ФИНАНСОВАЯ АКАДЕМИЯ
ПРИ ПРАВИТЕЛЬСТВЕ РОССИЙСКОЙ ФЕДЕРАЦИИ

Кафедра «Математика и финансовые приложения»

В.Б. Гисин

Лекции по дискретной математике

(часть 2)

МОСКВА 2002 ГОД

Аннотация

Пособие предназначено студентам, изучающим дискретную математику, и преподавателям, проводящим занятия по указанной дисциплине. Дисциплина «Дискретная математика» является обязательной для студентов дневного отделения института «Антикризисное управление и математические методы в экономике». Настоящее пособие содержит вторую часть курса лекций по дискретной математике. В этой части излагаются комбинаторика (включая производящие функции, рекуррентные последовательности и числа Фибоначчи) и теория графов. Значительное место уделено финансовым приложениям изучаемого материала.

Оглавление

8. Элементы комбинаторики

8. 1. Предварительные сведения

Целью комбинаторики является изучение комбинаторных конфигураций – конструкций, построенных из элементов некоторого (как правило, конечного) множества в соответствии с заданными правилами. Важный класс комбинаторных задач составляют задачи перечисления и, в частности, определения числа конфигураций данного типа. Простейшими примерами комбинаторных конфигураций являются *перестановки*, *размещения* и *сочетания*, представляющие собой различного рода выборки.

Конечное множество, содержащее n элементов, в комбинаторике принято называть n-множеством. Различают *упорядоченные* и *неупорядоченные* выборки из n-множества, выборки *с возвращением* и выборки *без возвращения*. К числу основных задач комбинаторики относятся задачи подсчета числа выборок заданного типа и заданного объема из произвольного n-множества.

Решение многих комбинаторных задач основывается на двух фундаментальных правилах, называемых соответственно правилами суммы и произведения.

Правило суммы. Если X_1 *и* X_2 – непересекающиеся конечные множества, содержащие соответственно n_1 *и* n_2 элементов, то объединение $X_1 \cup X_2$ содержит $n_1 + n_2$ элементов.

Сформулированное правило можно распространить на случай произвольного числа слагаемых:

если множества $X_1, X_2, \ldots, X_k$ образуют разбиение множества X, то $n = n_1 + n_2 + \ldots + n_k$, где $n = |X|$ – число элементов множества X, а $n_i = |X_i|$ – число элементов множества X_i, $i = 1, 2, \ldots, k$.

Правило суммы можно приспособить и для подсчета числа элементов объединения двух множеств с непустым пересечением:

$$|X_1 \cup X_2| = |X_1| + |X_2| - |X_1 \cap X_2|. \qquad (1)$$

В самом деле, множества $X_1 \backslash X_2$ и $X_1 \cap X_2$ образуют разбиение множества X_1, поэтому

$$|X_1| = |X_1 \backslash X_2| + |X_1 \cap X_2|.$$

Аналогично

$$|X_2| = |X_2 \backslash X_1| + |X_1 \cap X_2|.$$

Кроме того, множества $X_1 \backslash X_2$, $X_2 \backslash X_1$ и $X_1 \cap X_2$ образуют разбиение множества $X_1 \cup X_2$, так что

$$|X_1 \cup X_2| = |X_1 \backslash X_2| + |X_2 \backslash X_1| + |X_1 \cap X_2|.$$

Соотношение (1) является следствием трех последних формул.

Формула (1) обобщается на произвольное число слагаемых с помощью так называемого *принципа включения и исключения* (см. п. 8.4).

Пример. Найдем число правильных несократимых дробей со знаменателем 100. Числителем правильной несократимой дроби со знаменателем 100 должно быть целое число в промежутке от 1 до 100, не имеющее среди своих делителей чисел 2 и 5. Обозначим через X_2, X_5 и X_{10} множества чисел из промежутка от 1 до 100, делящихся соответственно на 2, на 5 и на 10. Легко видеть, что $|X_2| = 100/2 = 50$ и $|X_5| = 100/5 = 20$ и $|X_{10}| = 100/10 = 10$. Так как $X_2 \cap X_5 = X_{10}$, то по формуле (1) получаем $|X_2 \cup X_5| = 50 + 20 - 10 = 60$. Таким образом, среди чисел от 1 до 100 имеется 60 чисел, делящихся на 2 или 5, и $100 - 60 = 40$ чисел, взаимно простых с числом 100. Следовательно, среди правильных дробей со знаменателем 100 ровно 40 являются несократимыми.

Правило произведения. Будем рассматривать упорядоченные наборы (кортежи) вида $(a_1, a_2,\ldots, a_k)$ заданной длины k. Предположим, что элемент a_1 может быть выбран n_1 способами; при фиксированном a_1 элемент a_2 может быть выбран n_2 способами; при фиксированных a_1 и a_2 элемент a_3 может быть выбран n_3 способами и т. д. (при фиксированных $a_1, a_2,\ldots, a_{k-1}$ элемент a_k может быть выбран n_k способами). Тогда число различных кортежей равно произведению $n_1 n_2 \cdot\ldots\cdot n_k$. В частности, по правилу произведения для конечных множеств $X_1, X_2, \ldots, X_k$ имеем:

$$|X_1 \times X_2 \times \ldots \times X_k| = |X_1| \times |X_2| \times \ldots \times |X_k|.$$

Пример. Используя правило произведения и правило суммы, найдем число различных трехзначных чисел, не содержащих одинаковых цифр, и число различных трехзначных чисел, содержащих хотя бы две одинаковые цифры.

Пусть a_1, a_2, a_3 – цифры трехзначного числа. Первую цифру a_1 можно выбрать девятью способами (в качестве нее можно взять любую цифру, кроме нуля); при фиксированной первой цифре вторую цифру a_2 можно выбрать также девятью способами (в качестве нее можно взять любую цифру, кроме a_1); наконец, при фиксированных первой и второй цифрах третью цифру можно выбрать восемью способами. По правилу произведения число трехзначных чисел, не содержащих одинаковых цифр, равно $9{\cdot}9{\cdot}8 = 648$. Всего имеется 900 различных трехзначных чисел. Каждое из них либо содержит две одинаковых цифры, либо нет. Следовательно, имеется 900 – 648 = 252 различных трехзначных числа, содержащих хотя бы две одинаковые цифры.

