$, где q пробегает последовательно значения от 0 до $N - 1$. Соответственно значения аргумента были заданы точно, а значения функции получены со случайной ошибкой опыта $(\tilde{y} - y)$, имеющей плотность вероятности, согласно закону центрального нормального распределения.

Затем трансформируем исходную выборку с $\mathbf{x}_q$ в тождественную ей выборку с $\mathbf{v}_q$, т. е. $\langle \mathbf{x}_q, \tilde{\mathbf{y}}_q \rangle \rightarrow \langle \mathbf{v}_q, \tilde{\mathbf{y}}_q \rangle$; $q=0, N-1$. Множество значений функции отклика $\tilde{\mathbf{y}}_q$ составляет $N \times 1$ -вектор $\tilde{\mathbf{y}} = (y_0, y_1, y_2, \dots, y_{N-1})'$, или вектор-столбец из отдельных значений $\tilde{y}_q$. Этот вектор является элементом некоторого функционального евклидова пространства $\mathcal{E}^N$. (Его евклидова метрика обусловлена здесь тем, что в качестве целевого параметра применяется сумма квадратов разностей координат, т. е. согласно базовому методу наименьших квадратов.) В свою очередь, множество значений радиус-векторов $\mathbf{v}_q'$ в факторном пространстве составляет $N \times (K+1)$ -матрицу $V = \{\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{N-1}\}'$, где $N > K + 1$. Эта прямоугольная матрица есть характеристика, определяемая только исходным расположением точек $\mathbf{x}_q$ в плане эксперимента в $\mathcal{A}^n$. Она, т. е. V , традиционно определяется как *матрица планирования*.

С учётом (497) и этих понятий, уравнения для минимизации суммы квадратов отклонений $(\tilde{\mathbf{y}}_q - \dot{\mathbf{y}}_q)$ в $\mathcal{E}^N$ типа (413), (414) принимают весьма упрощённый и в итоге чисто алгебраический вид:

$$\psi = \psi(\dot{\mathbf{h}}) = (\tilde{\mathbf{y}} - \dot{\mathbf{y}})' \cdot (\tilde{\mathbf{y}} - \dot{\mathbf{y}}) = (\tilde{\mathbf{y}} - V \cdot \dot{\mathbf{h}})' \cdot (\tilde{\mathbf{y}} - V \cdot \dot{\mathbf{h}}) = \min, \quad (506)$$

$$\frac{1}{2} \cdot \frac{d\psi}{d\dot{\mathbf{h}}} = V' \cdot \tilde{\mathbf{y}} - V' \cdot V \cdot \dot{\mathbf{h}} = \mathbf{0}. \quad (507)$$

Множество точек $\mathbf{x}_q$ должно быть таково, чтобы $(K+1) \times (K+1)$ -матрица внутренней гомомультимпликации $V' \cdot V$ не была плохо обусловленной ($\det V' \cdot V \approx 0$) или вырожденной ($\det V' \cdot V = 0$). Естественно, что для центрированных планов это выполняется. Поэтому из полученного линейного уравнения (507) в итоге имеем обязательно однозначное решение для $(K+1) \times 1$ -вектора коэффициентов линейной регрессии:

$$\dot{\mathbf{h}} = (V' \cdot V)^{-1} \cdot V' \cdot \tilde{\mathbf{y}} = V_L^{-1} \cdot \tilde{\mathbf{y}}, \quad (508)$$

где V_L^{-1} — однозначная левая обратная $(K+1) \times N$ -матрица для V , так как $V_L^{-1} \cdot V = I_{(K+1) \times (K+1)}$. Но при умножении справа она же играет роль квазиобратной матрицы Мура — Пенроуза $V^+ = V_L^{-1}$ (§ 2.1). Поэтому $V \cdot V_L^{-1} = V \cdot V^+ = \overset{\leftarrow}{V} V'$ — есть симметричный $N \times N$ -ортопроектор, проецирующий в функциональном пространстве $\mathcal{E}^N$ на образ $\langle \text{im } V \rangle$ параллельно ядру $\langle \ker V' \rangle$, т. е. попросту ортогонально $\langle \text{im } V \rangle$.

Этот ортопроектор в функциональном пространстве $\mathcal{E}^N$ в полной сумме дополняется другим симметричным $N \times N$ -ортопроектором $\overrightarrow{V}V'$, проецирующим на ядро $\langle \ker V' \rangle$ параллельно образу $\langle \text{im } V \rangle$ — попросту ортогонально $\langle \ker V' \rangle$. Откуда имеем: $\overleftarrow{V}V' + \overrightarrow{V}V' = I_{N \times N}$. Такого рода проекторы подробным образом рассматривались ранее в гл. 2.

С их явной помощью, а также с применением вышеуказанной левой обратной матрицы, можно весьма наглядно, т. е. чисто геометрическим путём, в целом интерпретировать как линейную регрессию, так и сам базовый метод наименьших квадратов. Геометрическая интерпретация, в принципе, вполне аналогична той, которая даётся для однозначного квазиразрешения и минимодульной невязки *несовместного* линейного алгебраического уравнения (см. § 4.5), но с конкретизацией к объектам рассматриваемой общей задачи. В качестве несовместного линейного уравнения здесь фигурирует выражение:

$$\underset{N \times (K+1)}{V} \cdot \underset{(K+1) \times 1}{\dot{\mathbf{h}}} = \underset{N \times 1}{\tilde{\mathbf{y}}}. \quad (509)$$

При этом имеем:

$$\text{rang } V = \text{rang } V' = \text{rang } \{V \cdot V'\} = \text{rang } \{V' \cdot V\} = K+1. \quad (510)$$

Отсюда $\langle \ker V \rangle = \mathbf{0}$. Поэтому из формулы (508) с использованием левой квазиобратной матрицы V_L^{-1} вычисляется однозначное квазиразрешение $\dot{\mathbf{h}}$. Вектор невязки (§ 4.5) этого линейного уравнения имеет вид:

$$\delta(\dot{\mathbf{h}}) = V \cdot \dot{\mathbf{h}} - \tilde{\mathbf{y}} = \dot{\mathbf{y}} - \tilde{\mathbf{y}}. \quad (511)$$

Именно при квазиразрешении (508) обеспечивается минимум евклидовой нормы вектора невязки $\delta(\dot{\mathbf{h}})$ уравнения (509), а геометрически в $\mathcal{E}^N$ — ортогональность вектора невязки $\delta(\dot{\mathbf{h}}) = \dot{\mathbf{y}} - \tilde{\mathbf{y}}$ и $\langle \text{im } V \rangle$, т. е. всем N векторам-строкам $\mathbf{v}_q'$ в матрице планирования V :

$$\underset{N \times 1}{\dot{\delta}} = - \underset{N \times N}{\overrightarrow{V}V'} \cdot \underset{N \times 1}{\tilde{\mathbf{y}}}. \quad (512)$$

В результате можно сделать главный геометрический вывод, что найденные выше в результате этого линейного регрессионного анализа коэффициенты полнолинейной модели (497) обеспечивают минимум суммы квадратов отклонений $(\tilde{y}_q - \dot{y}_q)$, взятых по N субосям y_q в $\mathcal{E}^N$. Аналогичное было показано ранее в § 4.7 в простейшем *одномерном* варианте метода наименьших квадратов. Подставив в (506) найденный вектор коэффициентов модели, получаем выражение для минимальной суммы квадратов отклонений опытных от регрессионных значений функции отклика, в том числе через вектор минимодульной невязки:

$$q_{\min} = \left[\left(I - V \overset{\leftarrow}{V'} \right) \tilde{\mathbf{y}} \right]' \cdot \left[\left(I - V \overset{\leftarrow}{V'} \right) \tilde{\mathbf{y}} \right] = \left(\vec{V} \overset{\rightarrow}{V'} \tilde{\mathbf{y}} \right)' \cdot \left(\vec{V} \overset{\rightarrow}{V'} \tilde{\mathbf{y}} \right) = \tilde{\mathbf{y}}' \cdot \vec{V} \overset{\rightarrow}{V'} \cdot \tilde{\mathbf{y}}. \quad (513)$$

Итак, внешняя гомомультипликация матрицы планирования $\{V \cdot V'\}$ есть *сингулярная* симметричная $N \times N$ -матрица, так как здесь $N > K+1$, где K — количество переменных факторов u_p , а $K+1$ — общее количество коэффициентов в модели регрессии типа (497). Именно благодаря этому факту существуют характеристические ортопроекторы, выполняющие необходимое ортопроецирование в функциональном пространстве $\mathcal{E}^N$.

В случае *централизованной линейной регрессии*, когда в функциях и в их моделях (494)–(497) заведомо известно, что $a_0 = b_0 = 0$, очевидно, $(K+1) \times 1$ -радиус-векторы $\mathbf{v}_q$ заменяются на $K \times 1$ -радиус-векторы $\mathbf{u}_q$, а в уравнениях (506)–(513) соответственно заменяют $N \times (K+1)$ -матрицу планирования $V = \{\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{N-1}\}'$ ранга $K+1$ на $N \times K$ -матрицу планирования $U = \{\mathbf{u}_0, \mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{N-1}\}'$ ранга K и $(K+1) \times 1$ -вектор-коэффициент $\dot{\mathbf{h}}$ заменяют на $K \times 1$ -вектор-коэффициент $\dot{\mathbf{b}}$.

Например, из (508) следует решение для модели (496) в форме:

$$U' \cdot U \cdot \dot{\mathbf{b}} = U' \cdot \tilde{\mathbf{y}} \quad \rightarrow \quad \dot{\mathbf{b}} = (U' \cdot U)^{-1} \cdot U' \cdot \tilde{\mathbf{y}} = U_L^{-1} \cdot \tilde{\mathbf{y}}. \quad (514)$$

Эта система и система линейных нормальных уравнений Гаусса (504) для *централизованной полифакторной линейной регрессии* типа $\tilde{y}|u$, тождественны друг другу, так как действуют формулы соответствия:

$$\begin{aligned} U' \cdot U &= \sum_{q=0}^{N-1} \mathbf{u}_q \mathbf{u}_q'; \quad U' \cdot \tilde{\mathbf{y}} = \sum_{q=0}^{N-1} \tilde{y}_q \mathbf{u}_q \quad \rightarrow \\ &\rightarrow \quad \dot{\mathbf{b}} = (U' \cdot U)^{-1} \cdot U' \cdot \tilde{\mathbf{y}} = \left\{ \sum_{q=0}^{N-1} \mathbf{u}_q \mathbf{u}_q' \right\}^{-1} \cdot \left(\sum_{q=0}^{N-1} \tilde{y}_q \mathbf{u}_q \right). \end{aligned} \quad (515)$$

В любом из этих 2-х вариантов регрессии (нецентрализованной и централизованной) внутренняя гомомультипликация $W = W' = V'V$ или $W = W' = U'U$ традиционно определяется как *информационная матрица* для заданного плана эксперимента (матрицы планирования). Если столбцы матрицы V или U попарно *ортогональны* друг другу, то она в целом *ортогональная по столбцам*. Для неё $W = D$ есть диагональная информационная матрица размера $(K+1) \times (K+1)$ или $K \times K$. Такая матрица легко обращается, например, в формулах (508) и (514). Именно в этом и состоит *главное преимущество* связанных с такими ортогональными матрицами планов экспериментов.

Информационная матрица $W = V'V$ имеет блочную структуру

$$W = \begin{Bmatrix} N & \sum_{q=0}^{N-1} \mathbf{v}_q' \\ \sum_{q=0}^{N-1} \mathbf{v}_q & \sum_{q=0}^{N-1} \mathbf{v}_q \mathbf{v}_q' \end{Bmatrix} \quad (516)$$

$$\text{размера} \begin{pmatrix} 1 \times 1 & 1 \times K \\ K \times 1 & K \times K \end{pmatrix}.$$

Подставив блок (516) в решение (508), получаем те же формулы (500) и (501) для коэффициентов регрессионной модели (496), что говорит о тождественности двух изложенных подходов к линейной регрессии. Информационная матрица $W = U'U$ также после подстановки блочной формы в решение (508) даёт (504).

Однако есть принципиальные различия в применении обеих форм решений—традиционной и общелинейной. Например, нетрудно видеть значительную разницу здесь в размерах рабочих матриц в формулах, выражающих искомые векторные коэффициенты регрессии. В первом варианте их размер меньше. Однако во втором варианте существенное упрощение именно вычисления векторного коэффициента регрессии в целом и его скалярных компонентов по отдельности обеспечивают вышеуказанные специальные *ортогональные матрицы планирования* V или U (попарно ортогональные по столбцам). Таким образом, выбор формы решения из этих двух зависит от конкретных обстоятельств.

В свою очередь, матрица W^{-1} , фигурирующая в формулах (508) и (514), формирует, с учётом σ^2 , *ковариационную матрицу*:

$$S = W^{-1} \cdot \sigma^2 \{ \tilde{y} \} = S'. \quad (517)$$

Формально она возникает при вычислении матричной дисперсии регрессионного коэффициента $\mathbf{h}$ в (508) или $\dot{\mathbf{b}}$ в (514). Например:

$$\begin{aligned}\mu\{(\dot{\mathbf{h}} - \mathbf{h}) \cdot (\dot{\mathbf{h}} - \mathbf{h})'\} &= \mu\left\{\left[(V'V)^{-1} \cdot V' \cdot (\tilde{\mathbf{y}} - \mathbf{y})\right] \cdot \left[(V'V)^{-1} \cdot V' \cdot (\tilde{\mathbf{y}} - \mathbf{y})\right]'\right\} = \\ &= (V'V)^{-1} \cdot V' \cdot \mu\{(\tilde{\mathbf{y}} - \mathbf{y}) \cdot (\tilde{\mathbf{y}} - \mathbf{y})'\} \cdot V \cdot (V'V)^{-1} = \\ &= (V'V)^{-1} \cdot V' \cdot [I \cdot \sigma^2\{\tilde{\mathbf{y}}\}] \cdot V \cdot (V'V)^{-1} = \\ &= (V'V)^{-1} \cdot \sigma^2\{\tilde{\mathbf{y}}\} = W^{-1} \cdot \sigma^2\{\tilde{\mathbf{y}}\} = S = S'.\end{aligned}$$

Здесь используется также тот факт, что из формулы (508) следует:

$$\mu\{\dot{\mathbf{h}}\} = (V'V)^{-1} \cdot V' \cdot \mathbf{y} = V_L^{-1} \cdot \mathbf{y}.$$

Матрица S , согласно плану, обуславливает *плановые* дисперсии и ковариации для ошибок скалярных коэффициентов линейной регрессии $\sigma^2\{\dot{\mathbf{b}}_p\} = s_{pp} \cdot \sigma^2\{\tilde{y}\}$, $\mathbf{v}\{\dot{\mathbf{b}}_r, \dot{\mathbf{b}}_s\} = s_{rs} \cdot \sigma^2\{\tilde{y}\} = \mathbf{v}\{\dot{\mathbf{b}}_s, \dot{\mathbf{b}}_r\}$. (518)

Плановые коэффициенты линейной корреляции для ошибок $\dot{\mathbf{b}}_r$ и $\dot{\mathbf{b}}_s$:

$$\lambda\{\dot{\mathbf{b}}_p, \dot{\mathbf{b}}_p\} = 1, \quad \lambda\{\dot{\mathbf{b}}_r, \dot{\mathbf{b}}_s\} = \lambda\{\dot{\mathbf{b}}_s, \dot{\mathbf{b}}_r\} = s_{rs} / \sqrt{s_{rr}} \cdot \sqrt{s_{ss}}. \quad (519)$$

Кроме того, для *ортогональных матриц планирования* или *планов* в факторном пространстве должны иметь место простые соотношения:

$$W = V'V = D, \quad W^{-1} = D^{-1} \rightarrow S = \sigma^2\{\tilde{y}\} \cdot D^{-1}. \quad (520)$$

Тогда, в силу (519), коэффициенты модели $y(\mathbf{x})$, отвечающие плану V , вычисляются с ошибками, *статистически* независимыми друг от друга.

В отличие от *выборочного* коэффициента корреляции (404) из § 4.7 коэффициенты (519) являются математически ожидаемыми понятиями. Они зависят только от матрицы планирования (где элементы точные) и определяют степень линейной связи ошибок от коэффициентов $\dot{\mathbf{b}}_p$ в разностной модели $\hat{y} = \hat{y}(\mathbf{x})$. Для аналитических целевых функций коэффициенты их разностных степенных моделей 2-го порядка всегда как-то смешаны с коэффициентами высших порядков, в зависимости от разностной схемы оценки с прямоугольной матрицей плана. Несмотря на неравенство коэффициентов модели своим истинным прототипам в степенном разложении функции $y = y(\mathbf{x})$ разностные методы неплохо работают при численной оптимизации, что фактически используется в планировании эксперимента и в плано-вычислительных методах.

Метод наименьших квадратов и поэтому нормальная регрессия на его основе инвариантны к линейным, в том числе к масштабным преобразованиям координат базиса. Инвариантность следует, например, из формул (500) и (501) для модели типа (494). При $\mathbf{u} = V_{(u \rightarrow z)} \mathbf{z}$ имеем преобразования: $\mathbf{b}_{(z)}' = V_{(u \rightarrow z)}^{-1} \cdot \mathbf{b}_{(u)}$, $\mathbf{b}_{0(z)} = \mathbf{b}_{0(u)}$. Информационная W и ковариационная S матрицы преобразуются также ковариантно.

§ 7.3. Планирование эксперимента при оптимизации целевой функции отклика и дискуссионные вопросы

Оптимизация целевых функций отклика от нескольких скалярных переменных $\tilde{y} = \tilde{y}(\mathbf{x}) = \tilde{y}(x_1, x_2, \dots, x_n)$ с привлечением результатов экспериментальной части и полифакторного регрессионного анализа составляет ныне важный практически раздел в общей теории решений экстремальных задач. Традиционно этот раздел называют в специальной литературе *планированием эксперимента* [см., например, 15, 43, 46, 47]. Исторически первую работу в этой области опубликовал Жозеф Жергонн (1815 г.) — известный геометр и автор принципа двойственности [64].

В новейшее время это направление развивается с середины XX века. Пионерская публикация (1951 г.) принадлежит Дж. Боксу и К. Уилсону [54]. В ней, в частности, сразу были предложены весьма эффективные композиционные планы (n -ЦКП) для экспериментальной оптимизации с моделью функции отклика 2-го порядка. Кроме того, именно Дж. Бокс впервые рассмотрел применение и особенности процесса планирования эксперимента на промышленных объектах в сфере производства [55].

Однако планирование эксперимента в своём нынешнем виде пока не характеризуется в целом надлежащей математической строгостью. Так, в частности, в нём при оценке адекватности модели регрессии весьма некорректно применяются критерий Фишера — Снедекора, игнорируется соотношение случайной и систематической ошибок и влияние на него интервалов варьирования факторов. Но что более всего удивительно: не придаётся особого значения различию размерностей переменных и, вместе с тем, используется ряд неких критериев оптимальности планов, имеющих какой-либо смысл лишь для евклидовой метрики факторного пространства. Подобные противоречия в нынешнем содержании теории планирования эксперимента вызывают до сих пор широко известное недоверие к нему со стороны высокопрофессиональных математиков. Поэтому в последующих параграфах главы параллельно с изложением материала, по возможности, устраняются имеющиеся противоречия.

Рассмотрение начнём с того, что во всех статистических процедурах планирования эксперимента (как главная гипотеза) априори заложено то, что ошибка нахождения целевой функции $\mathbf{r} = \tilde{\mathbf{y}}(\mathbf{x}) - \mathbf{y}(\mathbf{x})$ подчиняется закону нормального распределения с параметрами: $\mu(\mathbf{r}) = 0$, $\sigma^2(\mathbf{r}) = \sigma^2$. Традиция, по сути, идёт от общепринятой процедуры статистической обработки экспериментальных данных. Поскольку сама оптимизация целевой функции отклика — это многошаговый процесс, требующий на каждом новом шаге повтора статистических оценок, то целесообразно всё же как-то проверять, надёжна ли вообще эта главная гипотеза.

Хорошо известно, что закон распределения конкретной случайной величины *исходя лишь из её конечной выборки* абсолютно достоверно установить нельзя. Но для ошибок измерений, имеющих в комплексе совершенно случайный характер, на практике априори принимают **закон центрального нормального распределения Гаусса** [29]:

$$p(r) = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot \exp(-r^2 / 2\sigma^2), \quad -\infty < r < +\infty, \quad (521)$$

где $r = (\tilde{y} - y)$ — случайная ошибка, $p(r)$ — плотность распределения её вероятности, σ^2 — дисперсия нормального распределения. График этой функции имеет характерную форму колокола и в литературе называется *кривой ошибок*. Площадь под ней, очевидно, единичная: $\int_{-\infty}^{+\infty} p(r) dr = 1$.

Из данного нормирующего условия теоретически следует коэффициент в (521). В зависимости от величины σ кривая ошибок растягивается или сжимается в сравнении со *стандартным нормальным распределением*:

$$p\left(\frac{r}{\sigma}\right) = \frac{1}{\sqrt{2\pi}} \cdot \exp\left[-\frac{1}{2}\left(\frac{r}{\sigma}\right)^2\right] = \frac{1}{\sqrt{2\pi}} \cdot \exp(-u^2 / 2) = p(u), \quad (522)$$

где $u = r/\sigma$ — *нормированная случайная ошибка*.

Вероятность того, что ошибка r *по величине* не выходит за значение $\Delta > 0$ ($r < \Delta$), даёт **функция вероятности Лапласа 1-го рода**

$$\Phi_1\left(\frac{\Delta}{\sigma}\right) = \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{\Delta/\sigma} \exp\left[-\frac{1}{2}\left(\frac{r}{\sigma}\right)^2\right] d\left(\frac{r}{\sigma}\right) = \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{\Delta/\sigma} \exp(-u^2 / 2) du. \quad (523)$$

Вероятность того, что ошибка r *по абсолютной величине* не выходит за значение $\Delta > 0$ ($|r| < \Delta$), даёт **функция вероятности Лапласа 2-го рода**

$$\Phi_2\left(\frac{\Delta}{\sigma}\right) = \frac{2}{\sqrt{2\pi}} \cdot \int_0^{\Delta/\sigma} \exp\left[-\frac{1}{2}\left(\frac{r}{\sigma}\right)^2\right] d\left(\frac{r}{\sigma}\right) = \frac{2}{\sqrt{2\pi}} \cdot \int_0^{\Delta/\sigma} \exp(-u^2 / 2) du. \quad (524)$$

Интегралы (498) и (499) не выражаются в элементарных функциях, но при этом, как и функция (497), представлены достаточно точно и широко в табулированных формах. Имеют место очевидные формулы их связи, ввиду симметрии функции плотности вероятности $p = p(u)$:

$$\Phi_1(u) = \frac{1}{2} \cdot [\Phi_2(u) + 1], \quad \Phi_2(u) = -\Phi_2(-u) = 2\Phi_1(u) - 1. \quad (525, 526)$$

Особо отметим, что закон нормального распределения Гаусса (521) является *предельной научной абстракцией*. Это хорошо видно хотя бы из факта неограниченности величины σ . На самом же деле случайные ошибки опыта практически всегда ограничены каким-либо физическим образом. Но именно данный закон является главной теоретической основой для всех применяемых стандартных статистических процедур обработки опытных результатов измерений и наблюдений, когда общие случайные ошибки вызваны большой совокупностью причин, чей вклад по отдельности весьма мал. Строгое теоретическое объяснение этому факту придала **центральная предельная теорема Ляпунова** (1900 г.). А именно, в этой теореме сформулировано достаточное условие, при котором сумма малых случайных отклонений с ростом их количества имеет в итоге асимптотически нормальное распределение. Достаточным оказалось то, чтобы в этой сумме частные отклонения не доминировали над всеми остальными случайными отклонениями [31].

Чисто математически закон нормального распределения может быть строго выведен, например, из естественного требования, что наиболее вероятное значение величины, для которой имеется выборка отдельных случайных значений, есть их среднее арифметическое. Далее применяют основанные на этом законе методы *математической статистики* [44].

Пусть в некоторой, например, начальной точке $\mathbf{x} \in \mathcal{A}^n$ (т. е. при точных значениях частных переменных $x_1, x_2, \dots, x_n$) экспериментально получена выборка отдельных значений целевой функции отклика $\langle \tilde{y}_j \rangle$ объема m , где $j = 1, m$.