Факториалы. Напомним, что факториалом целого положительного числа n (обозначение $n!$) называется произведение $1{\cdot}2{\cdot}\ldots{\cdot}n$. По определению считается, что $0! = 1$. Факториал как функция натурального аргумента может быть однозначно определен рекурсивно:

$$0! = 1;\ (n + 1)! = n!{\cdot}(n + 1).$$

Приближенное значение $n!$ при больших n дает следующая *формула Стирлинга*:

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n. \qquad (2)$$

Символ ~ означает здесь, что отношение

$$\sqrt{2\pi n}\left(\frac{n}{e}\right)^n / n!$$

стремится к единице при n, стремящемся к бесконечности. Хотя разность правой и левой частей в формуле Стирлинга неограниченно возрастает, относительная ошибка быстро убывает. Например, уже при $n = 10$ относительная ошибка составляет менее одного процента:

$$10! = 3628800; \qquad \sqrt{2\pi \cdot 10} \cdot \left(\frac{10}{e}\right)^{10} \approx 3598696.$$

8. 2. Размещения и перестановки

Размещения с повторениями объема k из n-множества X, или размещения с повторениями из n элементов по k, – это упорядоченные выборки вида $(a_1, a_2, \ldots, a_k)$, где a_i, $i = 1, 2, \ldots, k$, – произвольные элементы множества X.

Каждый элемент a_i может быть выбран n способами. По правилу произведения общее число размещений с повторениями из n элементов по k равно n^k.

Размещения (без повторений) объема k из n-множества X, или *размещения* из n элементов по k, – это упорядоченные выборки вида $(a_1, a_2, \ldots, a_k)$, где a_i, $i = 1, 2, \ldots, k$, – различные элементы множества X.

Элемент a_1 может быть выбран n способами. При $n \geq 2$ элемент a_2 может быть выбран $n - 1$ способами: в качестве a_2 можно взять любой элемент из $(n - 1)$-множества $X \backslash \{a_1\}$. По правилу произведения число упорядоченных пар различных элементов вида (a_1, a_2) равно $n(n - 1)$. Аналогично (при $n \geq 3$) элемент a_3 может быть произвольно выбран из $(n - 2)$-множества $X \backslash \{a_1, a_2\}$, так что число упорядоченных троек (a_1, a_2, a_3) равно $n(n - 1)(n - 2)$.

Число размещений из n по k, $n \geq k$, обозначается через A_n^k. Как мы видели, $A_n^1 = n$. Естественно считать, что

$$A_n^0 = 1 \qquad (3)$$

(имеется ровно одна, не содержащая ни одного элемента, т. е. пустая выборка).

Несложно показать, что при $k < n$ имеет место соотношение

$$A_n^{k+1} = (n - k) A_n^k. \qquad (4)$$

В самом деле, упорядоченную выборку вида $(a_1, a_2, \ldots, a_k, a_{k+1})$ можно рассматривать как упорядоченную пару, в которой первым элементом служит размещение $(a_1, a_2, \ldots, a_k)$, а вторым – a_{k+1}. Размещение из n элементов по k может быть составлено A_n^k способами, элемент a_{k+1} может быть произвольно выбран из $(n - k)$-множества $X \backslash \{ a_1, a_2, \ldots, a_k\}$. Формула (3) получается теперь по правилу произведения.

Используя формулы (3) и (4), получаем:

$$A_n^1 = n\,,\ A_n^2 = n(n-1)\,,\ A_n^3 = n(n-1)(n-2)\,,\ \ldots\,,$$

и, в общем случае:

$$A_n^k = n(n-1)\cdot\ldots\cdot(n-k+1)\,. \tag{5}$$

Предыдущую формулу можно записать следующим образом:

$$A_n^k = \frac{n!}{(n-k)!}\,. \tag{6}$$

Пример. Найдем число всех отображений и число инъективных отображений из k-множества X в n-множество Y.

Перенумеруем элементы множества X, $X = \{a_1, a_2, \ldots, a_k\}$. Всякое отображение f из X в Y однозначно определяется упорядоченным набором элементов множества Y вида $(b_1, b_2, \ldots, b_k)$, в котором $b_i = f(a_i)$. Значит, число всех отображений из X в Y равно числу размещений с повторениями из n элементов по k, то есть числу n^k. Таким образом, получаем формулу:

$$|Y^X| = n^k,$$

где через Y^X обозначено множество всех отображений из X в Y.

Если $k > n$, инъективных отображений из X в Y нет. При $k \leq n$ отображение f из X в Y инъективно, если в наборе $(b_1, b_2, \ldots, b_k)$ все элементы различны, т. е. набор $(b_1, b_2, \ldots, b_k)$ является размещением без повторений. Следовательно, число инъективных отображений из X в Y равно числу размещений A_n^k.

Размещения объема n, составленные из элементов n-множества X называются *перестановками* элементов множества X.

Задать перестановку элементов множества X – это значит линейно упорядочить элементы этого множества. Число всех перестановок из n элементов обозначается через P_n. Используя формулу (5) при $k = n$, получаем:

$$P_n = n!. \tag{7}$$

Пример. Перечислим перестановки элементов множества $X = \{1, 2, 3\}$:

123; 132; 213; 231; 312; 321.

В заключение рассмотрим *перестановки с повторениями*. Пусть $X = \{a_1, a_2,\ldots, a_k\}$ – некоторое k-множество. Перестановкой с повторениями типа $(n_1, n_2,\ldots, n_k)$ называется упорядоченный набор элементов множества X вида $(x_1, x_2,\ldots, x_n)$, в котором элемент a_1 встречается n_1 раз, элемент a_2 встречается n_2 раз, …, элемент a_k встречается n_k раз. Естественно, $n_1 + n_2 +\ldots+ n_k = n$. Число перестановок типа $(n_1, n_2,\ldots, n_k)$ обозначается через $P(n_1, n_2,\ldots, n_k)$.

Пример. Перечислим все перестановки типа (2, 2) из элементов множества $X = \{a, b\}$:

aabb; *abab*; *abba*; *baab*; *baba*; *bbaa*.

Чтобы получить формулу для вычисления $P(n_1, n_2,\ldots, n_k)$, воспользуемся следующей метафорой. Элементы множества X будем считать буквами, а перестановки с повторениями – словами. Для определенности будем считать, что X содержит три буквы a, b, c. Рассмотрим произвольное слово w длины n, содержащее n_1 букв a, n_2 букв b и n_3 букв c. Подсчитаем число перестановок букв, при которых слово w не меняется. Буквы a можно произвольно переставлять на n_1 местах, на которых они расположены. Это можно сделать $n_1!$ числом способов. Точно так же буквы b можно переставить $n_2!$ способами, буквы c – $n_3!$ способами. По правилу произведения число перестановок букв, при которых слово w не меняется, равно $n_1!n_2!n_3!$. Таким образом, для каждого слова имеется $n_1!n_2!n_3!$ перестановок букв, которые его не меняют. Общее число перестановок букв равно $n!$. Следовательно, число слов, т. е. перестановок типа (n_1, n_2, n_3), равно

$$\frac{n!}{n_1!n_2!n_3!}.$$

Предыдущее рассуждение легко переносится на общий случай и приводит к следующей формуле:

$$P(n_1,n_2,\ldots,n_k)=\frac{n!}{n_1!n_2!\ldots n_k!}. \qquad (8)$$

Пример. Пакет акций состоит из акций четырех типов: 5 акций типа A, 3 акции типа B и по одной акции типов C и D.

Сколькими способами можно распродать этот пакет, продавая по одной акции ежедневно?

Каждая распродажа однозначно определяется словом длины 10, составленном из пяти букв A, трех букв B, одной буквы C и одной буквы D (буква, стоящая на месте i, означает, что в день i была продана акция соответствующего типа). Значит, число всевозможных распродаж составляет

$$P(5,3,1,1) = \frac{10!}{5!\cdot 3!\cdot 1!\cdot 1!} = 5040.$$

8. 3. Сочетания

Сочетания объема k из n-множества X – это неупорядоченные выборки вида $\{a_1, a_2,\ldots, a_k\}$, где a_i, $i = 1, 2,\ldots, k$, – попарно различные элементы множества X.

В соответствии с определением сочетания объема k – это k-подмножества множества X. Число сочетаний объема k из n-множества X обозначают через C_n^k или через $\binom{n}{k}$ и называют *числом сочетаний из n по k*.

Легко видеть, что $C_n^0 = 1$, $C_n^1 = n$, $C_n^n = 1$. Чтобы получить формулу для числа сочетаний в общем случае, докажем сначала, что при $k \le n$ справедливо следующее соотношение:

$$A_n^k = C_n^k \cdot P_k.$$

В самом деле, пусть X – n-множество. Каждому размещению

$$(a_1, a_2,\ldots, a_k)$$

поставим в соответствие сочетание

$$\{a_1, a_2, \ldots, a_k\}.$$

Одно и то же сочетание $\{a_1, a_2, \ldots, a_k\}$ соответствует всем размещениям, полученным перестановкой элементов $a_1, a_2, \ldots, a_k$. Число таких перестановок равно P_n. Таким образом, вся совокупность размещений по k элементов разбивается на C_n^k классов, каждый из которых содержит P_n размещений. Отсюда и следует доказываемая формула.

Используя (6), (7) получаем выражение для числа сочетаний «в факториалах»:

$$C_n^k = \frac{n!}{k!(n-k)!}. \qquad (9)$$

Часто используется и следующее соотношение:

$$C_n^k = \frac{n(n-1)\cdot\ldots\cdot(n-k+1)}{k!}. \qquad (10)$$

Пример. Найдем число сочетаний из 52 по 26.

В соответствии с (9) имеем:

$$C_{52}^{26} = \frac{52!}{(26!)^2}.$$

Используя формулу Стирлинга можно получить приблизительную оценку:

$$\ln C_{52}^{26} \approx 0{,}5(\ln 2\pi + \ln 52) + 52(\ln 52 - 1) -$$

$$- (\ln 2\pi + \ln 26) - 52(\ln 26 - 1)$$

$$= 53\ln 2 - 0{,}5\ln(52\cdot 2\pi) \approx 33{,}84.$$

Значит,

$$C_{52}^{26} \approx e^{33,84} \approx 4,98 \cdot 10^{14}$$

(прямое вычисление дает $C_{52}^{26} \approx 4,96 \cdot 10^{14}$).

Пример. Найдем число монотонно возрастающих отображений из множества

$$X = \{1, 2, \ldots, k\}$$

в множество

$$Y = \{1, 2, \ldots, n\}.$$

Любое такое отображение f задает набор чисел

$$f(1) < f(2) < \ldots < f(k).$$

Обратно, любое k-подмножество $Y' = \{y_1, y_2, \ldots, y_k\}$ множества Y однозначно определяет монотонно возрастающее отображение из X в Y, для которого $f(X) = Y'$. В самом деле, упорядочим элементы множества Y' по возрастанию так, что $Y' = \{y_1, y_2, \ldots, y_k\}$, причем $y_1 < y_2 < \ldots < y_k$. Полагая $f(i) = y_i$, получаем соответствующее отображение. Таким образом, число монотонно возрастающих отображений из X в Y равно числу k-подмножеств множества Y, т. е. числу C_n^k – числу сочетаний из n по k.

Рассмотрим теперь вопрос о числе сочетаний с повторениями. *Сочетания с повторениями* объема k из n-множества X – это неупорядоченные выборки вида $a_1\, a_2 \ldots a_k$, где a_i, $i = 1, 2, \ldots, k$, – произвольные элементы множества X, среди которых могут быть и совпадающие. Примерами сочетаний с повторениями из элементов a, b, c, d по три могут служить выборки aaa; aab; dab.

Предположим для определенности, что $X = \{1, 2,\dots, n\}$. Сочетание с повторениями из элементов множества X однозначно определяется упорядоченным набором неотрицательных целых чисел $(x_1, x_2,\dots, x_n)$, где x_i, $i = 1, 2,\dots, n$, равно числу вхождений в выборку элемента i. Например, выборка 1, 3, 2, 1, 3, 3 из множества $X = \{1, 2, 3, 4\}$ задается набором (2, 1, 3, 0), поскольку 1 встречается в ней два раза, 2 – один раз, 3 – три раза, и 4 – ни разу. Сумма $x_1 + x_2 +\dots+ x_n$ – это объем выборки. Таким образом, число сочетаний с повторениями из n элементов по k равно числу решений уравнения

$$x_1 + x_2 +\dots+ x_n = k \qquad (11)$$

в целых неотрицательных числах. Сделаем замену неизвестных. Пусть $y_i = x_i + 1$. Относительно новых неизвестных уравнение (11) запишется в следующем виде:

$$y_1 + y_2 +\dots+ y_n = n + k. \qquad (12)$$

Нас интересует число решений этого уравнения в целых положительных числах. Если взять строчку из $n + k$ точек и в промежутках между ними поставить $n - 1$ разделительную черточку, то можно считать y_1 числом точек в первой группе, y_2 – числом точек во второй группе и т. д.:

$$\underbrace{\bullet\bullet\bullet\bullet}_{y_1}|\underbrace{\bullet\bullet\bullet\bullet\bullet}_{y_2}|\cdots|\underbrace{\bullet\bullet\bullet\bullet}_{y_n}\ .$$