Для каждого закона распределения есть основные количественные параметры — моменты различных порядков. Например, для функции типа (521) — это *начальные* и, вместе с тем, *центральные моменты* (относительно начала координат и, вместе с тем, центра распределения). Закон нормального распределения однозначно характеризуется только моментами 1-го и 2-го порядка.

$$\mu\{r\} = \mu\{\tilde{y} - y\} = \int_{-\infty}^{+\infty} r \cdot p(r) dr = 0 \quad (527)$$

— математическое ожидание для r , или *относительный* момент 1-го порядка. Его выборочная оценка вообще есть среднее арифметическое. Но тут — для централизованной величины с $\mu = 0$ она совпадает с моментом, так как $\overline{(\tilde{y} - \tilde{y})} = 0$. В общем же случае для $\tilde{y}$ выборочная оценка есть среднее арифметическое $\tilde{\tilde{y}}$, но при этом $\mu(\tilde{y}) = \mu(\tilde{\tilde{y}}) = y$.

$$\mu\{|r|\} = \mu\{|\tilde{y} - y|\} = \int_{-\infty}^{+\infty} |r| \cdot p(r) dr = 2 \int_0^{+\infty} r \cdot p(r) dr = \sqrt{2/\pi} \cdot \sigma = \delta \quad (528)$$

— математическое ожидание для $|r|$, или *абсолютный момент 1-го порядка*. Его выборочная оценка, согласно формулам (489), (490), есть *несмещённое абсолютизированное выборочное среднее*:

$$d\{\tilde{y}\} = \frac{1}{\sqrt{m(m-1)}} \cdot \sum_{j=1}^m |\tilde{y}_{q,j} - \bar{\tilde{y}}_q| = \frac{1}{\sqrt{m(m-1)}} \cdot \overline{|\tilde{y}_{q,j} - \bar{\tilde{y}}_q|}$$

$$\left(\mu\{d\{\tilde{y}\}\} = \delta\{\tilde{y}\} = \sqrt{2/\pi} \cdot \sigma\{\tilde{y}\} < \sigma\{\tilde{y}\} \right).$$

$$\mu\{r^2\} = \mu\{(\tilde{y} - y)^2\} = \int_{-\infty}^{+\infty} r^2 \cdot p(r) dr = 2 \int_0^{+\infty} r^2 \cdot p(r) dr = \sigma^2 \quad (529)$$

— математическое ожидание для r^2 , или момент 2-го порядка, или *дисперсия*. Её выборочная оценка, согласно формулам (485), (486), есть *несмещённая выборочная дисперсия*:

$$s^2 = \frac{1}{m-1} \cdot \sum_{j=1}^m \left(\tilde{y}_{q,j} - \bar{\tilde{y}}_q \right)^2.$$

Математические ожидания для d есть $\delta = \sqrt{2/\pi} \cdot \sigma$, для s^2 есть σ^2 , т. е. они взаимосвязаны. Следовательно, с возрастанием m параметры d и s в случае нормального распределения $\tilde{y}$ должны стремиться точно к такой же взаимосвязи. Отсюда вытекает *предельная формула*:

$$\frac{\pi}{2} = \lim_{m \rightarrow \infty} \left\{ m \cdot \frac{\sum_{j=1}^m (\tilde{y}_j - \bar{\tilde{y}})^2}{\left[\sum_{j=1}^m |\tilde{y}_j - \bar{\tilde{y}}| \right]^2} \right\}. \quad (530)$$

(Данная формула может подвергаться экспериментальной проверке при больших объёмах выборки и выполнении условия случайности!)

С алгебраической точки зрения, отношение d^2/s^2 ограничивается сверху неравенством для средних положительных величин (§ 4.1) как

$$d^2/s^2 \leq 1 \leftrightarrow d/s \leq 1 \leftrightarrow d \leq s. \quad (531)$$

Согласно этому *алгебраическому* подходу, отношение d/s есть индекс равномерности абсолютных отклонений $\tilde{y}$ от $\bar{\tilde{y}}$. Однако, согласно *статистическому* подходу, отношение d/s есть индекс нормальности отклонений $\tilde{y}$ от y , а, следовательно, и нормальности распределения случайной величины $\tilde{y}$. Причём он работает тем лучше, чем больше m :

$$d^2/s^2 \rightarrow 2/\pi \approx 0,637 \quad (m \rightarrow \infty), \quad (532)$$

$$d/s \rightarrow \sqrt{2/\pi} \approx 0,798.$$

Теоретической основой для этого является вышеуказанная взаимосвязь их математических ожиданий $\delta = \sqrt{2/\pi} \cdot \sigma$, или

$$\pi \left(\frac{\delta}{2} \right)^2 = \left(\frac{\sigma}{\sqrt{2}} \right)^2 \leftrightarrow \delta^2 = \frac{2}{\pi} \sigma^2 \leftrightarrow \delta^2 / \sigma^2 = 2 / \pi. \quad (533)$$

Геометрически 1-я *квадратурная формула* означает, что площадь круга радиуса $\delta/2$ равна площади квадрата со стороной $\sigma / \sqrt{2}$, где параметры взяты из несобственных интегралов (528), (529) с пределами от 0 до $+\infty$.

С учётом вышеизложенного, параметр d^2/s^2 можно применять для оценки нормальности распределения случайной величины. Например, возьмём в качестве выборки классические результаты ряда измерений заряда электрона, выполненных Милликеном в 1910 — 1914 гг. Одна из последних серий экспериментов дала следующую таблицу значений заряда в электростатических единицах, с учётом коэффициента 10^{-13} , при $m = 20$ (данные опытов взяты из [31]):

4781	4775	4764	4789
4795	4772	4776	4764
4769	4791	4771	4774
4792	4782	4789	4778
4779	4767	4772	4791

$$\bar{\tilde{y}} = \frac{\sum_{j=1}^{20} \tilde{y}_j}{20} \approx 4778,55;$$

$$\sum_{j=1}^{20} |\tilde{y}_j - \bar{\tilde{y}}| = 164,10 \rightarrow d = 164,10/19 \approx 8,637, \quad d^2 \approx 74,595;$$

$$\sum_{j=1}^{20} (\tilde{y}_j - \bar{\tilde{y}})^2 = 1828,95 \rightarrow s^2 = 1828,95/19 \approx 96,26;$$

$$d^2/s^2 \approx 74,595/96,26 \approx 0,775.$$

Полученное значение индекса нормальности распределения $\tilde{y}(\mathbf{x})$ находится между алгебраическим верхним пределом 1, согласно (531), и математическим ожиданием 0,637, согласно (532). Следовательно, в имеющейся выборке случайных значений в этом конкретном примере не имеется какого-то заметного противоречия гипотезе о нормальном распределении $\tilde{y}$. Это установлено здесь на основе индекса d^2/s^2 .

Применение моментов более высокого порядка (> 2), с целью оценки справедливости той же гипотезы о нормальном распределении случайной ошибки, вряд ли целесообразно, ввиду гораздо большего накопления статистических отклонений в их выборочных значениях.

* * *

Иной способ проверки гипотезы о нормальности распределения $\tilde{y}$, а именно с принимаемой *доверительной вероятностью* $\mathcal{P}$, реализуется даже при относительно небольшом объёме m частной выборки $\langle \tilde{y}_j \rangle$ с использованием t -критерия согласия Стьюдента. В данном критерии изначально заложен закон нормального распределения для случайной величины $\tilde{y}$. Но при этом само T -распределение Стьюдента в отличие от нормального распределения вполне *конкретно*. Обычное отношение, которое подчинено T -распределению Стьюдента, — это статистика типа:

$$u = \frac{\bar{\tilde{y}} - y}{\sigma / \sqrt{m}} / \sqrt{\frac{s^2 \cdot (m-1)}{\sigma^2} / (m-1)} = \frac{\bar{\tilde{y}} - y}{s / \sqrt{m}}. \quad (534)$$

Статистика является, по сути, отношением Стьюдента, поскольку $(\bar{\tilde{y}} - y) / (\sigma / \sqrt{m})$ есть по главной гипотезе нормированная нормально распределённая случайная величина — по закону (521), а $[s^2 \cdot (m-1) / \sigma^2]$ есть по ней же случайная величина, имеющая χ^2 -распределение Пирсона с числом степеней свободы $k = m - 1$. Соответственно эта статистика имеет T -распределения Стьюдента с тем же числом степеней свободы $k = m - 1$. Она имеет плотность распределения вероятности $p_k = p_k(u)$. Данная функция с ростом k неуклонно приближается к нормированной функции (522) для закона нормального распределения, при $k \rightarrow \infty$ они полностью совпадают. (На практике при $k > 30$ они уже довольно мало различимы между собой [15, 46].)

Интегралы вероятности от неё выражаются функциями 2-х видов:

$$T_1(t, k) = \int_{-\infty}^t p_k(u) du = \frac{1}{2} \cdot [T_2(t, k) + 1] \quad (535)$$

— для *одностороннего* t-критерия (функция Стюдента 1-го рода),

$$T_2(t, k) = 2 \int_0^t p_k(u) du = 2 \cdot T_1(t, k) - 1 \quad (536)$$

— для *двустороннего* t-критерия (функция Стюдента 2-го рода).

Интегралы вероятности (535), (536) при возрастании числа степеней свободы неуклонно приближаются к функциям Лапласа (523), (524). Соответственно для нормально распределённой $\tilde{y}$ имеем оценки для u в (534) через статистику u :

$$\mathcal{P}(u < t) = T_1(t, k), \quad \mathcal{P}(|u| < t) = T_2(t, k), \text{ где } t > 0.$$

Когда в задаче интересны пределы отклонений от u с доверительной вероятностью $\mathcal{P}$ в обе стороны от неё, тогда, естественно, применяется двусторонний критерий Стюдента. В таком случае с доверительной вероятностью $\mathcal{P}$ имеет место основное неравенство вида:

$$\bar{y} - t_2(\mathcal{P}, m-1) \cdot \frac{s}{\sqrt{m}} < y < \bar{y} + t_2(\mathcal{P}, m-1) \cdot \frac{s}{\sqrt{m}}. \quad (537)$$

Для проверки главной гипотезы о нормальности распределения $\tilde{y}$ с доверительной вероятностью $\mathcal{P}$ применим статистическое сравнение двух выборочных средних для $\tilde{y}$ с использованием того же t-критерия согласия Стюдента. Пусть, например, в начальной точке $\mathbf{x}_{(0)} \in \mathcal{A}^n$ экспериментально получены две независимые выборки $\langle \tilde{y}_{j(1)} \rangle$ и $\langle \tilde{y}_{j(2)} \rangle$ равного объема r (где $2r = m$). Величина $(\bar{\tilde{y}}_{(1)} - \bar{\tilde{y}}_{(2)})$ должна быть распределена нормально с $\mu = 0$ и дисперсией $\sigma\sqrt{2/r}$, см. (491), (493). При этом величина $(\bar{\tilde{y}}_{(1)} - \bar{\tilde{y}}_{(2)}) / (\sigma\sqrt{2} / \sqrt{r})$ имеет нормированное нормальное распределение. Для суммарной выборки $\langle \tilde{y} \rangle$ объёма $m = 2r$ несмещённая выборочная дисперсия равна s^2 . Величина $[s^2 \cdot (m-1) / \sigma^2]$ распределена как χ^2 с числом степеней свободы $k = m - 1$. образуем из этих величин статистику u , которая должна иметь T-распределение Стюдента с числом степеней свободы $k = m - 1$:

$$u = \frac{\bar{\tilde{y}}_{(1)} - \bar{\tilde{y}}_{(2)}}{\sigma\sqrt{2}/\sqrt{r}} / \sqrt{\frac{s^2 \cdot (m-1)}{\sigma^2} / (m-1)} = \frac{\bar{\tilde{y}}_{(1)} - \bar{\tilde{y}}_{(2)}}{s\sqrt{2}/\sqrt{r}} = \frac{\bar{\tilde{y}}_{(1)} - \bar{\tilde{y}}_{(2)}}{2s/\sqrt{m}}. \quad (538)$$

Поскольку знак выражения $(\bar{\tilde{y}}_{(1)} - \bar{\tilde{y}}_{(2)})$ априори неизвестен, то для оценки этой величины с доверительной вероятностью $\mathcal{P}$ используем двусторонний критерий Стьюдента:

$$|\bar{\tilde{y}}_{(1)} - \bar{\tilde{y}}_{(2)}| < t_2(\mathcal{P}, m-1) \cdot \frac{s\sqrt{2}}{\sqrt{r}} = t_2(\mathcal{P}, m-1) \cdot \frac{2s}{\sqrt{m}}. \quad (539)$$

Если эта разность двух выборочных средних попадает в указанный симметричный интервал, то с доверительной вероятностью $\mathcal{P}$ можно принять, что нет противоречия гипотезе о нормальности распределения случайной величины $\tilde{y}$, но пока только в центральной точке $\mathbf{x}_{(0)}$.

Продemonстрируем на конкретном примере проверку гипотезы по вероятностному неравенству (539), используя вышеуказанную таблицу экспериментальных значений заряда электрона, разделённую на два равных и статистически независимых блока объёмом $r = 10$. Имеем:

$$\bar{\tilde{y}}_{(1)} = 4780,3, \quad \bar{\tilde{y}}_{(2)} = 4776,8, \quad \bar{\tilde{y}}_{(1)} - \bar{\tilde{y}}_{(2)} = +3,5;$$

$$s = \sqrt{96,26} = 9,81, \quad s\sqrt{2}/\sqrt{r} = 9,81 \cdot 1,41/3,16 = 4,39.$$

Выпишем расчётные значения члена $t_2 \cdot (s\sqrt{2}/\sqrt{r})$ в неравенстве типа (539) при различных уровнях доверительной вероятности $\mathcal{P}$:

$\mathcal{P} = 0,99$:	$2,861 \cdot 4,39 = 12,56$;
$\mathcal{P} = 0,95$:	$2,093 \cdot 4,39 = 9,19$;
$\mathcal{P} = 0,90$:	$1,729 \cdot 4,39 = 7,59$;
$\mathcal{P} = 0,80$:	$1,328 \cdot 4,39 = 5,83$;
$\mathcal{P} = 0,70$:	$1,066 \cdot 4,39 = 4,68$;
$\mathcal{P} = 0,60$:	$0,861 \cdot 4,39 = 3,78$;
$\mathcal{P} = 0,50$:	$0,688 \cdot 4,39 = 3,02$.

То, что неравенство (539) выполняется даже при довольно низком уровне $\mathcal{P} = 0,60$, свидетельствует здесь о том, что проверяемая гипотеза практически достоверно не противоречит разбросу в общей выборке экспериментальных значений целевой функции $\tilde{y}$. Причём дальнейшее повышение уровня $\mathcal{P}$ тем более не опровергнет эту главную гипотезу, так как параметр t_2 будет только возрастать, увеличивая допустимый интервал для разности двух выборочных средних.

Вообще, при использовании понятия *доверительная вероятность* $\mathcal{P}$ здесь и далее нужно учитывать имеющийся в общепринятом подходе к проверке статистических гипотез явный парадокс. Он проявляется при подобного рода статистических оценках с применением $\Phi(u)$, $T(u)$ и других стандартных функций вероятности. Этот парадокс заключается в следующем. Чем при меньшем нижнем *предельном* уровне $\mathcal{P}$ данная статистическая гипотеза не опровергается, тем она более надёжная; и обратно — чем при большем нижнем *предельном* уровне $\mathcal{P}$ данная статистическая гипотеза не опровергается, тем она менее надёжная! Таким образом, в пределе при $\mathcal{P} \rightarrow 1$ любая статистическая гипотеза, проверяемая по этому критерию, не будет опровергнута (*даже неверная*), поскольку при этом t -критерий и соответственно допустимый интервал для случайной величины стремятся теоретически к бесконечности.

Именно поэтому на практике выбирается действительно априори необходимый уровень $\mathcal{P}$. Например, в научных экспериментах или при плано-экспериментальной оптимизации целевой функции применяют обычно $\mathcal{P} = 0,95$ — как для одностороннего, так и для двустороннего критерия согласия.

В конкретном вышеуказанном примере статистическая гипотеза о нормальности распределении случайной величины $\tilde{y}$ при $\mathcal{P} = 0,95$ не опровергается, причём с большим запасом по надёжности. В другой же серии подобных экспериментов вполне возможно, что отклонение $(\bar{\tilde{y}}_{(1)} - \bar{\tilde{y}}_{(2)})$ могло бы быть побольше, но в любом случае оно должно отвечать неравенству (539) при уровне $\mathcal{P} = 0,95$. Хороший запас по надёжности как раз свидетельствует о том, что за границы допустимого интервала эта случайная величина при принятом уровне $\mathcal{P}$ практически никогда не выйдет.

Если же проверяемая статистическая гипотеза не подтверждается, даже при довольно высокой доверительной вероятности $\mathcal{P} = 0,95$, то далее нужно предпринимать корректирующие действия по улучшению качества случайной выборки $\langle \tilde{y} \rangle$. Например, выявляют и устраняют какие-то имеющиеся систематические ошибки в опытных значениях $\tilde{y}$, отбрасывают по стандартным процедурам явно малоправдоподобные опытные данные, повышают точность измерений или наблюдений. Более затратный способ — увеличение количества экспериментальных значений $\tilde{y}$, или объёма m выборки $\langle \tilde{y} \rangle$. Тогда левая часть в (539) будет теоретически уменьшаться относительно правой части, причём несколько медленнее, чем $\sim 1/\sqrt{m}$.

С другой стороны, изученное ранее выборочное отношение (532) при $m \rightarrow \infty$ должно приближаться к теоретическому пределу в силу того, что $\tilde{d}$ и s приближаются к математическим ожиданиям δ и σ . Уменьшение σ только за счёт изменения масштаба для величины $\tilde{y}$ формально также уменьшает абсолютную ошибку s в её значениях. Но это, разумеется, никак не влияет на соотношение левой и правой частей в (539) и на отношение (532). Отсюда хорошо видна непригодность *абсолютной ошибки*, например, $\tilde{d}$ или s для оценки точности значения $\tilde{y}$ и, наоборот, хорошо видна актуальность *относительной ошибки* (погрешности) типа $\alpha_1 = 100\tilde{d}/\tilde{y}\%$ или $\alpha_2 = 100s/\tilde{y}\%$ для той же цели.

Например, в указанной выше серии измерений заряда электрона фактически имеем $s = 9,81 \cdot 10^{-13}$, $\alpha_2 = 0,2\%$. В пределе при $\alpha \rightarrow 0$ начинает уже сказываться и превалировать *ошибка округления* для $\tilde{y}$, которая имеет своё распределение. Следовательно, при этом неизбежно нарушение закона нормального распределения ошибки опыта в значениях целевой функции $\tilde{y}$ в точке $\mathbf{x}_{(0)}$, хотя точность здесь сверхвысокая!

Отсюда в планировании эксперимента с имеющимся значительным разбросом опытных данных при оценке качества выборки $\langle \tilde{y} \rangle$ должны играть главнейшие роли, как правило, следующие два обстоятельства. Первое — соответствие выборки закону нормального распределения с принимаемой доверительной вероятностью $\mathcal{P}$ для реализации затем возможности использования в линейной полифакторной регрессии базирующихся исходно на нём стандартных статистических процедур. Второе — достижение приемлемой по уровню относительной ошибки, например, $\alpha = \alpha_2 \leq 5\%$ (по аналогии с принятым уровнем доверительной вероятности $\mathcal{P} = 0,95$), что обеспечивает соответственно приемлемый уровень общей ошибки для находимых в процессе оптимизации $\mathbf{y}(\mathbf{x})$ статистических параметров: директивных векторов, точки экстремума, всех значимых коэффициентов модели регрессии и т. д.

* * *

Согласно первому из этих обстоятельств, следующий этап в оценке качества полученного множества экспериментальных значений $\langle \tilde{y} \rangle$, или выборки сводится к аналогичному статистическому анализу, но в целом на области плана эксперимента, и сравнению значения характеристики распределения с таковым для исходной начальной точки $\mathbf{x}_{(0)} \in \mathcal{A}^n$. (Последняя является геометрическим центром плана эксперимента.)

Например, при статистическом анализе выборки вначале находится *выборочная несмещённая дисперсия воспроизводимости*:

$$s_r^2 = s_r^2\{\tilde{y}\} = \frac{1}{K} \cdot \sum_{q=0}^{N-1} (m_q - 1) \cdot s^2\{\tilde{y}_q\} = \frac{1}{M - N} \cdot \sum_{q=0}^{N-1} \sum_{j=1}^{m_q} (\tilde{y}_{q,j} - \tilde{y}_q)^2, \quad (540)$$

где N — общее количество точек в плане эксперимента, m_q — количество опытных значений $\tilde{y}$ в каждой q -й точке, K — число степеней свободы, определяемое здесь естественным образом как

$$K = \sum_{q=0}^{N-1} (m_q - 1) = \sum_{q=0}^{N-1} k_q = M - N, \quad (541)$$

где

$$M = \sum_{q=0}^{N-1} m_q \quad (542)$$

— суммарное количество экспериментальных значений $\tilde{y}$ на множестве точек плана. Понятно, что указанная дисперсия воспроизводимости определяется при условии, что хотя бы некоторые (а лучше все) точки имеют $m_q > 1$. Ранее в центре $\mathbf{x}_{(0)} \in \mathcal{A}^n$ была найдена *выборочная дисперсия* $s_0^2 = s^2\{\tilde{y}_0\}$ с числом степеней свободы $k_0 = m_0 - 1$.

Пусть $s_{\max}^2 = \max\langle s_0^2, s_r^2 \rangle$, $s_{\min}^2 = \min\langle s_0^2, s_r^2 \rangle$; k_1 и k_2 — их числа степеней свободы. Критериальное отношение типа

$$F = s_{\max}^2 / s_{\min}^2 \geq 1 \quad (543)$$

должно здесь иметь F -распределение Фишера — Снедекора с числами степеней свободы k_1 для числителя и k_2 для знаменателя [15, 46]. Пусть F^* — его критическое (*табулярное*) значение при доверительной вероятности $\mathcal{P} = 0,95$. Тогда, если $F \leq F^*$, то далее с вероятностью $\mathcal{P}$ принимаем гипотезу об *однородности* нормального распределения $\tilde{y}$ на области плана эксперимента ($\sigma = \text{const}$). И обратно, если $F > F^*$, то эта гипотеза с той же вероятностью отвергается.

В первом случае переходим к построению, анализу и применению для оптимизации функции отклика соответствующей регрессионной модели. Во втором случае сначала выполняются вышеупомянутые корректирующие действия, но теперь для точек плана эксперимента в целом. Например, можно увеличивать объёмы m_q частных выборок $\langle \tilde{y}_{q,j} \rangle$, $j = 1, m_q$ во всех или в отдельных точках плана эксперимента.

§ 7.4. Критерий адекватности плано-регрессионных моделей

Для оптимизации целевой функции отклика $\tilde{y} = \tilde{y}(\mathbf{x})$ в планировании эксперимента обычно используют её локальные (т. е. на области планов) регрессионные модели. Однако при *высокой* точности эксперимента возможно попросту прямое применение разностных моделей функции отклика, находимых по *минимальным планам эксперимента*, подобных ранее описанным в § 6.3—6.5 планам вычислений. В последнем случае критерий адекватности модели и функции уже рассматривался в § 6.2. Особенностью такого планирования эксперимента является достаточно малая доля случайной ошибки для $\dot{y}(\mathbf{x})$ в её общей суммарной ошибке. Поэтому далее рассмотрим самый общий случай, при котором модель регрессии содержит и случайную, и систематическую ошибки. Понятно, что последняя может на областях планов эксперимента быть главной причиной завышения реальной ошибки для $\tilde{y}(\mathbf{x})$ в сравнении с той, что оценивается по закону нормального распределения с $\sigma = \text{const}$ и $\mu = 0$.

К модели линейной регрессии $\dot{y} = \dot{y}(\mathbf{x})$ предъявляют требование её адекватности функции отклика $\tilde{y} = \tilde{y}(\mathbf{x})$ на области плана эксперимента или конкретнее — на множестве точек $\langle \mathbf{x}_q \rangle$, $q = \overline{0, N-1}$. Но решение о том, адекватна регрессионная модель функции отклика или нет, принимают с задаваемой доверительной вероятностью $\mathcal{P}$ по критерию согласия Фишера — Снедекора [15, 29, 46], выражаемого отношением дисперсии адекватности и дисперсии воспроизводимости, сравнением его с F^* :

$$\hat{F} = s_{\text{ad}}^2 \{ \tilde{y}, \dot{y} \} / s_r^2 \{ \tilde{y} \} > 1. \quad (544)$$

(Причём, если $\hat{F} \leq 1$, то сразу же принимают, что модель и функция адекватны, так как на фоне случайной ошибки явно отсутствует какое-либо заметное смещение модели от функции.) Несмещённая дисперсия воспроизводимости $s_r^2 \{ \tilde{y} \}$ с числом степеней свободы $K_r = M - N$ уже вычислялась выше, согласно формуле (540). В свою очередь, дисперсия адекватности естественным образом вычисляется через полную сумму квадратов имеющих отклонений $\tilde{y}$ от $\dot{y}$, с учётом формулы (513):

$$s_{\text{ad}}^2 = s_{\text{ad}}^2 \{ \tilde{y}, \dot{y} \} = \frac{1}{K_{\text{ad}}} \cdot q_{\min} = \frac{1}{M - Q} \cdot \sum_{q=0}^{N-1} \sum_{j=1}^{m_q} (\tilde{y}_{q,j} - \dot{y}_q)^2, \quad (545)$$

где Q — количество коэффициентов в модели линейной регрессии, $K_{\text{ad}} = M - Q$ — число её степеней свободы ($Q \leq K+1$ или $Q \leq K$).

Теоретически, с учётом формулы (513), оно составляет значение

$$K_{ad} = M - Q = \text{rang } \vec{V}V', \quad (546)$$

где V есть $N \times Q$ -матрица планирования эксперимента, $\text{rang } V = Q$.

В итоге по формуле (545) вычисляется в общем случае *смещённая систематически* выборочная дисперсия адекватности $S_{ad}^2 \{ \tilde{y}, \dot{y} \}$ модели и функции. Причиной такого смещения модели является то, что $\mu \{ \dot{y} \}$ может заметно отличаться от реальной $y = \mu \{ \tilde{y} \}$. Отсюда числитель в формуле (544), как статистика, имеет в общем случае *нецентральное* χ^2 -распределение Пирсона с числом степеней свободы $K_{ad} = M - Q$.

$$\mu \{ S_{ad}^2 \} = \sigma^2 \{ \tilde{y} \} + \frac{1}{M - Q} \cdot \sum_{q=0}^{N-1} m_q \cdot \Delta^2 y_q = \sigma^2 \{ \tilde{y} \} + \Delta^2 \langle y_{(N)} \rangle, \quad (547)$$

где $\Delta y_q = \left[\mu \{ \dot{y}_q \} - y_q \right]$ — частная систематическая ошибка для $\dot{y}_q$;

$$\Delta^2 \langle y_{(N)} \rangle = \frac{1}{M - Q} \cdot \mu \left\{ \tilde{y}' \cdot \vec{V}V' \cdot \tilde{y} \right\} = \frac{1}{M - Q} \cdot y' \cdot \vec{V}V' \cdot y, \quad (548)$$

т. е. её квадрат средней систематической ошибки на $\langle x_q \rangle$, или плановая дисперсия. Но знаменатель в (544), как статистика, тут имеет *центральное* χ^2 -распределение Пирсона с числом степеней свободы $K_r = M - N$:

$$\mu \{ S_r^2 \} = \sigma^2 \{ y \}. \quad (549)$$

Если бы в дисперсии (545) отсутствовала систематическая ошибка (548), то и статистика $\hat{F}$ в (544) имела *центральное* F-распределение. Но поскольку целевая функция отклика $\tilde{y}(x)$ аппроксимируется разностной моделью всего лишь до 2-го порядка, то систематическая ошибка в (545) практически есть всегда — либо больше, либо меньше. Отсюда и критериальная статистика $\hat{F}$, вычисляемая по формуле (544), имеет в общем случае также *нецентральное* F-распределение.

Казалось бы, критериальное неравенство $\hat{F} \leq F^*$ (с доверительной вероятностью $\mathcal{P}$) действительно могло бы применяться для оценки пригодности получаемой регрессионной модели $\dot{y} = \dot{y}(x)$ в целях дальнейшей оптимизации функции отклика. Однако это вовсе не так. Покажем, что именно для данной роли оно не годится.

В математической статистике под адекватностью чего-нибудь чему-нибудь понимают их тождественность с достаточно высокой вероятностью $\mathcal{P}$, но лишь в пределах отвечающей ей допустимой случайной ошибки. Тогда с увеличением случайной ошибки в (545) и (540) вероятность адекватности модели и функции будет только возрастать. Например, из (544), (540), (545), (548), (549) следует: чем больше $\sigma^2\{\tilde{y}\}$, тем лучше по критерию $\hat{F}$ адекватность модели; и наоборот, чем меньше $\sigma^2\{\tilde{y}\}$, тем хуже по критерию $\hat{F}$ адекватность модели. Вклад систематической ошибки в общую в первом случае уменьшается, а во втором случае увеличивается. В пределе даже имеем $\lim_{\sigma \rightarrow \infty} \hat{F} = \infty$, так как $\lim_{\sigma \rightarrow 0} s_r^2 = 0$, $\lim_{\sigma \rightarrow 0} s_{ad}^2 = \Delta^2 \langle y_{(N)} \rangle$. Но это уже полный абсурд! Как маргинальный пример, он выражает здесь то, что планово-вычислительные методы, рассмотренные в гл. 6, применяют заведомо неадекватные по $\hat{F}$ модели 1-го и 2-го порядка (для них формально $\hat{F} = \infty$, т. е. всегда $\hat{F} > F^*$), но, тем не менее, все эти модели применимы для оптимизации $y(\mathbf{x})$.

Иными словами, при традиционном планировании эксперимента имеется парадокс. Адекватная по F -критерию и поэтому пригодная регрессионная модель становится вдруг неадекватной по F -критерию и поэтому непригодной при уменьшении её суммарной ошибки за счёт повышения точности эксперимента (например, вследствие увеличения количества опытов в точках плана). И обратно — неадекватную модель можно сделать адекватной при увеличении её суммарной ошибки за счёт уменьшения точности эксперимента!? Избежать этого несуразного несоответствия можно, если принять нижеследующую концепцию:

- 1) различие в значениях функции и её регрессионной модели в каждой точке плана и на их полном множестве составляется из случайной и систематической ошибок, что статистически в пределе выражает (547) с частными вкладами средних квадратичных отклонений двух типов;
- 2) точность регрессионной модели при необходимости надо повышать за счёт уменьшения явно большей из ошибок двух типов, что должно оцениваться статистически через анализ их частных вкладов;
- 3) пригодность полученной конкретной регрессионной модели для оптимизации функции отклика $\tilde{y}(\mathbf{x})$ нужно оценивать по относительной ошибке модели на области плана типа (471), имеющей универсальную применимость — как в планово-вычислительных, так и в планово-экспериментальных процедурах оптимизации.

Начнём с пункта 3. Чтобы его реализовать практически, определим относительную ошибку регрессионной модели аналогично (471) как

$$\xi = \frac{\sqrt{\frac{1}{N} \cdot \sum_{q=0}^{N-1} \sum_{j=1}^{m_q} (\tilde{y}_{q,j} - \dot{y}_q)^2}}{\bar{\tilde{y}}_{\max} - \bar{\tilde{y}}_{\min}} \cdot 100\% \approx \frac{s_{\text{ad}}\{\tilde{y}, \dot{y}\}}{\bar{\tilde{y}}_{\max} - \bar{\tilde{y}}_{\min}} \cdot 100\%, \quad (550)$$

где $\bar{\tilde{y}}_{\max}$ и $\bar{\tilde{y}}_{\min}$ — максимальное и минимальное выборочные средние на множестве точек плана эксперимента $\langle x_q \rangle$. Нетрудно видеть, что при таком подходе к оценке адекватности модели точность планирования эксперимента согласована корректно с вкладами ошибок обоих типов. А именно, критерий ξ уменьшается при уменьшении и случайной, и систематической ошибки — порознь или вместе и соответственно обратно. Отсюда видна его применимость и в планово-вычислительных методах оптимизации (см. гл. 6), где случайная ошибка практически нулевая. Отсюда же (471) является как бы предельным аналогом (550).

Далее обратимся к пункту 2. Если критерий ξ слишком большой по величине, например, $\xi > 5\%$, то затем на основе полученных статистических данных необходимо выяснить: какая же из этих двух ошибок вносит больший вклад в общую ошибку? (Обе ошибки в сумме своих долей дают 1.) Логика здесь такова. Если бы общая ошибка была тут чисто случайной, то равновероятным являлось $s_r^2 = s_{\text{ad}}^2$ ($\hat{F} = 1$), т. е. при $\mathcal{P} = 0,50$. Напротив, если бы общая ошибка была тут чисто систематической, то было $s_r = 0$. Отсюда с условием равновероятности весьма корректной оценкой относительных вкладов обеих этих ошибок являются долевые коэффициенты вида:

$$\tilde{\gamma} = s_r^2 / s_{\text{ad}}^2, \quad \hat{\gamma} = (s_{\text{ad}}^2 - s_r^2) / s_{\text{ad}}^2 \quad (\tilde{\gamma} + \hat{\gamma} = 1). \quad (551)$$

Снижение доли систематической ошибки достигают уменьшением интервалов варьирования частных факторов в плане эксперимента. (Это отвечает сужению области плана.) Снижение доли случайной ошибки достигают, наоборот, увеличением интервалов варьирования частных факторов в плане (или, если это достигается не слишком затратно, — увеличением количества экспериментов в точках плана).

Разумеется, все подобные действия по повышению точности аппроксимации выполняют только тогда, когда модель регрессии $\dot{y}(\mathbf{x})$ имеет неудовлетворительную величину относительной ошибки (550), например, более 5%. Здесь попросту сохраняется аналогия с обычно принимаемым в планировании эксперимента уровнем доверительной вероятности $\mathcal{P} = 0,95$ и соответственно ненадёжности $\alpha = 1 - \mathcal{P} = 0,05$.

Если же в (551) $\tilde{\gamma}$ и $\hat{\gamma}$ близки друг к другу, но относительная погрешность модели (550) довольно высокая, то изменение интервалов варьирования факторов в плане неэффективно. В таком случае нужно повышать точность экспериментальных данных, т. е. либо увеличивать количество опытов (измерений, наблюдений) q_j в точках плана, либо повышать качество проведения экспериментов.

Рассмотрев общие и смежные вопросы, относящиеся к главному предмету обсуждения данной главы, перейдём к основным процедурам планирования эксперимента при поиске экстремума функции отклика. В принципе, здесь возможны два вида аппроксимационного подхода, а именно, чисто разностный и общий регрессионный.

Первый подход целесообразен только в том случае, когда достаточно уверенно можно прогнозировать, что относительный вклад случайной ошибки в общую ошибку модели будет достаточно мал. В этом, скажем, уникальном варианте оптимизацию функции отклика целесообразно осуществлять, применяя планы, ранее уже описанные в § 6.2, но с *единственным* экспериментальным значением в каждой точке. При этом *минимальные планы* характеризуются тем, что в них количество точек (тут ещё и общее количество экспериментов) равно максимально возможному количеству ненулевых коэффициентов в соответствующей им разностной модели целевой функции отклика $\hat{y} = \hat{y}(\mathbf{x})$ — либо 1-го, либо 2-го порядка: $N = K + 1 = N_{\min}$. Например, убедиться в том, что вклад случайной ошибки слишком мал можно уже на самом начальном этапе оптимизации, а затем перейти к использованию вышеуказанных планов. К аналогичному прогнозу можно прийти исходя из величины реально достигаемой относительной ошибки измерения (наблюдения) $\alpha_2 = 100s/\tilde{y}\%$ (см. § 7.3). Так, если случайная ошибка находится лишь на уровне сотых долей процента от величины оценки $\tilde{y}$, то, в сравнении с принятым допустимым уровнем (5%) для относительной ошибки (550), это составляет $\sim 1\%$, что в таком варианте вполне допустимо.

Второй подход является вполне обычным и даже общепринятым. Он исходит из линейной регрессии функции отклика на области плана эксперимента (см. § 7.2) с последовательным использованием её тех же самых разностных моделей — либо 1-го, либо неполного 2-го, либо полного 2-го порядка. Но коэффициенты этих полиномиальных моделей содержат дополнительно случайные ошибки. Ввиду такого снижения точности аппроксимации для функции в планово-экспериментальных методах в сравнении с ней же в планово-вычислительных методах, применение разностной модели функции отклика 1-го порядка здесь не оправдано. (Скорость сходимости градиентного метода и так мала.)

В *нормированных* планах центр приведённых координат есть $\mathbf{t}_0 = \mathbf{0}$, т. е. исходная переменная $\mathbf{x}$ нормируется так, чтобы она обнулялась в центре плана. (Центр плана может являться независимой проверочной точкой при обосновании статистических гипотез, например, гипотезы об однородности нормального распределения ошибки эксперимента.) Переменные x_i нормируют по формулам приведения в переменные t_i :

$$t_i = \frac{x_i - x_{i(0)}}{\Delta x_i} \quad \leftrightarrow \quad x_i = x_{i(0)} + \Delta x_i \cdot t_i, \quad (552)$$

где Δx_i — частные вариации переменных x_i (интервалы варьирования). Их смысл вполне очевиден, если обратиться, например, к рис. 13 — 15 (гл. 6). Тогда все частные нормированные переменные t_i изменяются только на шаг с интервалом +1 или -1 (см. далее на рис. 16 — 18).

При *централизованном регрессионном* планировании эксперимента, с целью изъятия в модели регрессии свободного члена, функция $y(\mathbf{x})$ нормируется так, чтобы в центре плана она была заведомо нулевой. Функция y от $\mathbf{x}$ последовательно нормируется по формулам приведения сначала в функцию от нормированной переменной $\mathbf{t}$, затем в функцию от вектор-фактора $\mathbf{u}$ (см. § 7.2):

$$\begin{aligned} \hat{y} = \hat{y}(\mathbf{x}) &\equiv \hat{\varphi}(\mathbf{t}) \rightarrow \Delta \hat{y} = \hat{\varphi}(\mathbf{t}) - \varphi(\mathbf{0}) = \hat{\psi}(\mathbf{t}) \equiv \hat{l}(\mathbf{u}) = \mathbf{u}' \cdot \hat{\mathbf{b}} \rightarrow \\ &\rightarrow \dot{y} = \dot{y}(\mathbf{x}) \equiv \dot{\varphi}(\mathbf{t}) \rightarrow \Delta \dot{y} = \dot{\varphi}(\mathbf{t}) - \tilde{\varphi}(\mathbf{0}) = \dot{\psi}(\mathbf{t}) \equiv \dot{l}(\mathbf{u}) = \mathbf{u}' \cdot \dot{\mathbf{b}}, \end{aligned} \quad (553)$$

где формально $\varphi(\mathbf{0}) = y_0 = g_0$ и поэтому $\hat{y} = y_0 + \Delta \hat{y}$, $\dot{y} = \bar{y}_0 + \Delta \dot{y}$.

Через (552), (553) частные переменные x_i , t_i и факторы u_k , а также коэффициенты $\dot{g}_i$, $\dot{g}_{ii}$, $\dot{g}_{ij}$ полиномиальной модели и коэффициенты $\dot{b}_k$ модели линейной регрессии $\Delta \dot{y} = \dot{l}(\mathbf{u})$ однозначно взаимосвязаны. Напомним: в планировании эксперимента принимают, что переменные (факторы) для целевой функции в точках плана всегда задаются точно.

Далее рассматриваются ортогональные и субортогональные планы и процедуры планирования эксперимента при оптимизации функции отклика исходя из минимизации количества точек в плане эксперимента и общего количества опытов (наблюдений, измерений) в точках при обеспечении необходимой точности регрессионной аппроксимации функции $y(\mathbf{x})$ её разностными моделями как неполного 2-го, так и полного 2-го порядка. (Планирование с разностными моделями 1-го порядка, естественно, здесь не рассматривается из-за низкой скорости сходимости градиентного метода и её непредсказуемости, в силу высоковероятных проблем, связанных с овражной сходимостью.)

§ 7.5. Экспериментальная оптимизация неполного 2-го порядка по плану осевой n -крест

В указанном аспекте довольно эффективен масштабно-градиентный метод, который хорошо преодолевает проблемы, вызванные *овражной сходимостью* (см. §§ 5.9 и 6.4). Исходя из этого метода плано-экспериментальная оптимизация $\tilde{y}(\mathbf{x})$ в координатном пространстве $\mathcal{A}^n$ базируется на аппроксимации целевой функции отклика $y(\mathbf{x})$ разностной моделью неполного 2-го порядка, но с коэффициентами, полученными в виде регрессионных оценок:

$$\begin{aligned} y(\mathbf{x}) &\approx \dot{y} = \dot{y}(\mathbf{x}) = \dot{g}_0 + \sum_{i=1}^n \dot{g}_i \cdot (x_i - c_i) + \frac{1}{2} \cdot \sum_{i=1}^n \dot{g}_{ii} \cdot (x_i - c_i)^2 = \\ &= \dot{g}_0 + \dot{\mathbf{g}} \cdot (\mathbf{x} - \mathbf{c}) + \frac{1}{2} \cdot (\mathbf{x} - \mathbf{c})' \cdot \dot{\mathbf{D}} \cdot (\mathbf{x} - \mathbf{c}) \equiv \dot{\phi}(\mathbf{t}) = \\ &= \bar{y}_0 + \Delta \dot{y} = \bar{y}_0 + \dot{\psi}(\mathbf{t}) \equiv \bar{y}_0 + \dot{l}(\mathbf{u}) = \bar{y}_0 + \mathbf{u}' \cdot \dot{\mathbf{b}}, \end{aligned} \quad (554)$$

где $\dot{g}_0 = \bar{y}_0$ — *несмещённая* оценка начального коэффициента для любой степенной модели, применяемая при *централизованной* регрессии,

$\dot{g}_i = \frac{\Delta \tilde{y}}{\Delta x_i}$ — элементы регрессионного $1 \times n$ -вектора градиента $\dot{\mathbf{g}}$,

$\dot{g}_{ii} = \frac{\Delta^2 \tilde{y}}{\Delta x_i^2}$ — элементы регрессионной $n \times n$ -матрицы Гессе $\dot{\mathbf{D}}$.

Наиболее естественный план эксперимента для реализации данного метода — это *осевой n -крест*. Все точки такого плана расположены по осям координат в парных направлениях, одна точка $\mathbf{x}_0$ находится в центре — см. рис. 14 (гл. 6). Например, он сразу же может применяться как *минимальный* план эксперимента тогда, когда случайная ошибка модели по прогнозу достаточно мала (см. § 7.4). С целью придания плану эксперимента свойства *ортогональности* (§ 7.2) центральная точка $\mathbf{x}_0$ в основном расчёте не принимается во внимание, а функцию отклика нормируют, согласно (553). Эта точка служит для независимой оценки свободного члена $\dot{g}_0$ (при нормировании функции отклика), а также для проверки принимаемых статистических гипотез. Тогда вычисление коэффициентов регрессии осуществляется по упрощённым формулам (504), (514). В регрессионной модификации масштабно-градиентного метода, как будет видно далее, это тождественно их вычислению по скалярным формулам типа (475).

В нормированном варианте план осевой n-крест имеет упрощённый вид, изображённый иллюстративно на рис. 16 при $n = 2$. Все точки $\mathbf{t}_q$ находятся на осях координат на отметках $+1$ и -1 и одна точка $\mathbf{t}_0 = \mathbf{0}$ — в центре. Схема нормированного плана эксперимента осевой n-крест и движения к экстремуму целевой функции по данному плану показаны условно на рис. 16 при $n = 2$ в декартовом базисе.

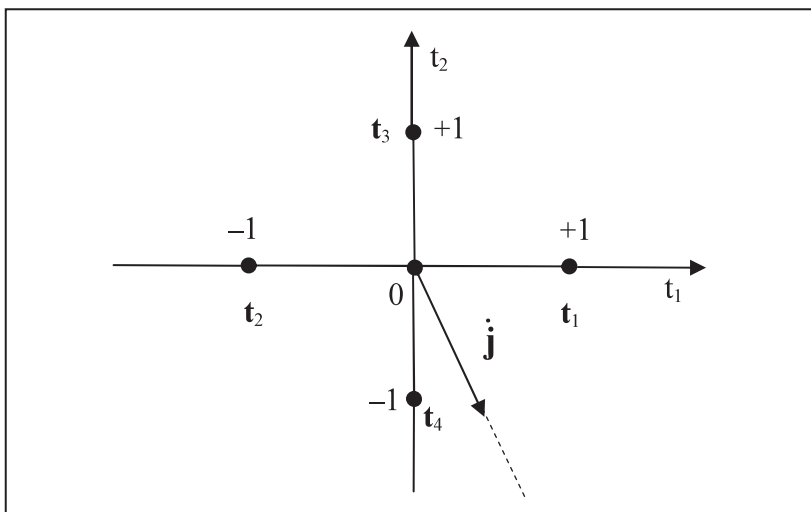


Рис. 16. Поиск экстремума целевой функции $\tilde{\varphi}(\mathbf{t}) \equiv \tilde{y}(\mathbf{x})$ при $n = 2$ по нормированному плану осевой n-крест.

Обратим внимание также на то, что и здесь директивный вектор $\dot{\mathbf{j}}$ неполного 2-го порядка находится всегда в тех же самых квадрантах, что и векторы градиента $\dot{\mathbf{g}}$ (при поиске максимума) или антиградиента $-\dot{\mathbf{g}}$ (при поиске минимума), что тут обусловлено чисто масштабным характером их взаимосвязи. (Например, на рис. 16 он условно находится в 4-м квадранте.) В данном примере имеем:

$\mathbf{t}_{(0)} = \mathbf{0} = (0, 0)$, $\mathbf{t}_{(1)} = (+1, 0)$, $\mathbf{t}_{(2)} = (-1, 0)$, $\mathbf{t}_{(3)} = (0, +1)$, $\mathbf{t}_{(4)} = (0, -1)$ — план;

$\Delta t_1 = \pm 1$ и $\Delta t_2 = \pm 1$ — частные вариации переменных t_i ;

$\tilde{\varphi}(\mathbf{t}_{(0)}) = \tilde{y}_0 = \dot{\mathbf{g}}_0$ — оценка начального коэффициента модели (554);

$\tilde{\varphi}(\mathbf{t}_{(1)}) = \tilde{y}_1$, $\tilde{\varphi}(\mathbf{t}_{(2)}) = \tilde{y}_2$, $\tilde{\varphi}(\mathbf{t}_{(3)}) = \tilde{y}_3$, $\tilde{\varphi}(\mathbf{t}_{(4)}) = \tilde{y}_4$ — значения целевой функции в точках плана. Причём, согласно (553), (554), имеем:

$$\dot{y} = \dot{y}(\mathbf{x}) \equiv \dot{\varphi}(\mathbf{t}) \leftrightarrow \dot{y} = \bar{\dot{y}}_0 + \Delta \dot{y}; \quad \Delta \dot{y} = \dot{\psi}(\mathbf{t}) \equiv \dot{l}(\mathbf{u}) = \mathbf{u}' \cdot \dot{\mathbf{b}}.$$

В частности, при $n = 2$, с учётом (554), имеем:

$$\begin{aligned}\dot{\mathbf{y}} &= \dot{\mathbf{y}}(\mathbf{x}) = \dot{\mathbf{g}}_0 + \dot{\mathbf{g}}_1 \cdot (\mathbf{x}_1 - \mathbf{c}_1) + \dot{\mathbf{g}}_2 \cdot (\mathbf{x}_2 - \mathbf{c}_2) + \\ &+ \frac{1}{2} \cdot \dot{\mathbf{g}}_{11} \cdot (\mathbf{x}_1 - \mathbf{c}_1)^2 + \frac{1}{2} \cdot \dot{\mathbf{g}}_{22} \cdot (\mathbf{x}_2 - \mathbf{c}_2)^2 \equiv \\ &\equiv \dot{\Phi}(\mathbf{t}) = \dot{\mathbf{g}}_0 + \dot{\mathbf{a}}_1 \cdot \mathbf{t}_1 + \dot{\mathbf{a}}_2 \cdot \mathbf{t}_2 + \frac{1}{2} \dot{\mathbf{a}}_{11} \cdot \mathbf{t}_1^2 + \frac{1}{2} \dot{\mathbf{a}}_{22} \cdot \mathbf{t}_2^2 \equiv \\ &\equiv \dot{l}(\mathbf{u}) = \bar{\mathbf{y}}_0 + \mathbf{u}' \cdot \dot{\mathbf{b}} = \bar{\mathbf{y}}_0 + \dot{\mathbf{b}}_1 \cdot \mathbf{u}_1 + \dot{\mathbf{b}}_2 \cdot \mathbf{u}_2 + \dot{\mathbf{b}}_3 \cdot \mathbf{u}_3 + \dot{\mathbf{b}}_4 \cdot \mathbf{u}_4.\end{aligned}$$

При *централизованной* линейной регрессии компоненты вектора $\dot{\mathbf{b}}$, или коэффициенты $\dot{\mathbf{b}}_k$ (где индекс k пробегает все значения от 1 до K) вычисляются при $n \geq 2$ по тождественным формулам типа (504), (514):

$$\dot{\mathbf{b}} = \left\{ \sum_{q=1}^{N-1} \mathbf{u}_q \mathbf{u}_q' \right\}^{-1} \cdot \left(\sum_{q=1}^{N-1} \Delta \tilde{\mathbf{y}}_q \mathbf{u}_q \right) = (\mathbf{U}' \cdot \mathbf{U})^{-1} \cdot \mathbf{U}' \cdot \Delta \tilde{\mathbf{y}}, \quad (K = 2n),$$

где $\Delta \tilde{\mathbf{y}}_q = \tilde{\mathbf{y}}_q - \bar{\mathbf{y}}_0$. Причём вторая из этих двух формул значительно проще, так как $(N-1) \times K$ -матрица планирования $\mathbf{U}$ для нормированного плана осевой n -крест (где $\Delta \tilde{\mathbf{y}}_0 = 0$) при $N-1 \geq K$ и $n \geq 2$ ортогональная. Например, при размерности $n = 2$ имеем 4×4 -матрицу планирования $\mathbf{U}$ структуры:

$\mathbf{U} = \{\mathbf{u}_q\}$	$\mathbf{u}_1 = \mathbf{t}_1$	$\mathbf{u}_2 = \mathbf{t}_2$	$\mathbf{u}_3 = \mathbf{t}_1^2$	$\mathbf{u}_4 = \mathbf{t}_2^2$	(555)
$\mathbf{u}_1$	+1	0	+1	0	
$\mathbf{u}_2$	-1	0	+1	0	
$\mathbf{u}_3$	0	+1	0	+1	
$\mathbf{u}_4$	0	-1	0	+1	

Отсюда $K \times K$ информационная и ковариационная матрицы вычисляются весьма просто: $\mathbf{W} = \mathbf{U}'\mathbf{U} = 2 \cdot \{\mathbf{I}\}$, $\mathbf{S} = \mathbf{W}^{-1} = \{\mathbf{I}\} / 2$. Вообще же при любой размерности задачи $n \geq 2$ матрица планирования и минимальная, и ортогональная: $N_{\min} = 2n+1 = K+1$, $M_{\min} = K + m_0$; $\mathbf{U}'\mathbf{U} = 2 \cdot \{\mathbf{I}\} = \mathbf{U}\mathbf{U}'$ ($m_q = 1$; $q = 1 \div K$). При этом она квадратная и имеет размер $2n \times 2n$.