Таким образом, число решений уравнения (12) равно числу способов расставить $n-1$ черточек в $n+k-1$ промежутках между точками. Число способов выбрать n – 1 промежутков (в которые ставятся черточки) из $n+k-1$ промежутков равно числу сочетаний из $n+k-1$ по $n-1$, то есть числу C_{n+k-1}^{n-1}. Следовательно, число сочетаний с повторениями из n элементов по k равно C_{n+k-1}^{n-1}.

Пример. Найдем число способов разложить 10 одинаковых монет по трем карманам. Каждый расклад монет представляет собой сочетание с повторениями из трех элементов по 10. Каждый «карман» входит в выборку столько раз, сколько монет в него положено. Таким образом, искомое число способов расклада равно $C_{12}^{2}=\dfrac{12\cdot 11}{2}=66$.

Используя сочетания можно получить новое обоснование формулы для числа перестановок с повторениям (8). В самом деле, пусть рассматриваются перестановки с повторениями типа $(n_1, n_2,\ldots, n_k)$ из элементов множества $X=\{x_1, x_2,\ldots, x_k\}$. Для определенности будем считать, что $k=3$. Чтобы составить перестановку с повторениями, нужно указать места, на которых расположены элементы x_1; места, на которых расположены элементы x_2, места, на которых расположены элементы x_3. Для расстановки элементов x_1 можно выбрать любые n_1

из $n_1 + n_2 + \ldots + n_k$ мест. Это можно сделать $C^{n_1}_{n_1+n_2+\ldots+n_k}$ способами. Элементы x_2 можно расставлять на незанятые места, которых остается $n_2 + \ldots + n_k$. Значит, места для элементов x_2 можно выбрать $C^{n_2}_{n_2+\ldots+n_k}$ способами. Продолжая подобным образом, приходим к тому, что элементы x_{k-1} должны быть расставлены на $n_{k-1} + n_k$ мест, а элементы x_k – на оставшиеся n_k мест. По правилу произведения получаем:

$$P(n_1, n_2, \ldots, n_k) = C^{n_1}_{n_1+n_2+\ldots+n_k} \cdot C^{n_2}_{n_2+\ldots+n_k} \cdot \ldots \cdot C^{n_k}_{n_k}.$$

Если теперь записать числа сочетаний в факториалах, то после сокращений получится (8).

Заметим, кстати, что $C_n^k = P(k, n-k)$.

8. 4. Некоторые свойства чисел C_n^k

Непосредственно из формулы (8) легко усмотреть, что при любых n и k, $0 \le k \le n$, верно равенство

$$C_n^k = C_n^{n-k}. \tag{13}$$

Если $0 \le k < n$, то

$$C_n^k + C_n^{k+1} = C_{n+1}^{k+1}. \tag{14}$$

С использованием (9) доказательство может быть получено непосредственной проверкой:

$$C_n^k + C_n^{k+1} = \frac{n!}{k!(n-k)!} + \frac{n!}{(k+1)!(n-k-1)!} = \frac{n!(k+1)+n!(n-k)}{(k+1)!(n-k)!} =$$

$$= \frac{n!(n+1)}{(k+1)!(n-k)!} = \frac{(n+1)!}{(k+1)!(n-k)!} = C_{n+1}^{k+1}.$$

Сумма

$$C_n^0 + C_n^1 + C_n^2 + \ldots + C_n^n \qquad (15)$$

– это число всех подмножеств n-множества (число пустых подмножеств C_n^0 плюс число одноэлементных подмножеств C_n^1 и т. д.). Как было установлено ранее, n-множество имеет 2^n подмножеств. Следовательно,

$$C_n^0 + C_n^1 + C_n^2 + \ldots + C_n^n = 2^n. \qquad (16)$$

Впрочем, с использованием (14) формула (16) может быть получена по индукции. При малых значениях n справедливость (16) легко проверяется непосредственно. Индуктивный шаг сводится к следующей несложной выкладке:

$$C_{n+1}^0 + C_{n+1}^1 + C_{n+1}^2 + \ldots + C_{n+1}^{n+1} =$$

$$= C_{n+1}^0 + (C_n^0 + C_n^1) + (C_n^1 + C_n^2) + \ldots + (C_n^{n-1} + C_n^n) + C_{n+1}^{n+1} =$$

$$= (C_n^0 + C_n^1 + \ldots + C_n^{n-1} + C_{n+1}^{n+1}) + (C_{n+1}^0 + C_n^1 + \ldots + C_n^n) =$$

$$= 2(C_n^0 + C_n^1 + C_n^2 + \ldots + C_n^n).$$

В последнем переходе использовалось то, что $C_{n+1}^{n+1} = 1 = C_n^n$ и $C_{n+1}^0 = 1 = C_n^0$.

Формула (14) позволяет последовательно заполнять строки таблицы, называемой *треугольником Паскаля*. Рассмотрим первые пять строк этой таблицы:

$$\begin{array}{c|ccccc} 0 & 1 \\ 1 & 1 & 1 \\ 2 & 1 & 2 & 1 \\ 3 & 1 & 3 & 3 & 1 \\ 4 & 1 & 4 & 6 & 4 & 1 \end{array}$$

В строке с номером n (указанном слева от вертикальной черты) расположены числа C_n^0, C_n^1, …, C_n^n. В соответствии с (14), складывая два соседних числа в строке n, мы получим число в строке $n + 1$, стоящее под правым слагаемым.

Рассмотрим *бином* $(x + y)^n$. Для любого целого неотрицательного числа n имеет место следующее тождество:

$$(x+y)^n = \sum_{k=0}^{n} C_n^k x^{n-k} y^k , \qquad (17)$$

или в развернутом виде:

$$(x+y)^n = C_n^0 x^n y^0 + C_n^1 x^{n-1} y^1 + \ldots + C_n^n x^0 y^n ,$$

называемое *формулой Ньютона*.

Доказательство этого факта несложно провести по индукции, используя (14). Формула Ньютона, очевидно, верна при $n = 0$. Чтобы выполнить индуктивный шаг, в равенстве

$$(x + y)^{n+1} = (x + y)^n(x + y)$$

заменим $(x+y)^n$ суммой из правой части (17). После раскрытия скобок и приведения подобных получаем:

$$(x+y)^{n+1} =$$

$$= xC_n^0 x^n y^0 + \sum_{k=1}^{n}\left(xC_n^k x^{n-k} y^k + yC_n^{k-1} x^{n-k+1} y^{k-1}\right) + yC_n^n x^0 y^n =$$

$$= x^{n+1} + \sum_{k=1}^{n}\left(C_n^k + C_n^{k-1}\right) x^{n-k+1} y^k + y^{n+1}$$

В соответствии с (14),

$$C_n^k + C_n^{k-1} = C_{n+1}^k,$$

так что

$$(x+y)^{n+1} = C_{n+1}^0 x^{n+1} y^0 + C_{n+1}^1 x^n y^1 + \ldots + C_{n+1}^{n+1} x^0 y^{n+1}$$

и, следовательно, формула Ньютона верна и при показателе степени равном $n+1$.

Подставка в формулу Ньютона конкретных числовых значений переменных позволяет получать различные тождества. Вот некоторые из них:

$$\sum_{k=0}^{n} C_n^k = 2^n;$$

$$\sum_{k=0}^{n} (-1)^k C_n^k = 0;$$

$$\sum_{k=0}^{n} 2^k C_n^k = 3^n.$$

В первом случае (полученное тождество совпадает с (16)) мы подставили в (17) $x = 1$, $y = 1$; во втором – $x = 1$, $y = –1$; в третьем – $x = 1$, $y = 2$.

8. 5. Принцип включения и исключения

Пусть X – некоторое n-множество, а p_1, p_2, …, p_k – свойства, которыми могут обладать элементы множества X. Каждый элемент множества X может обладать одним или несколькими свойствами из числа перечисленных или не обладать ни одним из них. Обозначим через n_i число элементов со свойством p_i, $i = 1, 2,\ldots, k$. Вообще, пусть $n_{i_1 i_2 \ldots i_r}$ – число элементов со свойствами p_{i_1}, p_{i_2}, …, p_{i_r}. Тогда число элементов n_0, не обладающих ни одним из перечисленных свойств, задается следующей формулой:

$$n_0 = n - \sum_i n_i + \sum_{i_1 < i_2} n_{i_1 i_2} + \ldots$$

$$\ldots + (-1)^{\mathrm{r}} \sum_{i_1 < i_2 < \ldots < i_r} n_{i_1 i_2 \ldots i_r} + \ldots + (-1)^n n_{12\ldots n}. \qquad (18)$$

Элемент a, не обладающий ни одним свойством, учитывается один раз в слагаемом n и ни одно другое слагаемое вклад не вносит. Элемент a, обладающий ровно одним свойством j, учитывается по одному разу в слагаемых n и $\sum\limits_i n_i$; его вклад в правую часть (18) составляет 1 – 1=0. Рассмотрим теперь элемент a, обладающий $s \geq 1$ свойствами $j_1,\ldots,j_s$.

Элемент a дает вклад 1 в каждое слагаемое $\sum\limits_{i_1<i_2<\ldots<i_r} n_{i_1 i_2 \ldots i_r}$ такое, что $\{i_1,\ldots, i_r\} \subset \{j_1,\ldots,j_s\}$. Для каждого фиксированного $r \le s$ число таких слагаемых равно числу r-подмножеств s-множества $\{j_1,\ldots,j_s\}$, т. е. числу сочетаний C_s^r. Следовательно, вклад элемента a в правую часть (18) равен

$$1 - C_s^1 + C_s^2 + \ldots + (-1)^r C_s^r + \ldots + (-1)^s C_s^s = (1-1)^s = 0.$$

Таким образом, правая часть (18) учитывает каждый элемент, не обладающий ни одним из указанных свойств ровно один раз, а каждый элемент, обладающий хотя бы одним свойством, – ноль раз. Но это и означает, что правая и левая части (18) равны.

В качестве применения метода включения и исключения рассмотрим известную *задачу о беспорядках*.

Перестановка $a_1 a_2 \ldots a_n$ чисел $1, 2, \ldots, n$ называется беспорядком, если $a_i \ne i$ для всех $i = 1, 2, \ldots, n$. Таким образом, беспорядок – это перестановка, в которой ни один элемент не занимает «своего» места. Обозначим число беспорядков через D_n.

Пример. Перечислим беспорядки из четырех элементов:

2143; 2341; 2413; 3142; 3412; 3421; 4123; 4312; 4321.

Значит, $D_4 = 9$.

На множестве всех перестановок из элементов $1, 2, \ldots, n$ определим набор свойств p_i, $i = 1, 2, \ldots, n$. Перестановка $a_1 a_2 \ldots a_n$ обладает свойством p_i, если $a_i = i$. Тогда число беспорядков – это число перестановок, не обладающих ни одним из указанных свойств. Число перестановок, оставляющих на месте ровно r фиксированных символов (например, $1, 2, \ldots, r$) равно $(n - r)!$. Выбрать r неподвижных символов можно C_n^r способами. Следовательно в формуле (18) каждое слагаемое $n_{i_1 i_2 \ldots i_r}$ равно $(n - r)!$, а общее число таких слагаемых C_n^r. Следовательно,

$$\sum_{i_1 < i_2 < \ldots < i_r} n_{i_1 i_2 \ldots i_r} = (n - r)!\,C_n^r = \frac{n!}{r!}.$$

Таким образом, в соответствии с формулой (18) имеем:

$$D_n = n! - \frac{n!}{1!} + \frac{n!}{2!} + \ldots + (-1)^r \frac{n!}{r!} + \ldots + (-1)^n \frac{n!}{n!}.$$

Последнее равенство можно переписать следующим образом:

$$\frac{D_n}{n!} = 1 - \frac{1}{1!} + \frac{1}{2!} + \ldots + (-1)^r \frac{1}{r!} + \ldots + (-1)^n \frac{1}{n!}. \qquad (19)$$

Как известно из анализа, правая часть (19) представляет собой приближение числа e^{-1} по формуле Тейлора. Следовательно, $\dfrac{D_n}{n!} \approx e^{-1}$ и, значит, $D_n \approx n!\,e^{-1}$. При этом абсолютная величина погрешности приближения не превосходит $\frac{1}{n+1}$.