Это легко проверить, если в матрице (555) выполнить перестановку столбцов 2 и 3, получив 2-х клеточную матрицу планирования U:

$U = \{u_q\}$	$u_1 = t_1$	$u_3 = t_1^2$	$u_2 = t_2$	$u_4 = t_2^2$
u_1	+1	+1	0	0
u_2	-1	+1	0	0
u_3	0	0	+1	+1
u_4	0	0	-1	+1

(556)

С возрастанием размерности n матрица планирования дополняется подобными диагональными 2×2 -блоками. С учётом (552) и того, что $\dot{y} = \tilde{y}_0 + \Delta \dot{y}$, из (514) для компонентов вектора градиента и матрицы Гессе легко получаются всегда однотипные для любого $n \geq 2$ скалярные формулы — аналоги (475). Имеем: $\dot{g}_0 = \tilde{y}_0$,

$$\left. \begin{aligned} \dot{g}_i &= \frac{\dot{a}_i}{\Delta x_i} = \frac{\tilde{y}_{1(i)} - \tilde{y}_{2(i)}}{2\Delta x_i}, \\ \frac{1}{2} \dot{g}_{ii} &= \frac{\dot{a}_{11}}{2\Delta x_i^2} = \frac{\tilde{y}_{1(i)} + \tilde{y}_{2(i)} - 2\tilde{y}_0}{2\Delta x_i^2}. \end{aligned} \right\} \quad (557)$$

$$\dot{\mathbf{g}} = (\dot{g}_1, \dots, \dot{g}_n); \quad \dot{\mathbf{D}} = \begin{Bmatrix} \dot{g}_{11} & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & \dot{g}_{nn} \end{Bmatrix}. \quad (558)$$

План на рис. 16 содержит 5 точек. В общем случае при $n \geq 1$ количество точек в плане осевой n-крест составляет $N = 2n + 1 = K + 1 = N_{\min}$, число опытов M определяется по формуле (542). В частности, при $n = 1$ процедура поиска экстремума сводится к одномерному планированию эксперимента 2-го порядка, т. е. к методу ньютоновского типа (§ 5.5). Теоретически из (520) и (557) нормальные дисперсии для статистически независимых здесь случайных ошибок коэффициентов составляют:

$$\begin{aligned} \sigma^2 \{\dot{g}_0\} &= \sigma^2 \{\tilde{y}\} / m_0, \quad \sigma^2 \{\dot{g}_i\} = \sigma^2 \{\tilde{y}\} / 2\Delta x_i^2, \\ \sigma^2 \{\dot{g}_{ii} / 2\} &= (1/m_0 + 1/2) \cdot \sigma^2 \{\tilde{y}\} / \Delta x_i^4. \end{aligned} \quad (559-561)$$

Коэффициенты вычисляются функционально независимо друг от друга.

Пошаговое движение к промежуточному или конечному экстремуму из точки $\mathbf{c}$ осуществляется при $n \geq 2$ по директивному вектору:

$$\dot{\mathbf{j}} = (-\dot{\mathbf{D}})^{-1} \cdot \dot{\mathbf{g}}' = - \begin{vmatrix} \dot{g}_1 / \dot{g}_{11} \\ \dots\dots\dots \\ \dot{g}_n / \dot{g}_{nn} \end{vmatrix}. \quad (562)$$

Этот вектор и радиус-вектор $\mathbf{x}$ по компонентам имеют одни и те же размерности. При изменениях масштабов по осям x_i они преобразуются ковариантно. Формально масштабы по осям здесь не влияют на скорость сходимости.

Отметим то обстоятельство, что укрупнением масштабов по осям x_i можно при необходимости добиваться пропорционального уменьшения отношений $\dot{g}_i / \dot{g}_{ii}$ в формуле (562), что важно, если знаменатели $\dot{g}_{ii}$ слишком тут малы и не точны. Причём с укрупнением масштабов знаменатели $\dot{g}_{ii}$ увеличиваются на порядок больше, нежели числители $\dot{g}_i$.

Весьма наглядным признаком движения к экстремуму функции $\tilde{y}(\mathbf{x})$ является знак всех $\dot{g}_{ii} \neq 0$. (Но разные знаки хотя бы для пары $\dot{g}_{ii}$ свидетельствуют о локальной седловине целевой функции.) Причём с приближением к искомому экстремуму $\tilde{y}(\mathbf{x})$ числители $\dot{g}_i$ неизбежно уменьшаются с замедлением вплоть до выполнения результирующего неравенства типа (449). Хорошим признаком эффективности метода является то, что приближение к экстремуму сопровождается только уменьшением значений всех $\dot{g}_i$, но все $\dot{g}_{ii}$ при этом весьма значимые.

Если матрица $\dot{\mathbf{D}}(\mathbf{x})$ в окрестности строгого экстремума $\tilde{y}(\mathbf{x})$ плохо обусловлена, т. е. при $\dot{\mathbf{g}} \rightarrow 0$ некоторые $\dot{g}_{ii} \approx 0$ (несмотря на укрупнение масштабов); или в окрестности нестрогого экстремума $\tilde{y}(\mathbf{x})$ вырождена, т. е. при $\dot{\mathbf{g}} \rightarrow 0$ некоторые $\dot{g}_{ii} = 0$, то этот метод в базовом варианте неэффективен. Тогда он приводит к слишком большому разбросу результата из-за неизбежных случайных ошибок в значениях целевой функции. Действенным способом разрешения этой проблемы может быть применение метода квадратичной регуляризации по Тихонову (см. § 5.10).

В качестве параметра квадратичной регуляризации, например, можно выбрать $\alpha = \max\{\dot{g}_i / \Delta x_i\} \rightarrow 0$. Тогда имеем достаточно устойчивую оценку директивного вектора и координат экстремума:

$$\dot{\mathbf{j}}_\alpha = -(\dot{\mathbf{D}} \pm \alpha \mathbf{I})^{-1} \cdot \dot{\mathbf{g}}'. \quad (563)$$

§ 7.6. Экспериментальная оптимизация 2-го порядка

Планово-экспериментальная оптимизация общеньютоновского типа (§ 5.10) в координатном пространстве $\mathcal{A}^n$ базируется на аппроксимации функции разностной моделью полного 2-го порядка с коэффициентами, полученными в виде регрессионных оценок:

$$\begin{aligned} y(\mathbf{x}) &\approx \dot{y}(\mathbf{x}) = \dot{g}_0 + \sum_{i=1}^n \dot{g}_i \cdot (x_i - c_i) + \frac{1}{2} \cdot \sum_{j=1}^n \sum_{i=1}^n \dot{g}_{ij} \cdot (x_i - c_i)(x_j - c_j) = \\ &= y(\mathbf{c}) + \dot{\mathbf{g}} \cdot (\mathbf{x} - \mathbf{c}) + \frac{1}{2} \cdot (\mathbf{x} - \mathbf{c})' \cdot \dot{\mathbf{G}} \cdot (\mathbf{x} - \mathbf{c}) \equiv \dot{\psi}(\mathbf{t}) = \\ &= \bar{y}_0 + \Delta \dot{y} = \bar{y}_0 + \dot{\psi}(\mathbf{t}) \equiv \bar{y}_0 + \dot{l}(\mathbf{u}) = \bar{y}_0 + \mathbf{u}' \cdot \dot{\mathbf{b}}, \end{aligned} \quad (564)$$

где $\dot{g}_0$ — *несмещённая оценка* начального коэффициента модели, полученная, как и ранее в модели (554), статистической оценкой, применяемой при централизованной регрессии; или *смещённая оценка*, находящаяся совместно с оценками других коэффициентов модели,

$\dot{g}_i = \frac{\Delta \tilde{y}}{\Delta x_i}$ — элементы разностного $1 \times n$ -вектора градиента $\dot{\mathbf{g}}$,

$\dot{g}_{ii} = \frac{\Delta^2 \tilde{y}}{\Delta x_i^2}$, $\dot{g}_{ij} = \frac{\Delta^2 \tilde{y}}{\Delta x_i \Delta x_j}$ — элементы разностной $n \times n$ -матрицы Гессе $\dot{\mathbf{G}}$

§ 7.6.1. Оптимизация по плану n-СКП

Наиболее наглядно для реализации процесса оптимизации целевой функции отклика с разностной моделью 2-го порядка выбрать n-СКП (*симметричный композиционный план*). Данный план — композиция из осевого плана n-крест и координатно плоскостного плана n-крест, а, по геометрической сути, — композиция из осевого плана n-крест и его 4-х угловых точек в каждой из C_n^2 координатных плоскостей $\langle x_i, x_j \rangle$. Но с целью приближения матрицы плана к ортогональности (см. далее), центральная точка $\mathbf{x}_0$ в основном расчёте не принимается во внимание, а функцию отклика нормируют, согласно (553). Точка $\mathbf{x}_0$ служит для независимой оценки члена g_0 (при нормировании функции), а также для проверки принимаемых статистических гипотез. Тогда, как было уже выше (§ 7.5), оценка коэффициентов модели 2-го порядка при *централизованной* регрессии осуществляется по формуле типа (514).

В нормированном варианте n -СКП имеет более упрощённый вид, отображённый на рис. 17 при $n = 2$. В данном случае он содержит как 4 осевые точки $\mathbf{t}_q$ на осях координат на отметках $+1$ и -1 , т. е. из плана осевой 2-крест, так и 4 угловые точки $\mathbf{t}_q$ на отметках $(\pm 1, \pm 1)$, т. е. из плана плоскостной 2-крест, а также одну точку в центре нового базиса.

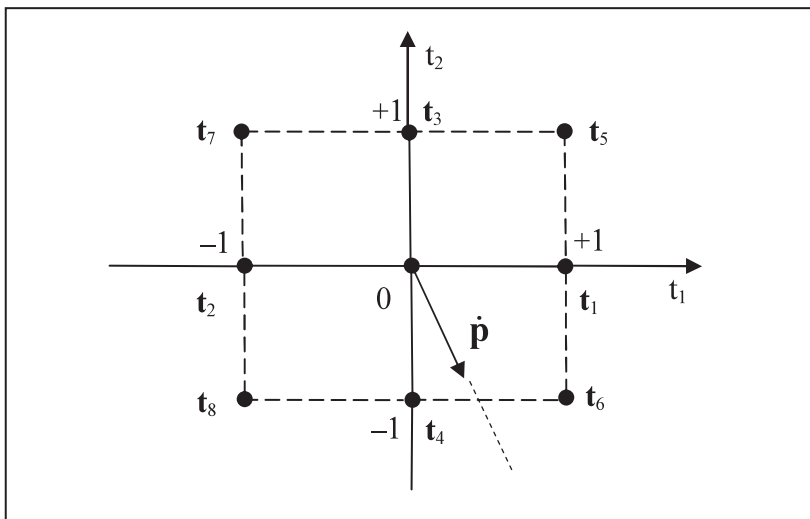


Рис. 17. Поиск экстремума целевой функции $\tilde{\varphi}(\mathbf{t}) \equiv \tilde{y}(\mathbf{x})$ при $n = 2$ по симметричному композиционному плану.

Директивный вектор $\dot{\mathbf{p}}$ находится, вероятнее, в тех же квадрантах, что и векторы градиента $\dot{\mathbf{g}}$ (при поиске максимума) или антиградиента $-\dot{\mathbf{g}}$ (при поиске минимума), или, по крайней мере, в одном из смежных с ними квадрантов. Это обусловлено тут всегда их знакоопределённым симметрично-линейным характером взаимосвязи. (На рис. 17 вектор $\dot{\mathbf{p}}$ условно находится в 4-м квадранте.) В данном примере имеем:

$\mathbf{t}_{(0)} = \mathbf{0} = (0, 0)$, $\mathbf{t}_{(1)} = (+1, 0)$, $\mathbf{t}_{(2)} = (-1, 0)$, $\mathbf{t}_{(3)} = (0, +1)$, $\mathbf{t}_{(4)} = (0, -1)$,

$\mathbf{t}_{(5)} = (+1, +1)$, $\mathbf{t}_{(6)} = (+1, -1)$, $\mathbf{t}_{(7)} = (-1, +1)$, $\mathbf{t}_{(8)} = (-1, -1)$ — план;

$\Delta t_1 = \pm 1$ и $\Delta t_2 = \pm 1$ — частные вариации переменных t_i ;

$\tilde{\varphi}(\mathbf{t}_{(0)}) = \tilde{y}_0 = \dot{\mathbf{g}}_0$ — оценка начального коэффициента модели (554);

$\tilde{\varphi}(\mathbf{t}_{(1)}) = \tilde{y}_1$, $\tilde{\varphi}(\mathbf{t}_{(2)}) = \tilde{y}_2$, $\tilde{\varphi}(\mathbf{t}_{(3)}) = \tilde{y}_3$, $\tilde{\varphi}(\mathbf{t}_{(4)}) = \tilde{y}_4$,

$\tilde{\varphi}(\mathbf{t}_{(5)}) = \tilde{y}_5$, $\tilde{\varphi}(\mathbf{t}_{(6)}) = \tilde{y}_6$, $\tilde{\varphi}(\mathbf{t}_{(7)}) = \tilde{y}_7$, $\tilde{\varphi}(\mathbf{t}_{(8)}) = \tilde{y}_8$ — значения целевой функции в точках плана. Причём, согласно (553), (554), имеем:

$$\dot{y} = \dot{y}(\mathbf{x}) \equiv \dot{\varphi}(\mathbf{t}) \leftrightarrow \dot{y} = \tilde{y}_0 + \Delta \dot{y}; \quad \Delta \dot{y} = \dot{\psi}(\mathbf{t}) \equiv \dot{l}(\mathbf{u}) = \mathbf{u}' \cdot \dot{\mathbf{b}}.$$

В частности, при $n = 2$, с учётом (564), имеем:

$$\begin{aligned}\dot{y} &= \dot{y}(\mathbf{x}) = \dot{g}_0 + \dot{g}_1 \cdot (x_1 - c_1) + \dot{g}_2 \cdot (x_2 - c_2) + \\ &+ \frac{1}{2} \cdot \dot{g}_{11} \cdot (x_1 - c_1)^2 + \frac{1}{2} \cdot \dot{g}_{22} \cdot (x_2 - c_2)^2 + \dot{g}_{12} \cdot (x_1 - c_1) \cdot (x_2 - c_2) \equiv \\ &\equiv \dot{\phi}(\mathbf{t}) = \dot{g}_0 + \dot{a}_1 \cdot t_1 + \dot{a}_2 \cdot t_2 + \frac{1}{2} \cdot \dot{a}_{11} \cdot t_1^2 + \frac{1}{2} \cdot \dot{a}_{22} \cdot t_2^2 + \dot{a}_{12} \cdot t_1 \cdot t_2 \equiv \\ &\equiv \dot{l}(\mathbf{u}) = \bar{y}_0 + \mathbf{u}' \cdot \dot{\mathbf{b}} = \bar{y}_0 + \dot{b}_1 \cdot u_1 + \dot{b}_2 \cdot u_2 + \dot{b}_3 \cdot u_3 + \dot{b}_4 \cdot u_4 + \dot{b}_5 \cdot u_5.\end{aligned}$$

При *централизованной* линейной регрессии компоненты вектора $\dot{\mathbf{b}}$, или коэффициенты $\dot{b}_k$ (где индекс k пробегает все значения от 1 до K) вычисляются при $n \geq 2$ по тождественным формулам типа (504), (514):

$$\dot{\mathbf{b}} = \left\{ \sum_{q=1}^{N-1} \mathbf{u}_q \mathbf{u}_q' \right\}^{-1} \cdot \left(\sum_{q=1}^{N-1} \Delta \tilde{y}_q \mathbf{u}_q \right) = (\mathbf{U}' \cdot \mathbf{U})^{-1} \cdot \mathbf{U}' \cdot \Delta \tilde{\mathbf{y}} \left(K = \frac{n \cdot (n+3)}{2} \right),$$

где $\Delta \tilde{y}_q = \tilde{y}_q - \bar{\tilde{y}}_0$ ($\Delta \tilde{y}_0 = 0$). Причём $(N-1) \times K$ -матрица планирования $\mathbf{U}$ для n -СКП при любом $n \geq 2$ *субортогональная* из-за неортогональности её столбцов с факторами t_1^2 . Например, при размерности $n = 2$ имеем 8×5 -матрицу планирования $\mathbf{U}$ нижеуказанной структуры:

$\mathbf{U} = \{\mathbf{u}_q\}$	$u_1 = t_1$	$u_2 = t_2$	$u_3 = t_1^2$	$u_4 = t_2^2$	$u_5 = t_1 \cdot t_2$
$\mathbf{u}_1$	+1	0	1	0	0
$\mathbf{u}_2$	-1	0	1	0	0
$\mathbf{u}_3$	0	+1	0	1	0
$\mathbf{u}_4$	0	-1	0	1	0
$\mathbf{u}_5$	+1	+1	1	1	+1
$\mathbf{u}_6$	+1	-1	1	1	-1
$\mathbf{u}_7$	-1	+1	1	1	-1
$\mathbf{u}_8$	-1	-1	1	1	+1

В самом же общем случае, т. е. при $n \geq 2$, матрица планирования U имеет блочную структуру вида:

$$U = \left\{ \begin{array}{cc} \mathbf{Q}_{2n \times 2n} & \mathbf{Z}_{2n \times C_n^2} \\ \mathbf{A}_{(1)4 \times 2n} & \mathbf{B}_{(1)4 \times C_n^2} \\ \mathbf{A}_{(2)4 \times 2n} & \mathbf{B}_{(2)4 \times C_n^2} \\ \dots & \dots \\ \mathbf{A}_{(C_n^2)4 \times 2n} & \mathbf{B}_{(C_n^2)4 \times C_n^2} \end{array} \right\} \quad (565)$$

$(N-1) \times K$ -матрица n -СКП (где $N-1 \geq K$) включает, кроме нулевого, ещё три независимых друг от друга *ортогональных* блока: Q , A и B . $2n \times 2n$ -блок Q отвечает осевому плану n -крест; в сумме $4C_n^2 \times 2n$ -блок A и $4C_n^2 \times C_n^2$ -блок B отвечают координатно-плоскостному плану n -крест. Независимый ортогональный блок B задаётся переменными типа $\{t_i \cdot t_j\}$. Матрицы блоков Q и B выделены в таблице (565) жирным шрифтом.

Вначале рассматриваем блоки Q и B как матрицы 2-х независимых централизованных ортогональных планов: осевого плана n -крест и координатно-плоскостного плана n -крест. Вычисление коэффициентов регрессии раздельно по этим планам обеспечивает для них наименьшие систематические ошибки (т. е. их смещения относительно истинных коэффициентов в разложении целевой функции в степенной ряд). Однако случайные ошибки для тех же коэффициентов регрессии незначительно уменьшаются в сравнении с таковой для целевой функции. Кроме того, ввиду ортогональности обоих раздельных планов тут довольно просто и дважды применяется формула (514) для вычисления коэффициентов. В итоге i и ii -е коэффициенты вычисляются по формулам (557) для осевого плана n -крест, ij -е коэффициенты вычисляются по формулам для координатно-плоскостного плана n -крест (в части блока B):

$$\dot{g}_{ij} = \dot{g}_{ji} = \frac{\dot{a}_{ij}}{\Delta x_i \cdot \Delta x_i} = \frac{(\tilde{y}_5 - \tilde{y}_6 - \tilde{y}_7 + \tilde{y}_8)_{(i,j)}}{4 \cdot \Delta x_i \cdot \Delta x_j}, \quad \sigma^2 \{\dot{g}_{ij}\} = \frac{\sigma^2 \{\tilde{y}\}}{4 \cdot \Delta x_i^2 \cdot \Delta x_j^2}. \quad (566)$$

Это упрощённый *первый вариант* вычисления коэффициентов регрессии по n -СКП. По степени простоты и наглядности он, пожалуй, самый простой из всех вариантов планирования эксперимента 2-го порядка. В итоге имеем:

$$\dot{\mathbf{g}} = (\dot{g}_i), \quad \dot{\mathbf{G}} = \{\dot{g}_{ij}\}.$$

Альтернативный ему *второй вариант* вычисления коэффициентов регрессии по n -СКП сводится к их совместному нахождению по формуле (514) исходя из матрицы плана в целом. Применение этой формулы здесь весьма просто, так как *субортогональная* матрица плана, хотя полностью не ортогональная, но производит довольно простую для последующего обращения *субдиагональную* информационную матрицу W :

$$W = \{U'U\}_{K \times K} = \left\{ (4n-2) \cdot I_{n \times n} \oplus [(4n-6) \cdot I_{n \times n} + 4 \cdot It_{n \times n}] \oplus 4 \cdot I_{C_n^2 \times C_n^2} \right\}, \quad (567)$$

где I — единичная матрица, It — тотально-единичная матрица, все элементы которой равны 1 (см. стр. 148). Причём, вполне очевидно, что $It_{n \times n}^2 = n \cdot It$. Матрица W со структурой (567) весьма просто обращается в ковариационную матрицу S с той же субдиагональной структурой:

$$S = W^{-1} = \{U'U\}_{K \times K}^{-1} = \left\{ (4n-2)^{-1} \cdot I_{n \times n} \oplus \left[\frac{b+c}{2d} \cdot I_{n \times n} - \frac{c}{2d} \cdot It_{n \times n} \right] \oplus 4^{-1} \cdot I_{C_n^2 \times C_n^2} \right\}, \quad (568)$$

где имеем следующие значения числовых параметров:

$$b = 4n - 5, c = 2, d = (4n - 5) \cdot (2n - 1) - (4n - 1) = (2n - 3) \cdot (4n - 3). \quad (569)$$

В (568) числовой параметр $2d$ есть, по сути, детерминант матрицы в квадратных скобках в (567). В свою очередь, числовые параметры b , c и d находятся в результате решения очевидного здесь уравнения:

$$[(4n-6) \cdot I_{n \times n} + 4 \cdot It_{n \times n}] \cdot \left[\frac{b+c}{2d} \cdot I_{n \times n} - \frac{c}{2d} \cdot It_{n \times n} \right] = I_{n \times n}.$$

Отсюда составляется и легко разрешается система из 2-х уравнений с 2-мя параметрами $b/2d$ и $c/2d$, указанными в (568) и (569):

$$\begin{cases} (4n-2) \cdot b - (n-1) \cdot 4 \cdot c = 2d, \\ -(4n-2) \cdot c + 4 \cdot e - (n-2) \cdot 4 \cdot c = 0. \end{cases}$$

Далее по n -СКП вычисляем коэффициенты регрессии и их дисперсии.

План на рис. 17 содержит 9 точек. В общем случае при $n \geq 2$ в n -СКП количество точек есть $N = 2n^2 + 1 > K + 1$; минимальное число опытов составляет $M_{\min} = 2n^2 + m_0$.

С учётом (552) и того, что $\dot{\mathbf{y}} = \bar{\mathbf{y}}_0 + \Delta\dot{\mathbf{y}}$, из (514) для компонентов вектора градиента и матрицы Гессе получаем все скалярные формулы.

При $n = 2$ имеем частные формулы: $\dot{\mathbf{g}}_0 = \bar{\mathbf{y}}_0$,

$$\left\{ \begin{aligned} \dot{\mathbf{g}}_1 &= \frac{\dot{\mathbf{a}}_1}{\Delta x_1} = \frac{(\tilde{y}_1 - \tilde{y}_2) + (\tilde{y}_5 + \tilde{y}_6 - \tilde{y}_7 - \tilde{y}_8)}{6\Delta x_1}, \\ \dot{\mathbf{g}}_2 &= \frac{\dot{\mathbf{a}}_2}{\Delta x_2} = \frac{(\tilde{y}_3 - \tilde{y}_4) + (\tilde{y}_5 + \tilde{y}_7 - \tilde{y}_6 - \tilde{y}_8)}{6\Delta x_2}, \\ \frac{1}{2}\dot{\mathbf{g}}_{11} &= \frac{\dot{\mathbf{a}}_{11}}{2\Delta x_1^2} = \frac{-6 \cdot \bar{\mathbf{y}}_0 + 3(\tilde{y}_1 + \tilde{y}_2) - 2(\tilde{y}_3 + \tilde{y}_4) + (\tilde{y}_5 + \tilde{y}_6 + \tilde{y}_7 + \tilde{y}_8)}{10\Delta x_1^2}, \\ \frac{1}{2}\dot{\mathbf{g}}_{22} &= \frac{\dot{\mathbf{a}}_{22}}{2\Delta x_2^2} = \frac{-6 \cdot \bar{\mathbf{y}}_0 + 3(\tilde{y}_3 + \tilde{y}_4) - 2(\tilde{y}_1 + \tilde{y}_2) + (\tilde{y}_5 + \tilde{y}_6 + \tilde{y}_7 + \tilde{y}_8)}{10\Delta x_2^2}, \\ \dot{\mathbf{g}}_{12} = \dot{\mathbf{g}}_{21} &= \frac{\dot{\mathbf{a}}_{12}}{\Delta x_1 \cdot \Delta x_2} = \frac{(\tilde{y}_5 - \tilde{y}_6) - (\tilde{y}_7 - \tilde{y}_8)}{4 \cdot \Delta x_1 \cdot \Delta x_2} = \frac{\tilde{y}_5 - \tilde{y}_6 - \tilde{y}_7 - \tilde{y}_8}{4 \cdot \Delta x_1 \cdot \Delta x_2}, \\ \dot{\mathbf{g}} &= (\dot{\mathbf{g}}_1, \dot{\mathbf{g}}_2); \quad \dot{\mathbf{G}} = \begin{Bmatrix} \dot{\mathbf{g}}_{11} & \dot{\mathbf{g}}_{12} \\ \dot{\mathbf{g}}_{12} & \dot{\mathbf{g}}_{22} \end{Bmatrix}. \end{aligned} \right.$$

При $n \geq 2$ имеем общие формулы (570): $\dot{\mathbf{g}}_0 = \bar{\mathbf{y}}_0$,

$$\left\{ \begin{aligned} \dot{\mathbf{g}}_i &= \frac{\dot{\mathbf{a}}_i}{\Delta x_i} = \frac{(\tilde{y}_{1(i)} - \tilde{y}_{2(i)}) + \sum_{m=1(m \neq i)}^n (\tilde{y}_5 + \tilde{y}_6 - \tilde{y}_7 - \tilde{y}_8)_{(m,i)}}{(4n-2) \cdot \Delta x_i}, \\ \frac{1}{2}\dot{\mathbf{g}}_{ii} &= \frac{\dot{\mathbf{a}}_{ii}}{2\Delta x_i^2} = -\frac{2n-1}{4n-3} \cdot \frac{\bar{\mathbf{y}}_0}{\Delta x_i^2} + \frac{b \cdot (\tilde{y}_{1(i)} + \tilde{y}_{2(i)}) - \sum_{j=1(j \neq i)}^n [c \cdot (\tilde{y}_{1(j)} + \tilde{y}_{2(j)}) + (b-c) \cdot (\tilde{y}_5 + \tilde{y}_6 + \tilde{y}_7 + \tilde{y}_8)_{(j,i)}] - 4 \cdot \sum_{m=j+1(m \neq i)}^n (\tilde{y}_5 + \tilde{y}_6 + \tilde{y}_7 + \tilde{y}_8)_{(j,m)}}{2d \cdot \Delta x_i^2}, \\ \dot{\mathbf{g}}_{ij} = \dot{\mathbf{g}}_{ji} &= \frac{\dot{\mathbf{a}}_{ij}}{\Delta x_i \cdot \Delta x_j} = \frac{(\tilde{y}_5 - \tilde{y}_6 - \tilde{y}_7 + \tilde{y}_8)_{(i,j)}}{4 \cdot \Delta x_i \cdot \Delta x_j}; \\ \dot{\mathbf{g}} &= (\dot{\mathbf{g}}_i), \quad \dot{\mathbf{G}} = \{\dot{\mathbf{g}}_{ij}\}. \end{aligned} \right.$$

(В процедуре 2-го порядка наибольшее влияние на точность вычисления директивных векторов и экстремума имеют i-ые и ii-ые коэффициенты.)