Пример. При $n = 4$ имеем $D_4 \approx 4!e^{-1} = 24e^{-1} \approx 8.83$. Напомним, что выше мы непосредственно нашли точное значение $D_4 \approx 9$.

9. Биномиальная модель ценообразования активов

9. 1. Биномиальная решетка

Предположим, что за каждый период времени, скажем час, стоимость некоторого актива S_0 увеличивается или уменьшается. Пусть $u > 1$ – повышающий коэффициент, а $d < 1$ – понижающий, так что стоимость актива за единицу времени либо увеличивается и становится равной S_0u, либо уменьшается и становится равной S_0d. Если стоимость актива повышалась p раз и понижалась q раз, то его стоимость станет равной $S_0u^pd^q$. При $ud = 1$ изменение стоимости актива может быть графически представлено так называемой *биномиальной решеткой*. На рис. 1 приведена трехпериодная биномиальная решетка.

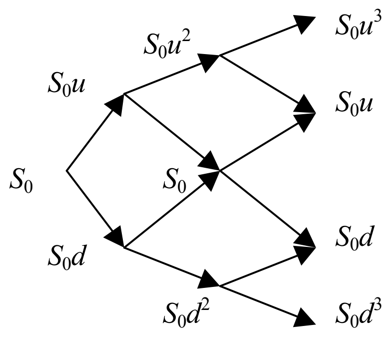


Рис. 1

Около каждого узла решетки указывается стоимость актива при соответствующем варианте развития событий. Максимальной стоимость актива окажется в том случае, когда в каждом периоде его стоимость повышалась. Вообще, каждому варианту развития событий, при которых стоимость актива окажется равной $S_0u^pd^q$, соответствует путь по стрелкам из начального узла S_0 в узел $S_0u^pd^q$. Сопоставим каждому пути последовательность букв d и u, в которой d соответствует понижению стоимости, а u – повышению. Число всевозможных путей из узла S_0 в узел $S_0u^pd^q$ равно числу способов расставить p букв u на $p + q$ местах, т. е. числу сочетаний C_{p+q}^{p}.

9. 2. Опционы. Основные понятия

Опционы относятся к числу так называемых *производных финансовых инструментов*.

Опцион «колл» (call option) дает его владельцу право (но, в отличие от контракта, не обязанность) купить определенный актив по фиксированной цене, называемой ценой исполнения, в установленное время в будущем или до наступления установленной даты. В роли актива могут выступать товар, ценная бумага, валюта и т. п. *Опцион «пут»* (put option) дает его владельцу аналогичное право продать актив. Реализация права на покупку (в случае опциона «колл») или продажу (в случае опциона «пут») называется исполнением опциона.

Различают *американские* и *европейски*е опционы. Если опцион может быть исполнен в любой момент до даты истечения срока действия, его называют американским. Европейский опцион может быть исполнен только в установленный срок. Не вдаваясь в детали опционной торговли, мы ограничимся рассмотрением европейских опционов.

Величина выигрыша от исполнения опциона «колл» равна

$$C = \max\{0, S - X\},$$

где S – цена актива в момент исполнения опциона, а X – цена исполнения опциона. В самом деле, если в момент исполнения опциона цена актива S превышает цену исполнения X, владелец опциона использует свое право купить опцион по цене ниже рыночной. При этом на разнице цен он получает выигрыш, равный $S - X$. В случае если на момент исполнения цена актива оказывается ниже цены исполнения, владелец опциона отказывается от покупки.

Пример. Стоимость актива в момент исполнения опциона «колл» равна 60, цена исполнения равна 50. Купив по цене $X = 50$ актив, продающийся на рынке по цене $S = 60$, владелец опциона получает выигрыш $C = 10$. Если бы стоимость актива в момент исполнения опциона оказалась ниже 50, владелец опциона не стал бы использовать свое право на покупку актива по цене $X = 50$.

Аналогичным образом величина выигрыша от исполнения опциона «пут» составляет

$$P = \max\{0, X - S\}.$$

За приобретение опциона, как и за всякое приобретение права без обязательств, нужно платить. Далее биномиальная модель используется для того, чтобы определить «справедливую» цену опциона, называемую *премией за опцион*.

9. 3. Однопериодная модель

Предположим, что инвестор в момент $t = 0$ покупает актив A по цене S_0. За единицу времени от $t = 0$ до $t = 1$ стоимость актива может повыситься или понизиться. Пусть u и d – повышающий и понижающий коэффициенты, r – безрисковая процентная ставка и выполняется следующее условие:

$$d < 1 + r < u.$$

Инвестор хочет избежать риска, связанного с изменением стоимости актива к моменту $t = 1$. С этой целью он продает некоторое количество k (быть может, дробное) опционов «колл» на актив A со сроком исполнения $t = 1$. Это означает, что инвестор принимает на себя обязательство продать k активов A в момент времени $t = 1$ по цене исполнения X. Требуется определить значение k и величину премии за опцион (в предположении, что транзакционные издержки отсутствуют).

Введем следующие обозначения (рис. 2):

S_0 – цена актива в момент времени $t = 0$;

$S_1(u) = S_0u$ – цена актива в момент времени $t = 1$ в случае повышения;

$S_1(d) = S_0d$ – цена актива в момент времени $t = 1$ в случае понижения;

C_0 – цена опциона в момент времени $t = 0$ (премия за опцион);

X – цена исполнения опциона;

$C_1(u) = \max\{0,\ S_0u - X\}$ – цена опциона в момент времени $t = 1$ в случае повышения;

$C_1(u) = \max\{0,\ S_0d - X\}$ – цена опциона в момент времени $t = 1$ в случае понижения.