Теоретически из (518), (519) и (570) вычисляются все дисперсионные характеристики для случайных ошибок коэффициентов регрессии:

$$\left\{ \begin{aligned} \sigma^2 \{\dot{g}_0\} &= \frac{\sigma^2 \{\tilde{y}\}}{m_0}, \quad \sigma^2 \{\dot{g}_i\} = \frac{\sigma^2 \{\tilde{y}\}}{\Delta X_i^2 \cdot [2 + 4(n-1)]}, \\ \sigma^2 \left\{ \frac{\dot{g}_{ii}}{2} \right\} &= \frac{\sigma^2 \{\tilde{y}\}}{\Delta X_i^4} \cdot \left[\left(\frac{2n-1}{4n-3} \right)^2 / m_0 + \frac{b}{2d} \right], \quad \lambda \{\dot{g}_{ii}, \dot{g}_{jj}\} = -\frac{c}{b}; \\ \sigma^2 \{\dot{g}_{ij}\} &= \frac{\sigma^2 \{\tilde{y}\}}{4 \cdot \Delta X_i^2 \cdot \Delta X_j^2}. \end{aligned} \right. \quad (571-574)$$

Как видно, между случайными ошибками для ii и jj -х коэффициентов имеется небольшая линейная корреляция, падающая с ростом n .

§ 7.6.2. Оптимизация по плану n -ЦКП Бокса — Уилсона

До настоящего времени наиболее известным и распространённым для планирования эксперимента по модели (564) является, как известно, n -ЦКП Бокса — Уилсона (*центральный композиционный план*). Впервые он был предложен в пионерской публикации этих авторов в 1951 г. [54]. Данный план эксперимента — это композиция из плана n -куб (или его дробной m -реплики) и плана осевой n -крест с плечом $\pm \varepsilon$ по осям координат. При $n = 2$: $\pm \varepsilon = \pm 1$, план формально совпадает с 2-СКП, но матрица планирования U в сравнении с ним содержит дополнительно 1-й столбец с $t_0 = +1$. При $n > 2$: $|\pm \varepsilon| > 1$ и вычисляется по специальной формуле — см. далее. Число точек в плане составляет $N = 2^n + 2n + 1$.

Входящий в n -ЦКП Бокса — Уилсона составной частью план n -куб вошёл в научный обиход статистики как *полный факторный план* с появлением *дисперсионного анализа* [15], основанного в 1-й трети XX века Р. Фишером. План n -куб применяется в математической статистике для выявления существенных факторов типа $u_p = t_i$ и $u_p = t_i \cdot t_j$. (Частные факторы $u_p = t_i$ могут иметь и количественный, и качественный характер.) В плано-экспериментальной оптимизации функции отклика n -куб позволяет получать оценки только коэффициентов $\dot{g}_i$ и $\dot{g}_{ij}$ в её модели (564), но не позволяет находить важные оценки $\dot{g}_{ii}$. Отсюда и возникает необходимость его дополнения планом осевой n -крест. Тот же вывод относится к любой *дробной m -реплике* плана n -куб. Это есть его некая часть типа m -куб (при $m < n$). Для n -ЦКП она порождает *генерирующее соотношение*, применяемое для вычисления факторов взаимодействий. В статистике m -реплика называется *дробным факторным планом*.

При $n = 2$ имеем $\varepsilon = 1$. Здесь 2-ЦКП совпадает с планом 2-СКП. Он отображён на рис. 18 в соответствующих новых обозначениях.

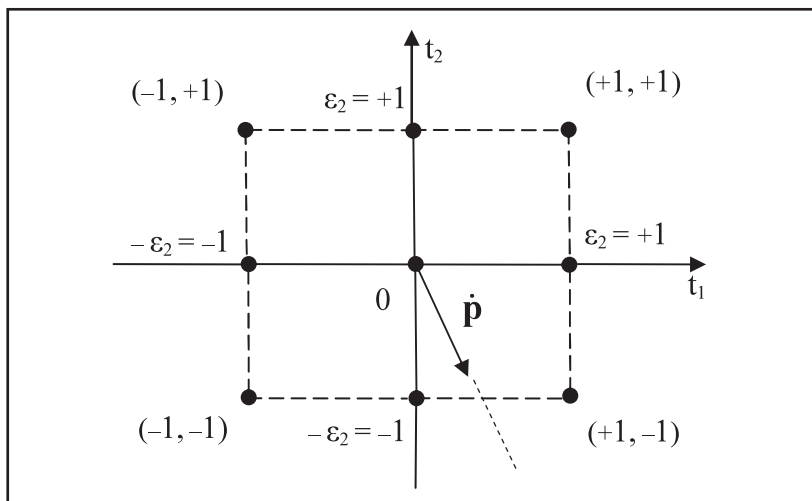


Рис. 18. Поиск экстремума целевой функции $\tilde{\varphi}(\mathbf{t}) \equiv \tilde{y}(\mathbf{x})$ при $n = 2$ по центральному композиционному плану Бокса — Уилсона.

Матрица планирования V (см. § 7.2) с переменными t_i характеризуется наличием столбца с $t_0 = +1$ и заменой t_i^2 на $(t_i^2 - d)$. Например, при $n = 2$:

$V = \{v_q\}$	t_0	t_1	t_2	$t_1^2 - d$	$t_2^2 - d$	$t_1 \cdot t_2$
v_0	+1	0	0	$-d$	$-d$	0
v_1	+1	$+\varepsilon$	0	$\varepsilon^2 - d$	$-d$	0
v_2	+1	$-\varepsilon$	0	$\varepsilon^2 - d$	$-d$	0
v_3	+1	0	$+\varepsilon$	$-d$	$\varepsilon^2 - d$	0
v_4	+1	0	$-\varepsilon$	$-d$	$\varepsilon^2 - d$	0
v_5	+1	+1	+1	$1 - d$	$1 - d$	+1
v_6	+1	+1	-1	$1 - d$	$1 - d$	-1
v_7	+1	-1	+1	$1 - d$	$1 - d$	-1
v_8	+1	-1	-1	$1 - d$	$1 - d$	+1

(575)

Исходная матрица планирования не ортогональная (при любом $n \geq 2$). Чтобы в общем случае перейти к ортогональной матрице планирования, в модели (564) и в матрице вместо факторов $u_p = t_1^2$ ($p = n + i$ лежат в интервале $n+1 \div 2n$) вводятся новые факторы $u_p^* = (t_1^2 - d)$, где $d = \sum_{q=1}^{N-1} t_{i(q)}^2 / N = (2^m + 2\varepsilon^2) / N$. При вышеуказанном преобразовании соответственно вместо $g_0 = a_0 = b_0$ вводят начальный коэффициент вида

$$\dot{a}_0^* = \dot{b}_0^* = \dot{a}_0 + d \cdot \sum_{i=1}^n \frac{1}{2} \dot{a}_{ii}. \quad (\text{Для дробного } n\text{-ЦКП: } N = 2^m + 2n + 1.)$$

Этот искусственный приём ортогонализации матрицы планирования устраняет нарушение независимости коэффициентов g_{ii} и g_{ij} в модели (564) именно по *n*-ЦКП Бокса — Уилсона. В результате обеспечивается независимость *всех* коэффициентов как для модифицированной модели линейной регрессии с новыми квадратичными факторами $u_p^* = (t_1^2 - d)$, так и для исходной модели (564). Дополнительный смысл подобной операции состоит в упрощении процедуры вычисления коэффициентов исходной модели путём использования матричной формулы типа (508).

С указанным подходом к проблеме начальная модель (564), с учётом нормирования переменных x_i по формуле (552), далее преобразуется в модифицированную модель линейной регрессии следующим образом

$$\begin{aligned} \dot{y}(x) &= \dot{g}_0 + \sum_{i=1}^n \dot{g}_i \cdot (x_i - c_i) + \frac{1}{2} \cdot \sum_{j=1}^n \sum_{i=1}^n \dot{g}_{ij} \cdot (x_i - c_i)(x_j - c_j) = \\ &\equiv \dot{\phi}(t) = \dot{g}_0 + \sum_{i=1}^n \dot{a}_i \cdot t_i + \frac{1}{2} \cdot \sum_{j=1}^n \sum_{i=1}^n \dot{a}_{ij} \cdot t_i \cdot t_j = \\ &= \dot{g}_0 + \sum_{i=1}^n \dot{a}_i \cdot t_i + \sum_{i=1}^n \frac{1}{2} \dot{a}_{ii} \cdot t_i^2 + \sum_{j=i+1}^n \sum_{i=1}^{n-1} \dot{a}_{ij} \cdot t_i \cdot t_j = \\ &= (\dot{a}_0 + \frac{d}{2} \cdot \sum_{i=1}^n \dot{a}_{ii}) + \sum_{i=1}^n \dot{a}_i \cdot t_i + \sum_{i=1}^n \frac{1}{2} \dot{a}_{ii} \cdot (t_i^2 - d) + \sum_{j=i+1}^n \sum_{i=1}^{n-1} \dot{a}_{ij} \cdot t_i \cdot t_j \equiv \\ &\equiv \dot{b}_0^* + \sum_{p=1}^n \dot{b}_p \cdot u_p + \sum_{p=n+1}^{2n} \dot{b}_p \cdot u_p^* + \sum_{p=2n+1}^K \dot{b}_p \cdot u_p, \end{aligned} \quad (576)$$

где $\dot{b}_0^* = \dot{a}_0^*$, $\dot{b}_p$ пробегают последовательно значения $\dot{a}_i, \dot{a}_{ii} / 2, \dot{a}_{ij}$.

При нормировании x_i в t_i по формуле (552) надо помнить, что шаги Δx_i в n -кубе и дробной m -реплике равны рёбрам. С другой стороны, вдоль осей x_i , т. е. в плане n -крест, шаги равны значениям $\varepsilon \cdot \Delta x_i$. Как известно (см., например, [54, стр. 145]), из требования ортогональности матрицы планирования V , сводящегося к ортогональности столбцов с факторами $u_{n+i}^* = (t_i^2 - d)$, для параметра ε при $n \geq 2$ и $m \leq n$ следует формула (577):

$$\varepsilon = \sqrt{\frac{1}{2} \cdot \left(\sqrt{2^m \cdot N} - 2^m \right)} = \sqrt{\frac{1}{2} \cdot \left(\sqrt{2^m \cdot (1 + 2n + 2^m)} - 2^m \right)} \leftrightarrow \frac{2^m}{N} = d^2.$$

Порядок m дробной реплики всегда меньше размерности n на число применяемых генерирующих соотношений — *генераторов* (одного или нескольких). Эти соотношения применяются именно для плана n -ЦКП с целью снижения количества точек в части n -куба с 2^n до 2^m , где $m < n$. Количество точек для дробного n -ЦКП Бокса — Уилсона сокращается до $N = 2^m + 2n + 1$. Это происходит за счёт того, что на каждое новое генерирующее соотношение из плана n -куб, во-первых, изымается столбец, как правило, начиная с t_n , и далее $t_{n-1}, t_{n-2}, \dots$, и, во-вторых, при этом сокращается в 2 раза количество строк ввиду их повторения. Далее каждый фактор взаимодействия переменных $u_p = t_i \cdot t_h$ умножается на генерирующее соотношение вида $t_h = t_k \cdot t_r \cdot \dots \cdot t_s$, образуемое из остающихся в дробной m -реплике факторов $t_1, t_2, \dots$, с естественным сокращением квадратов $t_r^2 = 1$. При этом получают *определяющие контрасты*, т. е. те же самые факторы взаимодействий переменных, но выраженные теперь через произведения всех оставшихся в дробной m -реплике факторов $t_1, t_2, \dots$. Последние не должны совпадать с $u_p = t_i$ и $u_p = t_i \cdot t_j$. Отсюда исходят требования к допустимым генерирующим соотношениям: они не должны в своих определяющих контрастах смешивать линейные факторы и факторы взаимодействий. Поэтому описанный редуccionный приём возможно использовать лишь только, начиная с $n = 5$. Так, например, при $n = 5$ применяют генерирующее соотношение типа $t_5 = t_1 \cdot t_2 \cdot t_3 \cdot t_4$ с определяющими контрастами в ряду $t_1 \cdot t_5 = t_2 \cdot t_3 \cdot t_4, t_2 \cdot t_5 = t_1 \cdot t_3 \cdot t_4, t_3 \cdot t_5 = t_1 \cdot t_2 \cdot t_4, t_4 \cdot t_5 = t_1 \cdot t_2 \cdot t_3$, которые не смешиваются друг с другом и с $t_1, t_2, t_3, t_4, t_1 \cdot t_2, t_1 \cdot t_3, t_1 \cdot t_4, t_2 \cdot t_3, t_2 \cdot t_4$ и $t_3 \cdot t_4$. При $n = 6$ и при $n = 7$ применяют одно и то же генерирующее соотношение $t_6 = t_7 = t_1 \cdot t_2 \cdot t_3 \cdot t_4 \cdot t_5$. При $n = 8$ возможно применение сразу 2-х соотношений $t_7 = t_1 \cdot t_2 \cdot t_3 \cdot t_4$ и $t_8 = t_1 \cdot t_2 \cdot t_5 \cdot t_6$ с $m = n - 2$.

Приведём примеры для начальных планов при $n = 3, 4$ и 5 , легко обобщаемых при $n > 5$.

Полный 3-ЦКП Бокса — Уилсона есть композиция из плана 3-крест с плечом $\varepsilon = 1,215$ по формуле (577), единичного плана 3-куб и нулевой центральной точки $t_0 = 0$.

Полный 4-ЦКП Бокса — Уилсона есть композиция из плана 4-крест с плечом $\varepsilon = 1,414$ по формуле (577), единичного плана 4-куб и нулевой центральной точки $t_0 = 0$.

Полный 5-ЦКП Бокса — Уилсона есть композиция из плана 5-крест с плечом $\varepsilon = 1,596$ по формуле (577), единичного плана 5-куб и нулевой центральной точки $t_0 = 0$.

Дробный 5-ЦКП Бокса — Уилсона есть композиция из плана 5-крест с плечом $\varepsilon = 1,547$ по формуле (577), единичного плана 4-куб и нулевой центральной точки $t_0 = 0$. Здесь используется, дробная 4-реплика и генерирующее соотношение $t_5 = t_1 \cdot t_2 \cdot t_3 \cdot t_4$.

Ниже даны формулы для оценок информационной и ковариационной матриц планов, а также формулы для оценок коэффициентов регрессии и их дисперсий применительно к модели линейной регрессии (596) и к исходной модели (564) на основе общих вычислительных формул типа (508), (517) и (518).

Диагональная информационная $(K+1) \times (K+1)$ -матрица имеет вид:

$$W = W' = V' \cdot V = D_{(K+1) \times (K+1)} =$$

$$= \begin{Bmatrix} w_0 & 0 & 0 & 0 \\ 0 & w_1 \cdot I_{n \times n} & 0 & 0 \\ 0 & 0 & w_2 \cdot I_{n \times n} & 0 \\ 0 & 0 & 0 & w_3 \cdot I_{C_n^2 \times C_n^2} \end{Bmatrix}, \quad (578)$$

$$\text{где } w_0 = \sum_{q=0}^{N-1} t_{0(q)}^2 = 2^m + 2n + 1 = N,$$

$$w_1 = \sum_{q=0}^{N-1} t_{i(q)}^2 = 2^m + 2\varepsilon^2 = d \cdot N = 2^m / d,$$

$$w_2 = \sum_{q=0}^{N-1} (t_i^2 - d)_{(q)}^2 = 2^m \cdot (1-d)^2 + 2 \cdot (\varepsilon^2 - d)^2 + (2n-1) \cdot d^2 = 2\varepsilon^4, \quad (579)$$

$$w_3 = \sum_{q=0}^{N-1} (t_i \cdot t_j)_{p(q)}^2 = 2^m.$$

Формулы (577) и (579) доказываются из, во-первых, независимости коэффициента $b_p = a_{ii} / 2$ от коэффициентов $b_0 = a_0$, $b_p = a_i$, $b_p = a_{ij}$ при любом значении ε и, во-вторых, независимости коэффициента $b_p = a_{ii} / 2$ от коэффициентов $b_p = a_{jj} / 2$ ($p = n+1 \div 2n$), но только при определённых значениях ε . Все 4 формулы для параметров получают суммированием квадратов факторов: $u_0 = +1$, $u_p = t_i$, $u_p^* = (t_i^2 - d)$, $u_p = t_i \cdot t_j$.

В соответствии с (508) при *нецентрализованной* линейной регрессии компоненты вектора $\dot{\mathbf{b}}$, или коэффициенты $\dot{b}_p$ (где индекс p пробегает значения от 0 до K) вычисляются при $n \geq 2$ как

$$\dot{\mathbf{b}} = (\mathbf{V}' \cdot \mathbf{V})^{-1} \cdot \mathbf{V}' \cdot \tilde{\mathbf{y}} \quad \left(K+1 = \frac{(n+1) \cdot (n+2)}{2} \right).$$

Отсюда, с учётом (576), имеем:

$$\dot{b}_0^* = \dot{a}_0^* = \sum_{q=0}^{N-1} \tilde{y}_q / w_0 = \dot{a}_0 + \frac{w_1}{w_0} \cdot \sum_{i=1}^n \frac{1}{2} \cdot \dot{a}_{ii}, \quad \dot{g}_0 = \dot{a}_0 = \dot{b}_0; \quad (580)$$

$$\dot{b}_p = \sum_{q=0}^{N-1} u_{p(q)} \cdot \tilde{y}_q / w_1 = \dot{a}_i, \quad u_p \equiv t_i, \quad \dot{g}_i = \frac{\dot{a}_i}{\Delta x_i}; \quad (581)$$

$$\dot{b}_p = \sum_{q=0}^{N-1} u_{p(q)}^* \cdot \tilde{y}_q / w_2 = \frac{\dot{a}_{ii}}{2}, \quad u_p^* \equiv t_i^2 - d, \quad \dot{g}_{ii} = \frac{\dot{a}_{ii}}{\Delta^2 x_i}; \quad (582)$$

$$\dot{b}_p = \sum_{q=0}^{N-1} u_{p(q)} \cdot \tilde{y}_q / w_3 = \dot{a}_{ij}, \quad u_p \equiv t_i \cdot t_j, \quad \dot{g}_{ij} = \frac{\dot{a}_{ij}}{\Delta x_i \cdot \Delta x_j}. \quad (583)$$

Диагональная ковариационная $(K+1) \times (K+1)$ -матрица имеет вид:

$$\mathbf{C} = \begin{Bmatrix} c_0 & 0 & 0 & 0 \\ 0 & c_1 \cdot \mathbf{I}_{n \times n} & 0 & 0 \\ 0 & 0 & c_2 \cdot \mathbf{I}_{n \times n} & 0 \\ 0 & 0 & 0 & c_3 \cdot \mathbf{I}_{C_n^2 \times C_n^2} \end{Bmatrix},$$

где $c_0 = w_0^{-1}$, $c_1 = w_1^{-1}$, $c_2 = w_2^{-1}$, $c_3 = w_3^{-1}$ вычисляются из (578).

Теоретически из (518)–(520) и (580)–(583) вычисляются дисперсии для случайных ошибок независимых коэффициентов регрессии:

$$\sigma^2 \{\dot{g}_0\} = \sigma^2 \{\tilde{y}\} \cdot c_0 \cdot (1 + n \cdot c_2 / c_3), \quad (584)$$

где также применяются формулы (577) и (579);

$$\sigma^2 \{\dot{g}_i\} = \sigma^2 \{\tilde{y}\} \cdot c_1 / \Delta^2 x_i, \quad (585)$$

$$\sigma^2 \{1/2 \cdot \dot{g}_{ii}\} = \sigma^2 \{\tilde{y}\} \cdot c_2 / \Delta^4 x_i, \quad (586)$$

$$\sigma^2 \{\dot{g}_{ij}\} = \sigma^2 \{\tilde{y}\} \cdot c_3 / \Delta^2 x_i \cdot \Delta^2 x_j. \quad (587)$$

Вычислительные параметры простейших планов Бокса — Уилсона приведены в таблице 4.

Таблица 4. Основные параметры n -ЦКП Бокса — Уилсона при $n \leq 8$

n	m	N	ε	$d \cdot 10$	$c_0 \cdot 10$	$c_1 \cdot 10$	$c_2 \cdot 10$	$c_3 \cdot 10$
2	2	9	1	6,667	1,111	1,667	5	2,5
3	3	15	1,215	7,303	0,667	0,913	2,296	1,250
4	4	25	1,414	8	0,4	0,5	1,25	0,625
5	5	43	1,596	8,627	0,232	0,270	0,770	0,312
5	4	27	1,547	7,700	0,370	0,481	0,873	0,625
6	6	77	1,761	9,117	0,130	0,142	0,520	0,156
6	5	45	1,722	8,433	0,222	0,264	0,564	0,312
7	7	143	1,909	9,461	0,070	0,074	0,376	0,078
7	6	79	1,885	9,001	0,127	0,141	0,389	0,156
8	8	273	2,045	9,684	0,037	0,038	0,286	0,039
8	6	81	2,001	8,889	0,123	0,139	0,312	0,156

* * *

В обоих изложенных плановых методах оптимизации 2-го порядка движение к экстремуму осуществляется по директивному вектору:

$$\dot{\mathbf{p}} = (-\dot{\mathbf{G}})^{-1} \cdot \dot{\mathbf{g}}'. \quad (588)$$

Его компоненты имеют те же самые размерности, что и компоненты $\mathbf{x}$. Поэтому при изменении масштабов x_i они преобразуются ковариантно.

Если матрица Гессе $\dot{\mathbf{G}}$ на области данного плана плохо обусловлена (т. е. $\det \dot{\mathbf{G}} \approx 0$), то для однозначности и устойчивости результатов целесообразно использовать метод регуляризации Тихонова (§ 5.10). Как параметр квадратичной регуляризации, например, можно выбрать $\alpha = \max \{ \dot{g}_i / \Delta x_i \} \rightarrow 0$. Тогда в итоге имеем устойчивую оценку:

$$\dot{\mathbf{p}}_\alpha = -(\dot{\mathbf{G}} \pm \alpha \mathbf{I})^{-1} \cdot \dot{\mathbf{g}}'. \quad (589)$$

Условие значимости любого коэффициента линейной регрессии в данной модели выражается через его плановое среднее квадратичное отклонение как:

$$|\dot{b}_p| \geq t_2(\mathcal{P}, k) \cdot |s(\dot{b}_p)|, \quad (590)$$

где t_2 — двусторонний критерий Стьюдента для принимаемого уровня доверительной вероятности при проверке гипотезы (как правило, 0,95) и числа k степеней свободы дисперсии $s^2(\dot{b}_p)$ — см. § 6.3. Число k есть количество независимых случайных величин $\tilde{y}_q$ в формуле для $\dot{b}_p$.

Незначимый коэффициент регрессии $\dot{b}_p$, если он только относится к *влияющему* фактору u_p , можно сделать значимым, увеличив интервал варьирования переменной Δx_i для факторов типов $u_p = t_i$ и $u_p = t_i^2$ или увеличив интервалы варьирования Δx_i и/или Δx_j для фактора $u_p = t_i \cdot t_j$. Разумеется, влияющие факторы u_p отбираются только в самом начале процесса оптимизации. Далее по мере приближения к окрестности стационарности (экстремума) целевой функции $y(\mathbf{x})$ и тем более на ней факторы типа $u_p = t_i$, согласно лемме Эйлера — Ферма, мало влияют, так как первые частные производные функции стремятся к нулю!