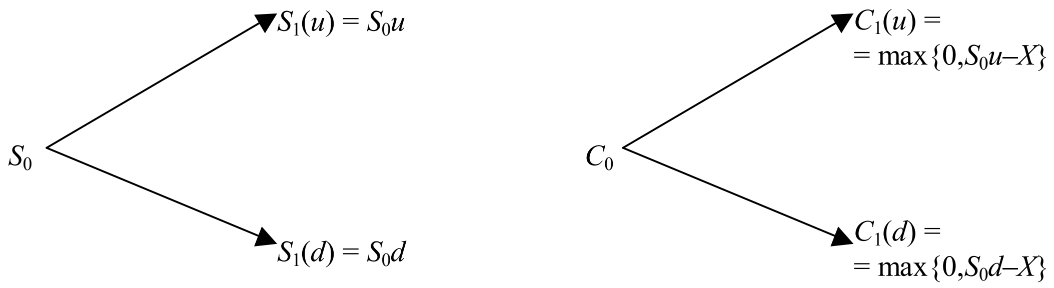


Рис. 2

На покупку актива инвестор затратил S_0, от продажи k опционов он получил kC_0. В результате он стал владельцем портфеля, состоящего из одного актива и k опционов. На формирование этого портфеля затрачено $S_0 - kC_0$. Величина $S_0 - kC_0$ составляет стоимость портфеля в момент времени $t = 0$. Если бы эта сумма была размещена под безрисковую процентную ставку, то к моменту времени $t = 1$ она бы выросла до

$$(1 + r)(S_0 - kC_0). \qquad (1)$$

С другой стороны, стоимость портфеля в момент времени $t = 1$ составит $S_1 - kC_1$, т. е.

$$S_1(u) - kC_1(u)$$

в случае повышения и

$$S_1(d) - kC_1(d)$$

в случае понижения. Если подобрать k так, чтобы выполнялось равенство

$$S_1(u) - kC_1(u) = S_1(d) - kC_1(d), \tag{2}$$

то портфель окажется безрисковым. Из (2) находим:

$$k = \frac{S_1(u) - S_1(d)}{C_1(u) - C_1(d)}. \tag{3}$$

При таком k стоимость портфеля в момент времени $t = 1$ окажется равной

$$\frac{S_1(d)C_1(u) - S_1(u)C_1(d)}{C_1(u) - C_1(d)}. \tag{4}$$

Так как портфель дает гарантированный доход, то в соответствии с принципом *безарбитражности* (прибыль без риска невозможна) его стоимость (4) должна совпадать наращенной суммой (1). Таким образом, получаем следующее равенство:

$$(1+r)\left(S_0 - \frac{S_1(u) - S_1(d)}{C_1(u) - C_1(d)} C_0 \right) = \frac{S_1(d)C_1(u) - S_1(u)C_1(d)}{C_1(u) - C_1(d)}.$$

После несложных преобразований, учитывая, что $S_1(u) = S_0u$ и $S_1(d) = S_0d$, находим:

$$C_0 = \frac{1}{1+r}\left(\frac{1+r-d}{u-d} C_1(u) + \frac{u-(1+r)}{u-d} C_1(d) \right). \tag{5}$$

Полагая

$$p = \frac{1 + r - d}{u - d},\ q = \frac{u - (1 + r)}{u - d}, \tag{6}$$

формуле (5) можно придать следующий вид:

$$C_0 = \frac{1}{1 + r}\left(pC_1(u) + qC_1(d)\right). \tag{7}$$

Пример. Пусть текущая стоимость актива равна 100 рублей. За год его стоимость может повыситься на 20% или понизиться на 15%. Безрисковая годовая ставка равна 10%. Цена исполнения опциона «колл» со сроком исполнения в конце года равна 105 руб. Определим величину премии за опцион.

Если цена актива вырастет и составит 1,2·100 = 120 руб., то владелец опциона воспользуется своим правом на покупку актива за 105 рублей. Следовательно, цена опциона $C_1(u)$ составит 120 –105 = 15 руб.

Если цена актива понизится и составит 0,85·100 = 85 руб., то владелец опциона свое право на покупку актива за 105 рублей использовать не станет, и цена опциона $C_1(d)$ будет равна нулю.

В соответствии с (6) и (7) находим коэффициенты:

$$p = \frac{1{,}1 - 0{,}85}{1{,}2 - 0{,}85} \approx 0{,}714,\ q = \frac{1{,}2 - 1{,}1}{1{,}2 - 0{,}85} \approx 0{,}286$$

и премию за опцион:

$$C_0 = \frac{1}{1{,}1}\left(0{,}714 \cdot 15 + 0{,}286 \cdot 0\right) = 9{,}74 \text{ руб.}$$

По формуле (3) можно найти число опционов, которые нужно продать, чтобы хеджировать риск, связанный с изменением цены актива за год:

$$k = \frac{120 - 85}{15 - 0} \approx 2{,}33\,.$$

Предположим, в начале года инвестор покупает 30 единиц актива A, заплатив 3000 рублей. Одновременно он продает 70 опционов с ценой исполнения 105 рублей. От продажи опционов инвестор получает $9{,}74 \cdot 70 = 681{,}80$ руб., которые размещает под 10%. К концу года сумма, полученная от продажи опционов, вырастет до 749,98 рублей.

Если стоимость актива к концу года вырастет, то, продав купленные им 30 единиц актива, инвестор может получить $120 \cdot 30 = 3600$ руб. При этом ему придется выполнить обязательства по 70 опционам, купив 70 единиц актива по 120 рублей. и продав их владельцам опционов по 105 рублей. На это придется потратить $120 \cdot 70 - 105 \cdot 70 = 1050$ руб.

Доход на конец года составит в этом случае

$$49{,}98 + 3600 - 1050 = 3299{,}98 \text{ руб.}$$

Если стоимость актива к концу года понизится, то, продав купленные им 30 единиц актива, инвестор может получить $85 \cdot 30 = 2550$ руб. Обязательств по опционам не возникнет, так что и в этом случае доход на конец года составит $749{,}98 + 2550 = 3299{,}98$ руб.

9. 4. Двух- и трехпериодные модели

Рассмотрим случай, когда срок исполнения опциона наступает в конце второго периода.

Введем следующие обозначения (рис. 3):

S_0 – цена актива в момент времени $t = 0$;

$S_1(d) = S_0d$, $S_1(u) = S_0u$ – цены актива в момент времени $t = 1$ в случае соответственно понижения или повышения цены;

$S_1(u,u) = S_0u^2$, $S_1(u,d) = S_0ud$, $S_1(d,u) = S_0du$, $S_1(d,d) = S_0d^2$ – цены актива в момент времени $t = 2$ в зависимости от варианта развития событий (повышение оба раза, повышение–понижение и т.д.);

C_0 – цена опциона в момент времени $t = 0$;

X – цена исполнения опциона;

$C_1(d)$, $C_1(u)$ – цены на опцион в момент времени $t = 1$ в случае соответственно понижения или повышения цены;

$C_2(x_1,x_2) = \max\{0,\ S_2(x_1,x_2) - X\}$ – цена опциона в момент времени $t = 2$ при изменении цены актива по сценарию (x_1,x_2) (например, $C_1(u,d) = \max\{0, S_2(u,d) - X\}$).