Критерием остановки общего процесса является ограничение всех коэффициентов регрессии типа $u_p = t_i$, согласно неравенству (590). Исходя из коэффициентов в точке стационарности $\mathbf{s}$ строят итоговую модель целевой функции отклика без линейных членов.

§ 7.7. Сравнительный анализ планов эксперимента

В таблице 5 дано количественное сравнение рассмотренных выше в §§ 7.5 и 7.6 планов эксперимента по числу точек N — без учёта числа точек в центре плана, которые нужны всегда для оценки дисперсии воспроизводимости (540).

Таблица 5. Сравнение планов эксперимента по числу точек

n	m	осевой n-крест N_1	плоскостной n-крест N_2	n-СКП N_3	n-куб или его m-реплика N_4	n-ЦКП N_5
2	2	5	4	9	4	9
3	3	7	12	19	8	15
4	4	9	24	33	16	25
5	5	11	40	51	32	43
5	4	—	—	—	16	27
6	6	13	60	73	64	77
6	5	—	—	—	32	45
7	7	15	84	99	128	143
7	6	—	—	—	64	79
8	8	17	112	129	256	273
8	6	—	—	—	64	81

Здесь $N_1 = 2n + 1 = N^+ + 1$, $N_2 = 4 \cdot n \cdot (n - 1)/2 = 2(n^2 - n) = N^\square$,

$N_3 = 2n + 4 \cdot n \cdot (n - 1)/2 + 1 = 2n^2 + 1 = N^+ + N^\square + 1$,

$N_4 = 2^n \vee 2^m = N^\blacksquare$, $N_5 = N^\blacksquare + N^+ + 1$,

где N^+ , $N^\square$ и $N^\blacksquare$ — число точек в *элементарных* планах: неполный осевой n-крест, плоскостной n-крест и n-куб или его дробная реплика.

В таблице 5 жирным шрифтом отмечены числовые характеристики для трёх вышерассмотренных методов планирования эксперимента, использующих модели неполного и полного 2-го порядка.

Из таблицы 5 видно следующее. Во-первых, план n -крест с работой по модели (554) требует значительно меньшего количества точек, нежели оба композиционных плана с работой по модели (564). Во-вторых, для работы с моделью (564) при $n < 5$ оба композиционных плана требуют сопоставимого количества точек; но при $n \geq 5$ имеется значительное преимущество у *дробных реплик* n -ЦКП, хотя, очевидно, *полные* n -ЦКП с дальнейшим ростом n по числу точек всё более проигрывают n -СКП. Здесь также, как в директивных численных методах поиска экстремума (гл. 5 и 6), скорость сходимости процедуры оптимизации определяется необходимым *общим* числом точек на двух повторяющихся стадиях: (1) — нахождение модели целевой функции (всех её коэффициентов с вычислением директивного вектора); (2) — одномерная оптимизация по направлению директивного вектора. С моделью неполного 2-го порядка (по масштабнo-градиентному методу) требуется в сравнении с моделью полного 2-го порядка (по обшeньютонoвскому методу) значительно меньше точек на 1-й стадии, но большее число точек на 2-й стадии. Кроме того, количество повтoров обшeй процедуры до достижения экстремума функции, очевидно, во втором случае значительно меньше. Можно прибегать к испытанному приёму: вначале применяется метод с моделью (554), а на заключительном этапе — метод с моделью (564).

В планово-экспериментальных процедурах на скорость сходимости всего процесса существенное влияние оказывает достигаемая точность при нахождении отдельных значений целевой функции в точках плана и на директории одномерной оптимизации. Эта точность характеризуется выборочной дисперсией $s^2\{\bar{y}\}$, определяемой в нулевой точке плана эксперимента. Например, если формально тут принять, что случайного отклонения нет, то скорость сходимости процесса будет тождественна таковой в численном прототипе данного метода экспериментальной оптимизации. Реально такая картина оптимизации возможна тогда, когда относительная ошибка при измерении функции *значительно* меньше относительной ошибки регрессионной модели (550), причём последняя укладывается в приемлемый уровень 5% (см. § 7.4). Обычно случайная ошибка модели вносит весьма заметный вклад в её общую ошибку. Согласно (551), этот вклад даже можно оценить через статистический подход. С другой стороны, систематическая ошибка модели определяется достигаемой точностью при вычислении её коэффициентов регрессии, что зависит от того, какой план эксперимента для этого используется.

Разумеется, начальный коэффициент модели $\dot{g}_0$ никак не влияет на сам директивный вектор, но его смещение приводит к параллельному сдвигу директивного вектора. Однако влияют все прочие коэффициенты регрессии. Поэтому $\dot{g}_0$ в масштабно-градиентном и общеньютоновском методе в §§ 7.5 и 7.6.1 определяется как *несмещённая* оценка $\dot{g}_0 = \tilde{y}_0$. Но по n-ЦКП Бокса—Уилсона, согласно (580), вычисляется *смещённая* оценка этого начального коэффициента, как правило $\dot{g}_0 \neq \tilde{y}_0$, которая для целевых функций, разлагаемых в степенной ряд порядка более 2-х всегда содержит непредсказуемую алгебраическую сумму членов ряда порядка 3-х и выше (см. далее). Отсюда директория также смещается!

Используя представление целевой функции отклика бесконечным степенным рядом Тейлора (с предположением об её аналитичности), получаем теоретически для планов эксперимента в §§ 7.5, 7.6.1 и 7.6.2 смещения всех коэффициентов регрессии моделей (554) и (564). Как простейшие составляющие, применяем представления коэффициентов тех же самых моделей функции, но исходя из *элементарных* планов эксперимента: a^+ для плана *неполный* осевой n-крест ($N^+ = 2n$), $a^\square$ для плана плоскостной n-крест ($N^\square = 4 \cdot n \cdot (n - 1)/2$), $a^\blacksquare$ для плана n-куб ($N^\blacksquare = 2^n$), — а также $\alpha_0 = y_0 = g_0$ как *точный* начальный коэффициент.

В случае полного плана n-крест ($N_1 = 2n + 1$) и модели (554) имеем:

— для начального коэффициента $\Delta \dot{a}_0 = 0$;

— для коэффициентов $\dot{a}_i$ и $\dot{a}_{ii}$ по формулам (557)

$$\left. \begin{aligned} \Delta \dot{a}_i &= a_i^+ - \alpha_i, \\ \Delta(\dot{a}_{ii} / 2) &= a_{ii}^+ / 2 - \alpha_{ii} / 2, \end{aligned} \right\} \quad (591, 592)$$

где

$$\begin{aligned} a_i^+ &= \alpha_i + \sum_{k=1}^{\infty} \alpha_{(i \dots ii)_{2k+1}} / (2k+1)!, \\ a_{ii}^+ / 2 &= \alpha_{ii} / 2 + \sum_{k=1}^{\infty} \alpha_{(ii \dots ii)_{2k+2}} / (2k+2)!, \end{aligned}$$

т. е. с первыми в рядах членами в виде

$$a_i^+ = \alpha_i + (\alpha_{iii}/3! + \alpha_{iiii}/5! + \alpha_{iiiii}/7! + \dots);$$

$$a_{ii}^+ / 2 = \alpha_{ii}/2 + (\alpha_{iiii}/4! + \alpha_{iiiii}/6! + \alpha_{iiiiiii}/8! + \dots).$$

В случае n -СКП ($N_3 = 2n^2 + 1$) и модели (564) имеем:
 — для начального коэффициента $\Delta \dot{a}_0 = 0$;

— для коэффициентов $\dot{a}_i$, $\dot{a}_{ii}$ и $\dot{a}_{ij}$ по формулам (570)

$$\left. \begin{aligned} \Delta \dot{a}_i &= \frac{2a_i^+ + 4(n-1) \cdot a_i^\square}{2 + 4(n-1)} - \alpha_i, \\ \Delta(\dot{a}_{ii} / 2) &= \left\{ b \cdot a_{ii}^+ + \sum_{j=l(j \neq i)}^n [-c \cdot a_{jj}^+ + (b-c) \cdot 2(a_{ii}^\square + a_{jj}^\square)_{(j,i)}] - \right. \\ &\quad \left. - 4 \cdot \sum_{m=j+1(m \neq i)}^n 2(a_{jj}^\square + a_{mm}^\square)_{(j,m)} \right\} / 2d - \alpha_{ii} / 2, \\ \Delta \dot{a}_{ij} &= a_{ij}^\square - \alpha_{ij}, \end{aligned} \right\} \quad (593—595)$$

где b , c и d имеют значения, согласно формулам (569),

a_i^+ и $a_{ii}^+ / 2$ — см. выше,

$$\begin{aligned} a_i^\square &= \alpha_i + \sum_{p=1}^{\infty} \alpha_{i[(ii)_p} / (2p+1)! + \\ &\quad + (n-1)^{-1} \sum_{h=1(h \neq i)}^n \sum_{k=1(q \geq 1)}^{\infty} \alpha_{i[(ii)_p(hh)_q]_{2k}} / (2p+1)! (2q)!, \\ a_{ii}^\square / 2 + a_{jj}^\square / 2 &= \alpha_{ii} / 2 + \alpha_{jj} / 2 + \sum_{k=1}^{\infty} \alpha_{[(ii)_p(jj)_q]_{2k+2}} / (2p)! (2q)! \end{aligned}$$

— для примера,

$$a_{ij}^\square = \alpha_{ij} + \sum_{k=1}^{\infty} \alpha_{ij[(ii)_p(jj)_q]_{2k}} / (2p+1)! (2q+1)!;$$

т. е. с первыми в рядах членами в виде

$$\begin{aligned} a_i^\square &= \alpha_i + \alpha_{iii} / 3! + \alpha_{iiii} / 5! + \dots + \\ &\quad + 1 / (n-1) \cdot [\alpha_{i11} / 2! + \dots + \alpha_{inn} / 2! + \alpha_{i1111} / 4! + \dots + \alpha_{iii11} / 3! 2! + \dots], \\ a_{ii}^\square / 2 + a_{jj}^\square / 2 &= \alpha_{ii} / 2 + \alpha_{jj} / 2 + \alpha_{iiii} / 4! + \alpha_{jjjj} / 4! + \alpha_{ijjj} / 2! 2! + \dots, \\ a_{ij}^\square &= \alpha_{ij} + \alpha_{ijii} / 3! + \alpha_{ijjj} / 3! + \alpha_{ijiiii} / 5! + \alpha_{ijjjj} / 3! 3! + \alpha_{ijjjjj} / 5! + \dots \end{aligned}$$

В случае n -ЦКП ($N_5 = 2^n + 2n + 1$) и модели (564) для вычисления смещений применяем исходные формулы для *сложных* коэффициентов регрессии (580)–(583), таблицу плана (575) и разложения в степенной ряд *простых* коэффициентов для *элементарных* планов.

Причём при дробном n -ЦКП в формулы для смещений вводятся дополнительные слагаемые в части составляющих от плана n -куб. Для их получения индексы истинных коэффициентов α в части от n -куба символически умножают на алгебраическую сумму из определяющих контрастов (с попутным сокращением квадратов $t_r^2 = 1$) [46, с. 134] — см. о контрастах в § 7.6.2. Для дробного n -ЦКП ($N_5 = 2^m + 2n + 1$) и модели (564) в главных формулах для смещений 4-х коэффициентов делается замена $2^n \rightarrow 2^m$. Далее вычисляем смещения (где $\alpha_0 = y_0 = g_0$).

Для коэффициентов $\dot{a}_i$, $\dot{a}_{ii}$, $\dot{a}_{ij}$ по формулам (581)–(583), с учётом плана, аналогичного (575), имеем:

$$\Delta \dot{a}_i = \frac{2\varepsilon \cdot (\varepsilon \cdot a_i^+) + 2^n \cdot a_i^\square}{w_1} - \alpha_i = \frac{2\varepsilon \cdot (\varepsilon \cdot a_i^+) + 2^n \cdot a_i^\square}{2\varepsilon^2 + 2^n} - \alpha_i, \quad (596)$$

$$\begin{aligned} \Delta(\dot{a}_{ii} / 2) &= \{-d \cdot \alpha_0 + (\varepsilon^2 - d) \cdot (2\varepsilon^2 \cdot a_{ii}^+ / 2 + 2\alpha_0) - \\ &- d \cdot [2\varepsilon^2 \cdot (\sum_{r=1}^n a_{rr}^+ / 2 - a_{ii}^+ / 2) + (2n - 2) \cdot \alpha_0] + \\ &+ (1 - d) \cdot 2^n \cdot a_0^\square\} / w_2 - \alpha_{ii} / 2 = \\ &= \{-d \cdot \alpha_0 + 2\varepsilon^4 \cdot a_{ii}^+ / 2 + (\varepsilon^2 - d) \cdot 2\alpha_0 - d \cdot 2\varepsilon^2 \cdot \sum_{r=1}^n a_{rr}^+ / 2 - \\ &- d \cdot (2n - 2) \cdot \alpha_0 + (1 - d) \cdot 2^n \cdot [\alpha_0 + \sum_{r=1}^n a_{rr}^+ / 2 + \Delta]\} / w_2 - \alpha_{ii} / 2 = \\ &= \{\alpha_0 \cdot [-d + 2(\varepsilon^2 - d) - (2n - 2) \cdot d + 2^n \cdot (1 - d)] + 2\varepsilon^4 \cdot a_{ii}^+ / 2 + \\ &+ [-d \cdot 2\varepsilon^2 + (1 - d) \cdot 2^n] \cdot \sum_{r=1}^n a_{rr}^+ / 2 + (1 - d) \cdot 2^n \cdot \Delta\} / w_2 - \alpha_{ii} / 2 = \\ &= \{2\varepsilon^4 \cdot a_{ii}^+ / 2 + (1 - d) \cdot 2^n \cdot \Delta\} / w_2 - \alpha_{ii} / 2 = a_{ii}^+ / 2 - \alpha_{ii} / 2 + d \cdot \Delta / \varepsilon^2, \quad (597) \end{aligned}$$

$$\Delta \dot{a}_{ij}^\square = a_{ij} - \alpha_{ij}, \quad (598)$$

где

$$a_0^{\blacksquare} = \sum_{q=2n+1}^{N-1} \tilde{y}_q / 2^n = \alpha_0 + \sum_{k=1}^{\infty} \sum_{q=1}^n \alpha_{(r_1 \dots r_1)_{2k_1} \dots (r_q \dots r_q)_{2k_q}} / 2k_1! \dots 2k_q! ,$$

$$\varepsilon \cdot a_i^+ = \varepsilon \cdot \alpha_i + \sum_{k=1}^{\infty} \varepsilon^{2k+1} \cdot \alpha_{(i \dots ii)_{2k+1}} / (2k+1)! ,$$

$$\varepsilon^2 \cdot a_{ii}^+ / 2 = \varepsilon^2 \cdot \alpha_{ii} / 2 + \sum_{k=1}^{\infty} \varepsilon^{2k+2} \cdot \alpha_{(ii \dots ii)_{2k+2}} / (2k+2)! ,$$

$$a_i^{\blacksquare} = \alpha_i + \sum_{k=1}^{\infty} \sum_{q=1}^n \alpha_{i(r_1 \dots r_1)_{2k_1} \dots (i \dots i)_{2k_{(i)}} \dots (r_q \dots r_q)_{2k_q}} / 2k_1! \dots (2k_{(i)}+1)! \dots 2k_q! ,$$

$$a_{ij}^{\blacksquare} = \alpha_{ij} + \sum_{k=1}^{\infty} \sum_{q=1}^n \alpha_{ij(r_1 \dots r_1)_{2k_1} \dots (i \dots i)_{2k_{(i)}} (j \dots j)_{2k_{(j)}} \dots (r_q \dots r_q)_{2k_q}} /$$

$$/ 2k_1! \dots (2k_{(i)}+1)! (2k_{(j)}+1)! \dots 2k_q! ,$$

т. е. с первыми в рядах членами в виде

$$a_0^{\blacksquare} = \alpha_0 + (\alpha_{11}/2! + \dots + \alpha_{nn}/2! + \alpha_{1111}/4! + \alpha_{1122}/2!2! + \alpha_{1133}/2!2! + \dots) ,$$

$$\varepsilon \cdot a_i^+ = \varepsilon \cdot \alpha_i + (\varepsilon^3 \cdot \alpha_{iii}/3! + \varepsilon^5 \cdot \alpha_{iiii}/5! + \varepsilon^7 \cdot \alpha_{iiiii}/7! + \dots) ,$$

$$\varepsilon^2 \cdot a_{ii}^+ / 2 = \varepsilon^2 \cdot \alpha_{ii} / 2 + (\varepsilon^4 \cdot \alpha_{iiii}/4! + \varepsilon^6 \cdot \alpha_{iiiii}/6! + \varepsilon^8 \cdot \alpha_{iiiiiii}/8! + \dots) ,$$

$$a_i^{\blacksquare} = \alpha_i + (\alpha_{i11}/2! + \dots + \alpha_{iii}/3! + \dots + \alpha_{inn}/2! + \alpha_{i1111}/4! + \alpha_{i1122}/2!2! + \dots) ,$$

$$a_{ij}^{\blacksquare} = \alpha_{ij} + (\alpha_{ij11}/2! + \dots + \alpha_{ijii}/3! + \dots + \alpha_{ijjj}/3! + \dots + \alpha_{ijnn}/2! + \alpha_{ij1111}/4! +$$

$$+ \alpha_{ij1122}/2!2! + \dots + \alpha_{ij11ii}/2!3! + \dots + \alpha_{ij11jj}/2!3! + \dots + \alpha_{ijiiii}/5! + \dots) ;$$

$$a_0^{\blacksquare} = \alpha_0 + \sum_{r=1}^n a_{rr}^+ / 2 + \Delta \approx \alpha_0 + \sum_{r=1}^n a_{rr}^+ / 2 \text{—с точностью до членов более}$$

3-го порядка, образующих ещё смещение Δ , что использовано в (597);

$$2\varepsilon^4 - d \cdot 2\varepsilon^2 + (1-d) \cdot 2^n = 2(\varepsilon^2 - d)^2 + (1-d)^2 \cdot 2^n + (2n-1) \cdot d^2 = w_2 ,$$

согласно левой части формулы (579), т. е. при любом значении ε , что прямо смешивает $\dot{a}_{ii}/2$ с $\alpha_{ii}/2$;

$[-d + 2(\varepsilon^2 - d) - (2n-2) \cdot d + 2^n \cdot (1-d)] = 0$ при любом значении ε , что уничтожает в $\dot{a}_{ii}/2$ смещение с α_0 ;

$[-d \cdot 2\varepsilon^2 + (1-d) \cdot 2^n] = 0$, но только при значении ε , согласно (577), или правой части (579), что уничтожает смещение $\dot{a}_{ii}/2$ с $\dot{a}_{jj}/2$.

Для $\dot{a}_0$ по формуле (580), с учётом плана типа (575), имеем:

$$\begin{aligned} \Delta \dot{a}_0 &= \frac{(\alpha_0 + 2n \cdot a_0^+ + 2^n \cdot a_0^\square) - (2\varepsilon^2 + 2^n) \cdot \sum_{r=1}^n a_{rr} / 2}{1 + 2n + 2^n} - \alpha_0 = \\ &= \frac{\alpha_0 \cdot (1 + 2n + 2^n) + 2^n \cdot [1 - (1 + 2n + 2^n)^2 \cdot n / \varepsilon^2] \cdot \Delta}{1 + 2n + 2^n} - \alpha_0 = \\ &= 2^n \cdot [1 / (1 + 2n + 2^n) - (1 + 2n + 2^n) \cdot n / \varepsilon^2] \cdot \Delta, \end{aligned} \quad (599)$$

где

$$\begin{aligned} a_0^+ &= \sum_{q=n+1}^{2n} \tilde{y}_q / 2n = \alpha_0 + \frac{1}{n} \cdot \sum_{r=1}^n \sum_{k=1}^{\infty} \varepsilon^{2k} \cdot \alpha_{(rr \dots rr)_{2k}} / (2k)!, \\ a_0^\square &= \sum_{q=2n+1}^{N-1} \tilde{y}_q / 2^n = \alpha_0 + \sum_{k=1}^{\infty} \sum_{q=1}^n \alpha_{(r_1 \dots r_1)_{2k_1} \dots (r_q \dots r_q)_{2k_q}} / 2k_1! \dots 2k_q!, \end{aligned}$$

т. е. с первыми в рядах членами в виде

$$a_0^+ = \alpha_0 + (\varepsilon^2 \cdot \alpha_{11}/2! + \dots + \varepsilon^2 \cdot \alpha_{nn}/2! + \varepsilon^4 \cdot \alpha_{1111}/4! + \varepsilon^4 \cdot \alpha_{2222}/4! + \dots) / n,$$

$$a_0^\square = \alpha_0 + (\alpha_{11}/2! + \dots + \alpha_{nn}/2!) + \alpha_{1111}/4! + \alpha_{1122}/2!2! + \alpha_{1133}/2!2! + \dots;$$

$$a_0^\square = \alpha_0 + \sum_{r=1}^n a_{rr}^+ / 2 + \Delta \text{ — см. выше,}$$

$$\sum_{r=1}^n a_{rr} / 2 = \sum_{r=1}^n a_{rr}^+ / 2 + n(1-d)2^n \cdot \Delta / 2\varepsilon^4 = \sum_{r=1}^n a_{rr}^+ / 2 + (n \cdot d / \varepsilon^2) \cdot \Delta,$$

$$\text{с учётом (577), } 2^n - (2\varepsilon^2 + 2^n) \cdot d \cdot n / \varepsilon^2 = 2^n [1 - (1 + 2n + 2^n)^2 \cdot n / \varepsilon^2].$$

Обратим внимание здесь на то, что в моделях функции 2-го порядка с *начальным членом*, формируемых по элементарным планам неполный п-крест и п-куб, начальный коэффициент модели смешан, как показано выше в степенных рядах, с коэффициентами 2-го и более порядка, откуда и возникает необходимость преобразований: $u_p \rightarrow u_p^*$ и $a_0 \rightarrow a_0^*$.

Проанализируем последовательно формулы смещений (591)–(599) на предмет имеющих отличий разностных коэффициентов модели от родственными истинными коэффициентами в разложении целевой функции в степенной ряд Тейлора. Иначе говоря, проанализируем возможные систематические смещения для коэффициентов регрессионных моделей типа (554) и (564) и соответственно целевой функции отклика.

В случае полного плана n -крест и модели (554), согласно (591) и (592), приведённые коэффициенты a_i смешиваются, главным образом, только с одноиндексным коэффициентом 3-го порядка малости $\alpha_{iii}/3!$; но приведённые коэффициенты a_{ij} смешиваются, главным образом, только с одноиндексным коэффициентом 4-го порядка малости $\alpha_{iiii}/4!$. (При этом начальный коэффициент модели $a_0 = \alpha_0 = g_0 = y_0$ не смещён.)

В случае n -СКП и модели (564), согласно (593—595), приведённые коэффициенты a_i смешиваются, главным образом, с коэффициентами 3-го порядка малости — дважды с одноиндексным $\alpha_{iii}/3!$ и единожды с каждым из ненулевых коэффициентов типа $\alpha_{ijj}/2!$; но, например, $a_{ij}/2$ смешиваются, главным образом, с коэффициентами 4-го порядка малости — дважды с одноиндексным $\alpha_{iiii}/4!$ и единожды с каждым из ненулевых коэффициентов типа $\alpha_{iiij}/2!2!$; приведённые коэффициенты a_{ij} смешиваются, главным образом, с одноиндексными коэффициентами 4-го порядка малости типа $\alpha_{iiii}/3!$ и $\alpha_{iiij}/3!$. (При этом начальный коэффициент модели $a_0 = \alpha_0 = g_0 = y_0$ не смещён.)

В случае n -ЦКП и модели (564), согласно (596), приведённые коэффициенты a_i смешиваются, главным образом, с коэффициентами 3-го порядка малости — с одноиндексным $\varepsilon^2 \cdot \alpha_{iii}/3!$ и с каждым из ненулевых коэффициентов типа $\alpha_{ijj}/2!$; согласно (598), приведённые коэффициенты a_{ij} смешиваются, главным образом, с ненулевыми коэффициентами 4-го порядка малости типа $\alpha_{ijjj}/2!$, $\alpha_{iiij}/3!$ и $\alpha_{iiii}/3!$; согласно же (597) и (599), приведённые коэффициенты $a_{ij}/2$ и начальный (смещённый) коэффициент a_0 смещаются соответственно на величины $(d/\varepsilon^2) \cdot \Delta$ и $2^n \cdot [(1 + 2n + 2^n) \cdot n / \varepsilon^2 - 1 / (1 + 2n + 2^n)] \cdot \Delta$, т. е. смешаны, главным образом, с каждым ненулевым коэффициентом 4-го порядка малости типа $\alpha_{iiii}/4!$ и $\alpha_{i_1 i_1 i_2 i_2}/2!2!$. При этом особенно предсказуемо сильно (из-за слишком большой величины коэффициента перед смещением Δ — примерно $2^n \cdot [(1 + 2n + 2^n) \cdot n / \varepsilon^2]$ смещается коэффициент a_0 , что так же смещает и оценку точки экстремума!

Таким образом, по указанным формулам возможно анализировать смещения всех коэффициентов моделей для рассмотренных планов.

Математическое ожидание для квадрата отклонения коэффициента регрессионной модели от своего истинного значения в степенном ряде Тейлора выражается через его дисперсию и смещение как

$$\mu \left\{ (\dot{g} - g)^2 \right\} = \Delta^2 \dot{g} + \sigma^2 (\dot{g}). \quad (600)$$

Эта зависимость сходна с таковой для суммы квадратов отклонений целевой функции (408) при линейной регрессии типа $\tilde{y}|x$ (§ 4.7) и для дисперсии адекватности (547) значений функции $\tilde{y}$ и $\dot{y}$ (§ 7.4). Отсюда видно обоюдное количественное влияние на достигаемую адекватность модели и целевой функции отклика и случайного, и систематического отклонений от истинных значений коэффициентов, а, следовательно, и на эффективность процесса экспериментальной оптимизации в целом. Поэтому особенно на *последних* этапах планово-экспериментальной оптимизации при выборе плана эксперимента с моделью 2-го порядка для оценки точки экстремума и для аппроксимации целевой функции такой моделью в её окрестности необходимо принимать во внимание статистическую оценку имеющегося реального соотношения вкладов случайной и систематической ошибок, согласно формулам (551).

Вышеизложенный сравнительный анализ двух вариантов процедур для планово-экспериментальной оптимизации 2-го порядка позволяет сделать практический вывод: *n*-СКП предпочтительнее при заметно большем вкладе систематической ошибки; *n*-ЦКП предпочтительнее при заметно большем вкладе случайной ошибки. В первом случае более минимизируется различие целевой функции и её степенной модели 2-го порядка; во втором случае более минимизируется имеющийся разброс экспериментальных значений целевой функции — в точках на области плана эксперимента. Но в любом из рассмотренных вариантов процедур ортогональность или субортогональность матрицы плана эксперимента весьма упрощает вычисление коэффициентов модели.

Помимо этого любая модель регрессии для функции отклика должна обеспечить удовлетворительный уровень относительной ошибки (550), например, не более 5%. Этому способствует как сужение интервалов варьирования факторов (но замедляет движение к экстремуму), так и повышение точности экспериментов (но наиболее затратным путём — увеличением числа опытов). Разумеется, планирование эксперимента при оптимизации целевой функции отклика требует очень взвешенного подхода ко всем таким возможностям и последовательного применения моделей 1-го и 2-го порядка. Кроме того, — это не чисто алгоритмическая процедура, а комплексный творческий процесс, базирующийся ещё на профессиональных знаниях и интуиции исследователя.

Приложение

Физико-математическая кунсткамера

Ниже приведён ряд вопросов и задач, ответы на которые можно получить, например, изучив содержание ряда отдельных глав данной монографии и более ранней книги автора «Тензорная тригонометрия. Теория и приложение» – М: Мир, 2004.

1. Продолжите однозначно арифметические действия:

$$2 \times 2 = 11 = 2 + 2;$$

$$11 \times 11 = 121, \quad 11 + 11 = 22;$$

$$121 \times 121 = ? \quad 121 + 121 = ?$$

Объясните, какая целевая функция в этой задаче оптимизирована и от какого аргумента?

2. Укажите простейший способ общего решения дифференциального уравнения вида:

$$z \cdot d^2z/dx^2 - 3 \cdot (dz/dx)^2 + 6 \cdot z^5 = 0, \text{ где } z(x) \neq 0.$$

Решите в общем виде задачу Ньютона, связав рекуррентно и явно коэффициенты обратного и прямого аналитических степенных рядов.

3. Дайте *вещественное* представление пары сопряжённых комплексных чисел, а также покажите основные вычислительные операции с ними с их обычными свойствами, не прибегая к мнимой единице. В чём здесь проявляются различия между сопряжёнными комплексными числами и особенности в вычислительных операциях с ними?

Покажите также, что любое *вещественное* алгебраическое уравнение степени n имеет полное и единственное *вещественное* общее решение с точностью до реально допустимых перестановок.

4. Выразите формулы Кардано ($n = 3$) и Феррари ($n = 4$) через малые и большие медианы вещественного алгебраического уравнения степени n в знакопеременной форме и с положительными коэффициентами.

5. В 1-й половине XIX века Урбен Леверье, применив при обращении $n \times n$ -матрицы созданный им алгоритм для вычисления её скалярных характеристических коэффициентов и её степеней, открыл на кончике пера неизвестную ранее планету Нептун. Для характеристических коэффициентов $n \times n$ -матрицы B и её целочисленных степеней порядка t при $1 \leq t \leq n$ докажите следующие утверждения:

а) Если $\text{tr } B = \text{tr } B^2 = \dots = \text{tr } B^i = \dots = \text{tr } B^{t-1} = \text{tr } B^t = +1$,

то $k(B, t) = 0$ и, в частности, при $t = n$: $\det B = 0$.

б) Если $\text{tr } B = \text{tr } B^2 = \dots = \text{tr } B^i = \dots = \text{tr } B^{t-1} = \text{tr } B^t = -1$,

то $k(B, t) = (-1)^t$ и, в частности, при $t = n$: $\det B = (-1)^n$.

в) Если $\text{tr } B = \text{tr } B^2 = \dots = \text{tr } B^i = \dots = \text{tr } B^{t-1} = \text{tr } B^t = t$, то $k(B, t) = +1$.

г) Если $-\text{tr } B = \text{tr } B^2 = \dots = (-1)^i \text{tr } B^i = \dots = (-1)^{t-1} \text{tr } B^{t-1} = (-1)^t \text{tr } B^t = t$,

то $k(B, t) = (-1)^t$.

д) Если $\text{tr } B = \text{tr } B^2 = \dots = \text{tr } B^i = \dots = \text{tr } B^{t-1} = \text{tr } B^t = n$, то $k(B, t) = C_n^t$.

е) Если $-\text{tr } B = \text{tr } B^2 = \dots = (-1)^i \text{tr } B^i = \dots = (-1)^{t-1} \text{tr } B^{t-1} = (-1)^t \text{tr } B^t = n$,

то $k(B, t) = (-1)^n \cdot C_n^t$.

6. Целочисленные матрицы хранят в себе довольно много загадок и закономерностей. Ниже приведён ряд примеров для $t \times t$ -матриц.

Докажите следующие формулы:

$$\det \begin{array}{|c|c|c|c|c|c|c|} \hline 1 & 1 & 0 & \dots & 0 & 0 & 0 \\ \hline 1 & 1 & 2 & \dots & 0 & 0 & 0 \\ \hline 1 & 1 & 1 & \dots & 0 & 0 & 0 \\ \hline \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \hline 1 & 1 & 1 & \dots & 1 & t-2 & 0 \\ \hline 1 & 1 & 1 & \dots & 1 & 1 & t-1 \\ \hline 1 & 1 & 1 & \dots & 1 & 1 & 1 \\ \hline \end{array} = 0. \quad (1)$$

$$\det \begin{pmatrix} 1 & -1 & 0 & \dots & 0 & 0 & 0 \\ 1 & 1 & -2 & \dots & 0 & 0 & 0 \\ 1 & 1 & 1 & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & 1 & 1 & \dots & 1 & -(t-2) & 0 \\ 1 & 1 & 1 & \dots & 1 & 1 & -(t-1) \\ 1 & 1 & 1 & \dots & 1 & 1 & 1 \end{pmatrix} = t! \quad (2)$$

$$\det \begin{pmatrix} t & 1 & 0 & \dots & 0 & 0 & 0 \\ t & t & 2 & \dots & 0 & 0 & 0 \\ t & t & t & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ t & t & t & \dots & t & t-2 & 0 \\ t & t & t & \dots & t & t & t-1 \\ t & t & t & \dots & t & t & t \end{pmatrix} = t! \quad (3)$$

$$\det \begin{pmatrix} -t & 1 & 0 & \dots & 0 & 0 & 0 \\ t & -t & 2 & \dots & 0 & 0 & 0 \\ -t & t & -t & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ (-1)^{t-2} t & (-1)^{t-3} t & (-1)^{t-4} t & \dots & -t & t-2 & 0 \\ (-1)^{t-1} t & (-1)^{t-2} t & (-1)^{t-3} t & \dots & t & -t & t-1 \\ (-1)^t t & (-1)^{t-1} t & (-1)^{t-2} t & \dots & -t & t & -t \end{pmatrix} = (-t)! \quad (4)$$

$$\det \begin{pmatrix} n & 1 & 0 & \dots & 0 & 0 & 0 \\ n & n & 2 & \dots & 0 & 0 & 0 \\ n & n & n & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ n & n & n & \dots & n & t-2 & 0 \\ n & n & n & \dots & n & n & t-1 \\ n & n & n & \dots & n & n & n \end{pmatrix} =$$

$$= t! \cdot C_n^t, \text{ в т. ч. } = 0 \text{ при } n < t. \quad (5)$$

$$\det \begin{array}{|c|c|c|c|c|c|c|} \hline -n & 1 & 0 & \dots & 0 & 0 & 0 \\ \hline n & -n & 2 & \dots & 0 & 0 & 0 \\ \hline -n & n & -n & \dots & 0 & 0 & 0 \\ \hline \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \hline (-1)^{t-2}n & (-1)^{t-3}n & (-1)^{t-4}n & \dots & -n & t-2 & 0 \\ \hline (-1)^{t-1}n & (-1)^{t-2}n & (-1)^{t-3}n & \dots & n & -n & t-1 \\ \hline (-1)^tn & (-1)^{t-1}n & (-1)^{t-2}n & \dots & -n & n & -n \\ \hline \end{array} =$$

$$= (-t)! \cdot C_n^t, \text{ в т. ч. } = 0 \text{ при } n < t. \quad (6)$$

7. Почему все указанные ниже алгебраические уравнения обязательно имеют комплексные сопряжённые корни и при этом с положительными вещественными частями?

$$y(x) = x^5 - 10x^4 + 40x^3 - 80x^2 + 90x - 64 = 0,$$

$$y(x) = x^5 - 10x^4 + 40x^3 - 70x^2 + 80x - 64 = 0,$$

$$y(x) = x^5 - 10x^4 + 40x^3 - 80x^2 + 75x - 60 = 0,$$

$$y(x) = x^5 - 25x^4 + 90x^3 - 640x^2 + 80x - 1 = 0,$$

$$y(x) = x^5 - 25x^4 + 160x^3 + 80x - 1 = 0.$$

8. Чем принципиально различаются два понятия: «рефлектор-тензор» и «срединный рефлектор тензорного угла» – в бинарных геометриях с квадратичной метрикой (несмотря на схожесть их выражения)?

9. Изобразите для обобщённых окружностей и псевдоокружностей на вещественной аффинной плоскости графики функций $y(x)$:

а) $|y|^n + |x|^n = |R|^n$,

$n = 0, 1/4, 1/3, 1/2, 1, 3/2, 2, 3, 4, \infty$;

б) $|y|^n - |x|^n = |R|^n$,

$n = 0, 1/4, 1/3, 1/2, 1, 3/2, 2, 3, 4, \infty$.

— Объясните, почему именно для евклидовых, квазиевклидовых и псевдоевклидовых плоскостей и пространств выбрано значение $n = 2$?

— Имеет ли какой-либо геометрический смысл выбор параметра n для аффинных плоскостей и пространств?

(Вопросы, в частности, имеют прямое отношение к обоснованию МНК и квадратичной регрессии, квадратичных метрик, а в физике – СТО.)

Дайте сравнительный анализ ряда *обобщённых* тригонометрических функций при целых значениях $n \geq 1$:

$$y/R = \sin \varphi \equiv \text{Th } \gamma, x/R = \cos \varphi \equiv \text{Sch } \gamma \rightarrow \text{Sh } \gamma \equiv \text{Tg } \varphi, \text{Ch } \gamma \equiv \text{Sec } \varphi;$$

$$|\cos \varphi|^n + |\sin \varphi|^n = 1 \quad (|\cos \varphi|^n = \cos^2 \varphi, |\sin \varphi|^n = \sin^2 \varphi);$$

$$|\text{Ch } \gamma|^n - |\text{Sh } \gamma|^n = 1 \quad (|\text{Ch } \gamma|^n = \text{ch}^2 \gamma, |\text{Sh } \gamma|^n = \text{sh}^2 \gamma)$$

(при $n = 2$ тут имеет место *адекватная* сферическо-гиперболическая аналогия типа: $y/R = \sin \varphi \equiv \text{th } \gamma, x/R = \cos \varphi \equiv \text{sch } \gamma \leftrightarrow$

$$\leftrightarrow \text{sh } \gamma \equiv \text{tg } \varphi, \text{ch } \gamma \equiv \text{sec } \varphi).$$

— Почему углы в квадратичных геометриях и тригонометриях, т. е. при $n = 2$ (евклидовой, квази- и псевдоевклидовой), как геометрические и алгебраические объекты имеют природу двухвалентных тензоров?

— В чём состоит формальное различие между абстрактной и адекватной аналогией в тензорных и скалярных квадратичных тригонометриях?

10. Исходя из вышеуказанной адекватной аналогии возникает некий гиперболический аналог ω сферического угла $\pi/4$:

$$\text{sh } \omega = \text{tg } \pi/4 \rightarrow \omega = \text{Arsh } 1 \approx 0,881 \text{ рад}, \pi/4 = \text{arctg } 1 \approx 0,785 \text{ рад}.$$

Оба числа (константы) представляются сходными степенными рядами:

$$\pi/4 = \text{arctg } 1 = 1 - 1/3 + 1/5 - 1/7 + \dots + (-1)^n \cdot 1/(2n+1) + \dots$$

(указанное разложение для числа $\pi/4$ известно как ряд Лейбница),

$$\omega = \text{Arsh } 1 = 1 - (1/2)/3 + (1 \cdot 3/2 \cdot 4)/5 - (1 \cdot 3 \cdot 5/2 \cdot 4 \cdot 6)/7 + \dots$$

$$\dots + (-1)^n \cdot [1 \cdot 3 \cdot 5 \dots (2n-1)/2 \cdot 4 \cdot 6 \dots 2n]/(2n+1) + \dots$$

Для них же имеются степенные ряды с более высокой скоростью сходимости. Например, для числа ω имеется ряд с дополнительным трансцендентным слагаемым $\ln 2$:

$$\begin{aligned} \omega = \text{Arch } 2^{1/2} = (3/2) \cdot \ln 2 - [1/(2 \cdot 2 \cdot 1 \cdot 2^1) + 1 \cdot 3/(2 \cdot 4 \cdot 2 \cdot 2^2) + \\ + 1 \cdot 3 \cdot 5/(2 \cdot 4 \cdot 6 \cdot 2 \cdot 3 \cdot 2^3) + \dots + \\ + 1 \cdot 3 \cdot 5 \dots (2n-1)/(2 \cdot 4 \cdot 6 \dots 2n \cdot 2n \cdot 2^n) + \dots]. \end{aligned}$$

— Покажите: почему число ω тоже трансцендентное, как и число $\pi/4$?

— Дайте геометрическую трактовку угла ω . Почему $\omega > \pi/4$?

11. Докажите соотношения для скалярных коэффициентов порядка t :

$$k(A_1 \cdot A_2', t) = k(A_1' \cdot A_2, t) = k(A_2 \cdot A_1', t) = k(A_2' \cdot A_1, t),$$

где A_1 и A_2 — $n \times m$ -матрицы.

12. Дайте матричную интерпретацию скалярной формулы:

$$\frac{\det B_{11}}{\det(C_{21} \cdot B_{11})} = \frac{\det(B_{11} \cdot C_{12})}{\det(C_{21} \cdot B_{11} \cdot C_{12})} \rightarrow \frac{\det B_{11}}{\det B_{21}} = \frac{\det B_{12}}{\det B_{22}} \Leftrightarrow$$

$$\Leftrightarrow \det B_{11} \cdot \det B_{22} = \det B_{12} \cdot \det B_{21},$$

где B и C — $r \times r$ -матрицы ранга r .

13. Укажите схему приведения квадратичной формы $y(x) = x' \cdot B \cdot x$, где $B \neq B'$, и эрмитовой формы $y(x) = x^* \cdot B \cdot x$, где $B \neq B^*$, к каноническому виду (к алгебраической сумме квадратов) ортогональным и унитарным преобразованием исходного базиса. Для каких типов $n \times n$ -матриц B эти формы всегда нулевые в любом допустимом базисе?

14. Для сингулярных матриц, задающих какой-либо *планар* или *линеор*, выпишите в едином обозначении все возможные характеристические проекторы — *ортогональные* и *косогональные*:

8 — для вещественной квадратной матрицы,

12 — для комплексной квадратной матрицы,

4 — для вещественной прямоугольной матрицы и 8 — для их пары,

6 — для комплексной прямоугольной матрицы и 12 — для их пары.

Составьте *таблицу умножения* характеристических проекторов.

Почему парные ортогональные и косогональные характеристические проекторы обмениваются собственной природой при трансформациях из квазиевклидовой в псевдоевклидову геометрию (тригонометрию) и обратно? Имеются ли геометрические различия между ортогональными и симметричными, косогональными и несимметричными проекторами в пространствах с квадратичной метрикой? Когда характеристические проекторы являются, по сути, аффинными?

15. Что объединяет известные плоские кривые: циклоиду и трактрису? Опишите тела их параметрического вращения.

Каким общим свойством обладают окружности и сферы, равнобочные гиперболы и гиперболоиды, цепные линии и катеноиды, а также циклоиды, трактрисы и псевдосферы? Каков их единый определяющий параметр?

16. Какой угол задаёт отрезки в гиперболической геометрии, а также движения в псевдоевклидовой геометрии и в теории относительности? Взаимосвязан ли он как-то с углом параллельности Лобачевского?

Поясните *тригонометрическое* отличие в математическом описании основных релятивистских эффектов: замедления времени Эйнштейна и сокращения протяжённости Лоренца.

17. В чём состоит суть *математического принципа относительности* в геометриях и как он корреспондируется с *физическим принципом относительности* в природе?

18. Когда целесообразно применение понятия *физико-математический изоморфизм*? В чём его суть?

Могут ли иметь место понятия: химико-математический изоморфизм, биолого-математический изоморфизм и т. п.? Какова первичная основа подобных понятий?

19. Какая имеется связь между общей кривизной мировой линии и 2-ым законом механики Ньютона? Какие частные виды кривизны мировой линии имеют место в четырёхмерном пространстве-времени Минковского? Как они отвечают характеру физического движения материальной точки?

Дайте формулировку 1- и 2-го закона механики Ньютона в терминах геометрии мировых линий в пространстве-времени Минковского.

20. Укажите аналоги релятивистских формул Эйнштейна для полной энергии $E = mc^2$ и для внутренней энергии покоя $E_0 = m_0c^2$ также для величин: полного импульса P и внутреннего импульса покоя P_0 материального объекта соответственно массой m и m_0 .

Если в четырёхмерном псевдоевклидовом пространстве-времени Минковского выбрать *тригонометрический* псевдodeкартов базис, в котором $c = 1$, то останется ли в нём какое-то *физическое* различие между характеристиками m , P и E ?

21. Знаконеопределённый метрический тензор задаёт метрику n -мерного пространства, причём либо квази-, либо псевдориманова с касательным в каждой его точке либо квази-, либо псевдоевклидовым пространством. Что требуется для задания *конкретной метрики* из этих двух вариантов и соответственно n -мерного пространства в целом с таким тензором?

Список литературы

1. Алексеев В. М., Галеев Э. М., Тихомиров В. М. Сборник задач по оптимизации. Теория. Примеры. Задачи. — М.: Физматлит, 2007.
2. Арутюнов А. В. Условия экстремума. Анормальные и вырожденные задачи. — М.: Факториал, 1997.
3. Аттетков А. В., Галкин С. В., Зарубин В. С. Методы оптимизации. — М.: МГТУ, 2003.
4. Бахвалов Н. С., Жидков Н. П., Кобельков Г. М. Численные методы. — М.: Наука, 1987.
5. Болтянский В. Г., Яглом И. М. Геометрические задачи на максимум и минимум. — М.: Наука, 1966.
6. Буслов В. А., Яковлев С. Л. Численные методы. I. Исследование функций. II. Решение уравнений. — С.-Петербург: С.-ПбГУ, 2001.
7. Васильев Ф. П. Методы решения экстремальных задач. — М.: Наука, 1981.
8. Васильев Ф. П. Численные методы решения экстремальных задач. — М.: Наука, 1988.
9. Владимиров В. С. Методы теории функций многих комплексных переменных. — М.: Наука, 1964, с. 48 - 49.
10. Гилл Ф., Мюррей У. Численные методы условной оптимизации. // Том 1 и 2.: Пер. с англ. — М.: Мир, 1977, с. 196 - 206.
11. Гилл Ф., Мюррей У., Райт М. Практическая оптимизация: Пер. с англ. — М.: Мир, 1985.
12. Глебов Н. И., Кочетов Ю. А., Плясунов А. В. Методы оптимизации. — Новосибирск: НГУ, 2000.
13. Демьянов В. Ф., Рубинов А. М. Приближённые методы решения экстремальных задач. — Л.: Изд-во ЛГУ, 1968.
14. Деннис Д., Шнабель Р. Численные методы безусловной оптимизации и решения нелинейных уравнений: Пер. с англ. — М.: Мир, 1988.
15. Джонсон Н., Лион Ф. Статистика и планирование эксперимента в технике и науке // Том 1. Методы обработки данных / Том 2. Методы планирования эксперимента: Пер. с англ. — М.: Мир, 1980 — 1981.

16. Дэниел К. Применение статистики в промышленном эксперименте: Пер. с англ. — М.: Мир, 1979.
17. Ильин В. А., Позняк Э. Г. Основы математического анализа. // Части 1 и 2. — М.: Наука, 2001.
18. Измаилов А. Ф. Чувствительность в оптимизации. — М.: Физматлит, 2006.
19. Измаилов А. Ф., Солодов М. В. Численные методы оптимизации. — М.: Физматлит, 2005.
20. Канторович Л. В., Крылов В. И. Приближённые методы высшего анализа. — М.: Физматгиз, 1962.
21. Карманов В. Г. Математическое программирование. — М.: Физматлит, 2000.
22. Кобзарь А. И. Прикладная математическая статистика. — М.: Физматлит, 2006.
23. Кудрявцев Л. Д. Курс математического анализа // Том 1, 2 и 3. — М.: Высшая школа, 1989.
24. Курант Р. Курс дифференциального и интегрального исчисления // Том 1 и 2. — М.: Наука, 1967 — 1970.
25. Ланкастер П. Теория матриц: Пер. с англ. — М.: Наука, 1982.
26. Маркушевич А. И. Краткий курс теории аналитических функций. — М.: Наука, 1978.
27. Нинул А. С. Тензорная тригонометрия. *Теория и приложения*. — М.: Мир, 2004.
28. Поля Г. Математика и правдоподобные рассуждения: Пер. с англ. — М.: Наука, 1978, т. 1, с. 141 - 205. В оригинале: Polya G. *Mathematics and Plausible Reasoning*. — New York: Princeton, 1934.
29. Пугачёв В. С. Теория вероятностей и математическая статистика. — М.: Наука, 1979.
30. Пшеничный Б. Н., Данилин Ю. М. Численные методы в экстремальных задачах. — М.: Наука, 1975.
31. Румшиский Л. З. Элементы теории вероятностей. — М.: Наука, 1970.
32. Стренг Г. Линейная алгебра и её применения. — М.: Мир, 1980.
33. Сухарев А. Г., Тимохов А. В., Фёдоров В. В. Курс методов оптимизации. — М.: Физматлит, 2005.
34. Тихомиров В. М. Рассказы о максимумах и минимумах. — М.: Наука, 1986.
35. Тихомиров В. М. Дифференциальное исчисление. *Теория и приложения*. — М.: МЦНМО, 2002.
36. Тихонов А. Н. О некорректных задачах линейной алгебры и устойчивом методе их решения // ДАН СССР — 1965, т. 163, № 3. с. 591 - 594.

37. Тихонов А. Н., Арсенин В. Я. Методы решения некорректных задач. — М.: Наука, 1979.
38. Уайлд Д. Методы поиска экстремума: Пер. с англ. — М.: Наука, 1967.
39. Фаддеев Д. К. Лекции по алгебре. — СПб.: Лань, 2002.
40. Фаддеев Д. К., Фаддеева В. Н. Вычислительные методы линейной алгебры. — М.: Физматгиз, 1963.
41. Фёдоров В. В. Теория оптимального эксперимента. — М.: Наука, 1971.
42. Фиакко А., Маккормик Г. Нелинейное программирование. // Методы последовательной безусловной минимизации: Пер. с англ. — М.: Мир, 1972.
43. Финни Д. Введение в теорию планирования эксперимента: Пер. с англ. — М.: Мир, 1970.
44. Хальд А. Математическая статистика с техническими приложениями: Пер. с англ. — М.: ИЛ, 1956.
45. Харди Г., Литтлвуд Д., Полиа Г. Неравенства: Пер. с англ. — М.: ИЛ, 1948. В оригинале: Hardy G., Littlewood J., Polya G. Inequalities. — London: Cambridge University, 1934.
46. Хартман К., Лецкий Э., Шеффер В. Планирование эксперимента в исследовании технологических процессов: Пер. с нем. — М.: Мир, 1977.
47. Хикс Ч. Основные принципы планирования эксперимента: Пер. с англ. — М.: Мир, 1967.
48. Шабат Б. В. Введение в комплексный анализ. — М.: Наука, 1969.
49. Шилов Г. Е. Математический анализ. Функции одного переменного. — М.: Наука, 1969 — 1970.
50. Шилов Г. Е. Математический анализ. Функции нескольких вещественных переменных. — М.: Наука, 1972.
51. Шор Н. З. Методы минимизации недифференцируемых функций и их приложения. — Киев: Наукова думка, 1979.
52. Bertsekas D. Nonlinear Programming. — Belmont: Athena, 1999.
53. Bonnans J., Gilbert J., Lemaréchal C., Sagastizábal C. Numerical Optimization. Theoretical and Practical Aspects. — Berlin: Springer-Verlag, 2003.
54. Box G., Wilson K. On the Experimental Attainment of Optimum Conditions // J. Roy. Stat. Soc., ser. B. — London, 1951, v. 13, № 1, p. 1 - 45.
55. Box G. Evolutionary Operation: A Method for Increasing Industrial Productivity // Appl. Stat. — London, 1957, v. 6, p. 81 - 101.
56. Cauchy A. Méthode générale pour la resolution des systemes d'equations simultanées // Compt. Rend. Acad. Sci. — Paris: 1847, v. 25, p. 536 - 538.

57. Eulero L. Institutiones calculi differentialis, AIS Petropolitanae, 1755.
Пер. с лат.: Леонард Эйлер. Дифференциальное исчисление. — М., Л.: ГИ Технико-теоретической литературы, 1949.
58. Fermat P. Methodus ad disquirendam maximam et minimam.
(From the letter to mathematician Gilles Roberval). — Tolosae: 1638. Опубликовано: Tolosae, Varia opera mathematica, 1679.
(Метод отыскания максимума и минимума.)
59. Fletcher R. Practical Methods of Optimization. v.1. Unconstrained Optimization. v.2. Constrained Optimization. — Chichester, New York, Brisbae, Toronto: John Wiley, 1980 — 1981.
60. Lagrange J. Theorie des fonctions analytiques // — Paris: 1797.
61. Leibniz G. Nova methodus pro maximis et minimis itemque tangentibus, quae nec fractas nec irracionales quantitates moratur, et singulare pro illis calculi genus // — Leipzig: Acta Eruditorum, 1684.
(Новый метод максимумов и минимумов, а также касательных, для которого не служат препятствием ни дробные, ни иррациональные величины, и особый для этого род исчисления.)
62. Mason R., Gunst R., Hess J. Statistical Design and Analysis of Experiments. — New York: Wiley-Interscience, 2003.
63. The Mathematical Papers of Isaac Newton — 8 volumes. — Cambridge: Cambridge University Press, 1967 — 1981.
64. Stigler S. Gergonne's 1815 Paper on the Design and Analysis of Polynomial Regression Experiments. — Madison: University Wisconsin, 1973, TR. 344, p. 1 - 22 (Historia Math., 1974, v. 1, p. 431 - 447).

Именной указатель

- Абель (Abel N.) 189
Алексеев В. М. 322
Аристотель (Αριστοτέλης) 5
Арсенин В. Я. 323
Арутюнов А. В. 322
Архимед (Αρχιμήδης) 6
Аттетков А. В. 322
- Банах (Banach S.) 225
Бахвалов Н. С. 322
Бертсекас (Bertsekas D.) 324
Бетховен 5
Бине (Binet J.) 220
Блок А. 5
Бокс (Box G.) 7, 269, 287, 324
Болтянский В. Г. 322
Бонанс (Bonnans J.) 324
Буслов В. А. 322
- Варинг (Waring E.) 152 - 159
Васильев Ф. П. 322
Вейерштрасс (Weierstraß K.) 6, 18, 25, 56
Вергилий 5
Виет (Viète F.) 145, 152 - 155
Владимиров В. С. 322
- Галеев Э. М. 322
Галкин С. В. 322
Галуа (Galois E.) 189
Гаусс (Gauss C.) 204, 205, 211, 255, 261, 262, 270, 271
Гесс (Hess J.) 325
Гессе (Hesse L.) 8, 59, 78, 89, 106
Гёльдер (Hölder L.) 151
Гилл (Gill Ph.) 322
Гильберт (Gilbert J.) 324
Глебов Н. И. 322
Грам (Gram J.) 92
- Греффе (Greffe K.) 153, 162
Гунст (Gunst R.) 325
- Даламбер (D'Alembert J.) 115
Данилин Ю. М. 323
Декарт (Des Cartes R.) 6, 62, 166
Демьянов В. Ф. 322
Деннис (J. Dennis) 322
Джонсон (Johnson N.) 322
Дэниел (Daniel C.) 323
- Евклид (Euclid) 161
- Жергонн (J. Gergonne) 269, 325
Жидков Н. П. 322
- Зарубин В. С. 322
Зейдель (Seidel L.) 229
- Измаилов А. Ф. 323
Ильин В. А. 323
- Канторович Л. В. 7, 236, 323
Кардано (Cardano G.) 314
Карманов В. Г. 323
Кобельков Г. М. 322
Кочетов Ю. А. 322
Коши (Cauchy A.) 7, 38, 230, 324
Крылов В. И. 323
Кудрявцев Л. Д. 323
Курант (Courant R.) 7, 103, 323
- Лагранж (Lagrange J.) 7, 26, 87, 90 - 98, 102, 108, 109, 176, 325
Ланкастер (Lankaster P.) 323
Лаплас (Laplace P.) 116, 270
Леверье (Le Verrier U.) 159, 315
Лейбниц (Leibniz G.) 6, 9, 22, 325
Лемарешаль (Lemaréchal C.) 324
Леонардо да Винчи 219

- Лецкий Э. К. 324
Лион (Leone F.) 322
Липшиц (Lipschitz R.) 213, 228
Литлвуд (Littlewood J.) 324
Лобачевский Н. И. 153, 162
Лопиталь (L'Hôpital G.) 44
Ляпунов А. М. 231, 271
- Макормик (Mc Cormick G.) 324
Маклорен (Mc Laurin C.) 151, 159
Маркушевич А. И. 323
Милликен (Millikan R.) 273
Мур (Moore E.) 74, 86, 100, 107, 108, 192, 194, 264
Мэйсон (Mason R.) 325
Мюррей (Murrey W.) 322
- Нейль (Neil W.) 35 - 37
Ньютон (Newton I.) 3, 6, 7, 50, 154, 158, 165, 223 - 226, 235 - 237, 325
- Пеано (Peano G.) 24, 25, 33
Пенроуз (Penrose R.) 74, 86, 100, 107, 108, 192, 194, 264
Пирсон (Pearson K.) 259, 274, 281
Плясунов А. В. 322
Позняк Э. Г. 323
Поля (Pólya G.) 323, 324
Пугачёв В. С. 323
Пшеничный Б. Н. 323
Пуанкаре (Poincaré H.) 227
- Райт (Wright S.) 322
Рахманинов С. В. 5
Релей (Rayleigh J.) 71, 196 - 203
Роберваль (Roberval G.) 6
Роден (Rodin A.) 9
Ролль (Rolle M.) 159
Рубинов А. М. 322
Румшиский Л. З. 323
- Сагастизабал (Sagastizábal C.) 324
Сильвестр (Silvester J.) 7, 61, 166
Снедекор (Snedecor G.) 269, 279
Снеллиус (Snellius W.) 31, 32, 72
Солодов М. В. 323
Стиглер (Stigler S.) 325
- Стренг (Streng G.) 323
Стьюдент (Student W.) 274 - 276
Сухарев А. Г. 323
- Тейлор (Taylor B.) 24, 25, 42, 313
Тимохов А. В. 323
Тихомиров В. М. 322, 323
Тихонов А. Н. 7, 104, 105, 109, 238, 251, 253, 290, 304, 323
- Уайлд (Wilde D.) 323
Уилсон (Wilson K.) 7, 269, 287, 324
- Фаддеев Д. К. 323, 324
Фаддеева В. Н. 324
Ферма (Fermat P.) 3, 6, 18, 19, 31, 32, 43, 58, 60, 324
Феррари (Ferrari L.) 314
Фёдоров В. В. 323
Фиакко (Fiacco A.) 324
Фибоначчи (Pisano L.) 220
Финни (Finney D.) 324
Фишер (Fisher R.) 269, 279, 297
Флетчер (Fletcher R.) 325
Фреше (Fréchet M.) 26
- Хальд (Hald A.) 324
Харди (Hardy G.) 324
Хартман (Hartmann K.) 324
Хикс (Hicks C.) 324
- Шабат Б. В. 324
Шефер (Schäfer W.) 324
Шилов Г. Е. 324
Шмидт (Schmidt E.) 92
Шнабель (Schnabel R.) 322
Шор Н. З. 324
Штейнер (Steiner J.) 6
Штурм (Sturm J.-Ch.) 160, 174
- Эйлер (Eulero L.) 5 - 7, 39, 57, 58, 60, 62, 115, 116, 231, 325
Эйнштейн (Einstein A.) 321
- Якоби (Jacobi K.) 68, 82, 105
Яковлев С. Л. 322

Предметный указатель

- аналитическая функция
 - от вещественного аргумента 25
 - от комплексного аргумента
 - адекватно 115
 - формально 119
- аппроксимационный анализ 246
- базис координатный
 - аффинный 56, 62
 - ковариантный 201
 - контравариантный 201
 - декартов 104, 191, 196
- вложенные многообразия 67
 - гиперповерхность 93
 - траектория (кривая) 79
 - q-плоскость 68, 78, 83, 90
 - q-поверхность 69, 82
- гладкость 20
 - кривой 41
 - поверхности 69, 83
 - функции-отображения
 - векторной 69, 83
 - скалярной 41
- градиент скалярной функции от векторной переменной 59
 - условный 76, 84
 - композиционный 98
- директивный вектор 228
 - 1-го порядка 230
 - 2-го порядка 236, 253, 304
 - неполного 235, 250, 290
- дискриминант уравнения 152
- дисперсионный анализ 297
- дисперсия 256, 270, 272
 - выборочная 257, 272
 - адекватности 280
 - воспроизводимости 279
 - плановая 209, 268, 281
- дифференциал зависимой и ограниченной переменной
 - порядка 1 68, 83
 - порядка 2 70, 88
 - свободный 68, 82
 - условный 68, 70, 82, 88
 - формальный 119, 122
- дифференциал функции 25
 - обратной 50
 - порядка p 50
 - от зависимой и ограниченной переменной 42, 47, 69, 84, 89
 - формальный
 - порядка 1 124, 141
 - порядка 2 124, 141
 - эффект расширения 126, 134
- дифференциальное уравнение Коши 231
- дифференциальные формы 58
- дифференцируемость 25, 26
 - векторной функции
 - непрерывная 57
 - от $q \times 1$ зависимой векторной переменной ранга q 67
 - от $m \times 1$ ограниченной векторной переменной ранга m 82
 - скалярной функции
 - непрерывная 18
 - однократная (по Фреше) 26
 - k -кратная (определение) 25

- от комплексного аргумента
 - адекватная 115
 - уравнения Д'Аламбера — Эйлера
 - векторные 115
 - матричные 116
 - формальная 117, 119, 121
- закон инерции квадратичных форм Сильвестра 61, 201
- закон нормального распределения Гаусса 204, 270
- значимость величины
 - векторной 71
 - матричной 71
 - полная 69
 - скалярной 44
- золотое сечение 219
- связь с числами Фибоначчи 220
- иерархические инварианты 167
- квадратной матрицы 151
- интеграл формальный 121, 122
- независимость от пути
- интегрирования 121, 122
- от пары векторных комплексных функций 122
- от пары скалярных комплексных функций 121
- интервал
 - закрытый 18
 - экстримный 215
 - поиск
 - удвоением шага 214
 - уполовиниванием шага 215
 - с золотым сечением 214, 215
- итерации 212
 - главные и побочные 228
- ковариация 258, 268
 - выборочная 208
 - плановая 209, 268
- компакт 18, 56
- комплексификация 114
- адекватная 114, 115
 - псевдоизированная 114, 116
- эрмитова 114, 117
- координатное пространство
- аффинное 56, 62
- евклидово 104, 191, 196
- факторное 260 - 263
- корреляционный анализ 208
- корреляция 207
- коэффициент линейной корреляции 208, 268
- коэффициенты разностной модели 247, 249, 252
- коэффициенты регрессионной модели 259, 260, 262, 264
 - дисперсия 268
 - смещение 307 - 311
 - условие значимости 304
- коэффициенты для $n \times n$ -матриц
- характеристические
 - матричные 1- и 2-го рода 195
 - скалярные 145, 158, 195
- кривая ошибок 270
- критерий адекватности 246, 283
- критерий Сильвестра 61
- χ^2 -критерий Пирсона 274, 281
- T-критерий Стьюдента 275
 - односторонний и двусторонний 275
 - табулярные значения 276
- F-критерий Фишера — Снедекора 280 - 282
- лемма Ферма 18
- лемма Эйлера — Ферма 58
- лемма 1 (о линеаризации) 75
- лемма 2 (о линеаризации) 83
- лемма 3 (об иерархии на экстримном интервале) 216
- линейная регрессия 259
- линейное уравнение 191
 - вырожденное 191, 194
 - невязка 191

- алгебро-геометрическое
неравенство 192
- решение и квазирешение
квадратичное 191
- нормальное 192, 194
- плохо обусловленное 238
- линеоры 197, 198, 202, 203
- собственные 197, 202
- биортогональные 203
- биортонормированные 203
- ортогональные 197
- ортонормированные 198
- математическая экономика 254
- математическое ожидание 256
- матрица
 - Гессе скалярной функции 59
 - условная 78, 89
 - композиционная 106
 - полуусловная 91
 - функции Лагранжа 91
- знакоопределённая 61
- отрицательно 61
- положительно 61
- полуопределённая 61
- информационная 267, 301
- квазиобратная Мура —
- Пенроуза 74, 86, 100, 107, 108, 192, 194, 264
- явная формула 195
- ковариационная 267, 302
- кососимметричная 160
- нормальная 160, 165
- обратная 74, 86
- левая 74, 264
- правая 86
- планирования 264
- ортогональная 267, 288, 299
- субортогональная 293, 295
- простая с вещественным спектром 201
- диагональная форма 201
- симметричная 160, 165, 196
- диагональная форма 196
- сингулярная 92
- нуль-клеточная 92
- нуль-нормальная 99, 102
- нуль-простая 102
- спектральное представление 198, 203
- трёхмерная частных вторых производных вектор-функции зависимой переменной 70
- ограниченной переменной 88
- Якоби вектор-функции зависимой переменной 68
- ограниченной переменной 83
- метод ломаных Эйлера 231
- метод наименьших квадратов Гаусса 204
- многомерный 210
- одномерный 205
- метод регуляризации Тихонова 104
- метод функций Ляпунова 231
- методы многомерной оптимизации аналитической 56
- условной с вектор-переменной зависимой 67
- на траектории 79
- прямой 79
- проективная версия 76, 78
- ограниченной 82
- множителей Лагранжа 87
- на гиперповерхности 93
- гиперплоскости 93
- проективная версия 84, 90
- с параметром предельные большим 98
- штрафной функции Куранта 103
- малым 98, 99, 103
- планово-вычислительной 245
- по минимальным планам 245
- квадрантный n -симплекс 247
- осевой n -крест 249
- n -АКП 252

- планово-экспериментальной 255
по плану осевой n -крест 286
по n -СКП 291
по n -ЦКП Бокса — Уилсона 297
- численные 227
главные и побочные
итерации 228
- 0-го порядка покоординатный
Зейделя 228, 229
- 1-го порядка градиентный
Коши 230
директивные модификации
крутого восхождения 231
скорейшего спуска 231
устойчивость по Ляпунову 231
- масштабно-градиентный 233
- 2-го порядка Ньютона 235
директивная модификация 236
квадратичная регуляризация по
Тихонову 238
- условной 239
гиперсимплексный 244
нормальных проекций 1-го
и 2-го порядка 242
с большим и малым
параметром 243
с линейно зависимой или
ограниченной переменной 242
- методы одномерной оптимизации
аналитической 17
- численной 212
дихотомии 217
золотого сечения 219
парных касательных 220
- 2-го порядка Ньютона 223
парабол 226
хорд 226
- модели целевой функции
адекватность 246, 280
разностные 245
1-го порядка 247
2-го порядка 252
неполного 249
- регрессионные линейные 259
- неравенство
средних величин 151
генеральное 146
инверсионная форма 164
корреляционное 207
псевдокорреляционное 209
- отношение Релея 196
бинарное 201
экстремумы 199, 203
- отображение 41, 44
векторное 69
гладкое 69, 74
квазиобратное 74
полнозначимое 69, 74
постоянного ранга q 69, 74
регулярное 69, 74
- изоморфное 74
- итерационное 225
сужающее 225
степень сжатия 225, 236
- обратное 45, 83
скалярное 41
- парабола
аппроксимирующая 26
маргинальная 35
Нейля прямая и обратная 35, 36
- план вычисления или
эксперимента 246, 254
минимальный 245
ортогональный 268
 n -СКП (симметричный
композиционный план) 292
 n -ЦКП Бокса — Уилсона 298
дробный 297
генераторы 300
контрасты 300
параметры 303
- центрированный 206, 211, 262
- элементарный 305
 n -куб и дробные реплики 297
неполный осевой n -крест 307
плоскостной n -крест 307

- правила Лейбница 22
правила Лопиталья 44
правила Ролля — Маклорена 159
правила Эйлера 57
принцип максимума модуля для аналитической функции 127, 135
проекторы 192 - 195
 матрицы Якоби от переменной зависимой 75
 ограниченной 84, 100, 101
 симметричные 75, 84, 100, 194
 ортогональные 194
- распределение статистическое**
Гаусса (нормальное) 270
однородное 255, 279
 центральное 255, 270
Пирсона 274, 281
 нецентральное 281
Стьюдента 274 - 275
Фишера — Снедекора 279
 нецентральное 281
- регрессионный анализ 207, 259
регулярность 20
 кривой 41
 поверхности 69, 83
 функции-отображения
 векторной 69, 83
 скалярной 41
резольвента пхп-матрицы 194
результат 2-х многочленов 152
- скорость сходимости 225, 236
стационарность функции
 маргинальная 35
 седловина 60
 стационарный перегиб 19, 21
 условная 72, 69, 84, 87, 96, 99
- теорема Банаха 225
теорема Вейерштрасса 18, 56
теорема Ферма 19
теорема Штурма 161
теорема Эйлера — Ферма 58
- теоремы об условной стационарности
 Лагранжа 87
 1 и 2 (проективные) 76, 84
 3 и 4 (предельные) 99, 106
теоремы 5 – 8 (формального анализа) 124, 126, 131, 134
теорема 9 (о средних) 146
теорема 10 (о корректности) 190
- уровень экстремума 27
 нецелочисленный $p > 1$ 33
 маргинального $1 < p < 2$ 34
 целочисленный $p \geq 2$ 22, 24
- целевая функция** 18
 отклика 254
- экстремум целевой функции**
 локальный и глобальный 19
 маргинальный 34
 неявной функции 54
 нестрогий 21
 обратной функции 49
 строгий 21
от вещественной переменной 18
 дискретной 39
от комплексных переменных 124, 131
от смешанных переменных 141
условный от
 векторной вещественной переменной
 зависимой 67
 ограниченной 82, 97
 клеточный 96
комплексных переменных
 внешне ограниченных 136
 внутренне ограниченных 139

Оглавление

К читателям	3
Resume	4
Введение	5
Используемые обозначения	10
Глава 1. Аналитическая безусловная оптимизация	
§ 1.1. Экстремумы целочисленных уровней $p \geq 2$ для функций от независимой скалярной переменной	17
§ 1.2. Экстремумы нецелочисленных уровней $p > 1$	33
§ 1.3. Маргинальные ($1 < p < 2$) и особые ($p \leq 1$) экстремумы	34
§ 1.4. Экстремумы для функций от дискретной переменной	39
§ 1.5. Экстремумы 1-й и 2-й степени для функций от зависимой скалярной переменной типа $\tilde{x}: \tilde{x} = x(u)$	40
§ 1.6. Экстремумы 1-й и 2-й степени для функций от зависимой скалярной переменной типа $\tilde{x}: u = u(\tilde{x})$	45
§ 1.6.1. Изопараметрические многочлены	47
§ 1.7. Экстремумы для функции от независимой скалярной переменной, заданной через обратную функцию	49
§ 1.7.1. Зеркальные изопараметрические многочлены	53
§ 1.8. Экстремумы для неявных функций от независимой скалярной переменной типа $x: f(y, x) = 0$	54
§ 1.9. Экстремумы уровня $p = 2$ для функций от независимой векторной переменной в аффинном пространстве $\mathcal{A}^n$	56
Глава 2. Аналитическая условная оптимизация	
§ 2.1. Условные экстремумы уровня $p = 2$ для функций от зависимой переменной типа $\tilde{x} \in \mathcal{A}^n: \tilde{x} = x(u)$, $u \in \mathcal{A}^q, q \leq n$	67
§ 2.2. Условные экстремумы уровня $p = 2$ для функций от ограниченной переменной типа $\tilde{x} \in \mathcal{A}^n: h(\tilde{x}) = 0$, $h \in \mathcal{A}^m, m \leq n$	83
§ 2.3. Клеточный условный экстремум	96
§ 2.4. Предельные методы решения задач на условный экстремум с малым или с большим параметром	97
§ 2.5. Условное характеристическое (вековое) уравнение	110

Глава 3. Аналитическая оптимизация целевых функций от разнообразных комплексных переменных

§ 3.1. Два альтернативных варианта комплексификации	113
§ 3.2. Формальное дифференцирование и интегрирование в комплексном пространстве	118
§ 3.3. Экстремумы для функций от пары независимых одномерных комплексных сопряжённых переменных	124
§ 3.4. Экстремумы для функций от пары независимых многомерных комплексных сопряжённых переменных	131
§ 3.5. Условные экстремумы для функций от ограниченных внешне комплексных переменных	136
§ 3.6. Условные экстремумы для функций от ограниченных внутренне комплексных переменных	139
§ 3.7. Экстремумы для функций от вещественных и комплексных сопряжённых переменных	141

Глава 4. Аналитическая оптимизация в общей и линейной алгебре

§ 4.1. Генеральное неравенство для средних величин	144
§ 4.2. Экстремальные корни алгебраического уравнения	151
§ 4.3. Инверсия генерального неравенства	164
§ 4.4. Полные требования к коэффициентам алгебраического уравнения для положительности всех его корней	165
§ 4.5. Нормальное решение и квазирешение вырожденного линейного уравнения – вещественного и комплексного	191
§ 4.6. Экстремумы отношения Релея	196
§ 4.7. Метод наименьших квадратов Гаусса в одномерном и многомерном вариантах	204

Глава 5. Численные методы оптимизации

§ 5.1. Общие положения	212
§ 5.2. Итерационная одномерная оптимизация	214
§ 5.3. Методы дихотомии и золотого сечения 0-го порядка	216
§ 5.4. Метод парных касательных 1-го порядка	220
§ 5.5. Метод Ньютона 2-го порядка и его разностные аналоги	223
§ 5.6. Итерационная многомерная оптимизация	227
§ 5.7. Покоординатные методы 0-го порядка	228
§ 5.8. Градиентный метод Коши 1-го порядка	230
§ 5.9. Масштабно-градиентный метод неполного 2-го порядка и его директивная модификация	233
§ 5.10. Общий метод Ньютона 2-го порядка и его директивная модификация	235
§ 5.11. Регуляризация по Тихонову в методах 2-го порядка	238

§ 5.12. Условная численная оптимизация	239
§ 5.12.1. Методы 1-го и 2-го порядка для функций от линейно зависимой или ограниченной векторной переменной	240
§ 5.12.2. Методы нормальных проекций 1 и 2-го порядка	242
§ 5.12.3. Методы с большим и малым параметром	243
Глава 6. Планово-вычислительная n-мерная оптимизация по минимальным планам	
§ 6.1. Общие положения	245
§ 6.2. Критерий адекватности планово-разностных моделей	246
§ 6.3. Вычислительная оптимизация по плану квадрантный n -симплекс с разностной моделью 1-го порядка	247
§ 6.4. Вычислительная оптимизация по плану осевой n -крест с разностной моделью неполного 2-го порядка	249
§ 6.5. Вычислительная оптимизация по плану n -АКП с разностной моделью 2-го порядка	251
Глава 7. Планово-экспериментальная n-мерная оптимизация по ортогональным планам	
§ 7.1. Общие положения	254
§ 7.2. Полифакторная линейная регрессия	259
§ 7.3. Гипотеза о нормальности распределения случайной ошибки и её статистическая проверка	269
§ 7.4. Критерий адекватности планово-регрессионных моделей	280
§ 7.5. Экспериментальная оптимизация неполного 2-го порядка по плану осевой n -крест	286
§ 7.6. Экспериментальная оптимизация 2-го порядка	291
§ 7.6.1. Оптимизация по плану n -СКП	291
§ 7.6.2. Оптимизация по плану n -ЦКП Бокса — Уилсона	297
§ 7.7. Сравнительный анализ планов эксперимента	305
Приложение. Физико-математическая кунсткамера	314
Список литературы	322
Именной указатель	326
Предметный указатель	328

Научное издание

Нинул Анатолий Сергеевич

ОПТИМИЗАЦИЯ ЦЕЛЕВЫХ ФУНКЦИЙ

Аналитика. Численные методы. Планирование эксперимента.

Оригинал-макет *О. К. Макаренко*

Оформление обложки *А. С. Лунин*

ИД № 01389 от 30.03.2000

Гигиенический сертификат № 77.99.10.953.Д.005466.07.03 от 25.07.2003

Подписано в печать 25.03.2009 г. Формат 60×90 1/16.

Бумага офсетная № 1. Печать офсетная.

Усл. печ. л. 21,0. Уч.-ихд. л. 23,1. Тираж 1000 экз.

Заказ № 53

Издательство физико-математической литературы (ФИЗМАТЛИТ)

123182 Москва, ул. Щукинская, д. 12, к.1

Отпечатано с готовых диапозитивов ГП «ОБЛИЗДАТ»

248640 г. Калуга, пл. Старый торг, 5.