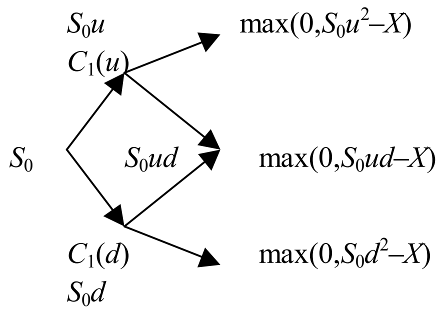


Рис. 3

Фрагменты двухпериодной решетки, представленные на рис. 4, являются однопериодными решетками.

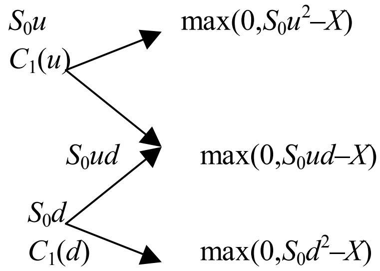


Рис. 4

Применяя к ним формулу (7) для расчета цены опционов в однопериодной модели, получаем:

$$C_1(u) = \frac{1}{1+r}\left(pC_2(u,u) + qC_2(u,d)\right);$$

$$C_1(d) = \frac{1}{1+r}\left(pC_2(d,u) + qC_2(d,d)\right).$$

Применим еще раз те же формулы к однопериодной решетке, представленной на рис. 5, и воспользуемся тем, что $C_2(d,u) = C_2(u,d)$.

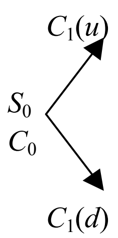


Рис. 5

Получаем:

$$C_0(u) = \frac{1}{1+r}\left(pC_1(u) + qC_1(d)\right) =$$

$$= \frac{1}{1+r}\left(p\frac{1}{1+r}\left(pC_2(u,u) + qC_2(u,d)\right) + \right.$$

$$\left. + \; q\frac{1}{1+r}\left(pC_2(d,u) + qC_2(d,d)\right)\right) =$$

$$= \frac{1}{(1+r)^2}\left(p^2C_2(u,u) + 2pqC_2(u,d) + q^2C_2(d,d)\right). \quad (8)$$

Пример. Пусть, как и в примере из предыдущего пункта, текущая стоимость актива равна 100 руб. Но теперь мы считаем, что изменения стоимости актива происходят дважды в год. За полгода стоимость актива может повыситься на 9,54% или понизиться на 7,80%. Полугодовой повышающий и понижающий коэффициенты равны соответственно u = 1,0954 и d = 0,9220. Заметим, что u^2 = 1,20 и d^2 = 0,85.

Безрисковая годовая ставка равна 10%. Полугодовая ставка определяется из равенства 1 + r = $\sqrt{1{,}1} = 1{,}0488$ и равна 4,88%. Цена исполнения опциона «колл» со сроком исполнения в конце года равна 105 руб. Определим величину премии за опцион.

Сначала вычислим коэффициенты p и q:

$$p = \frac{1{,}0488 - 0{,}9220}{1{,}0954 - 0{,}9220} \approx 0{,}731, \; q = \frac{1{,}0954 - 1{,}0488}{1{,}0954 - 0{,}9220} \approx 0{,}269.$$

Далее:

$$C_2(u,u) = 120 - 105 = 15;$$

$$C_2(u,d) = C_2(d,u) = max\{0,\ 100 - 105\} = 0;$$

$$C_2(d,d) = 0.$$

Таким образом (рис. 6):

$$C_0 = 0{,}731^2 \cdot 15/1{,}1 = 7{,}29 \text{ руб.}$$

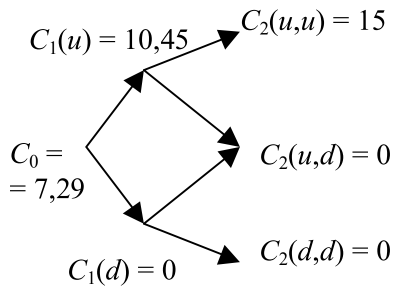


Рис. 6

Пример. Опуская подробности, проведем расчеты для трехпериодной модели. Исходные данные:

$S_0 = 100$ –текущая цена актива;

$u = 1{,}20^{1/3} = 1{,}0627$, $d = 0{,}85^{1/3} = 0{,}9473$ – повышающий и понижающий коэффициенты изменения цены актива за 4 месяца;

$r = 1{,}10^{1/3} = 1{,}0323$; 3,23% – безрисковая ставка за один период;

$X = 105$ – цена исполнения опциона (со сроком исполнения в конце года – через три периода).

Вычислим коэффициенты p и q:

$$p = \frac{1{,}0323 - 0{,}9473}{1{,}0627 - 0{,}9473} \approx 0{,}737, \qquad q = \frac{1{,}0627 - 1{,}0323}{1{,}0627 - 0{,}9473} \approx 0{,}263.$$

Используя очевидные обозначения, последовательно вычисляем:

$C_3(u,u,u) = 120 – 105 = 15;$

$C_3(u,u,d) = C_3(u,d,u) = C_3(d,u,u) = 106{,}27 – 105\} = 1{,}27;$

$C_3(u,d,d) = C_3(d,d,u) = C_3(d,u,d) = max\{0;\ 94{,}73 – 105\} = 0;$

$C_2(u,u) = (15·0{,}737 + 1{,}27·0{,}263)/1{,}0323 = 11{,}03;$

$C_2(u,d) = C_2(d,u) = 1{,}27·0{,}737/1{,}0323 = 0{,}91;$

$C_2(d,d) = 0;$

$C_1(u) = (11{,}03·0{,}737 + 0{,}91·0{,}263)/1{,}0323 = 8{,}11;$

$C_1(d) = 0{,}91·0{,}737 /1{,}0323 = 0{,}65;$

$C_0 = (8{,}11·0{,}737 + 0{,}65·0{,}263)/1{,}0323 = 5{,}96.$

Нетрудно заметить, что в общем виде формула для расчета премии за опцион в трехпериодной модели выглядит следующим образом:

$$C_0 = \frac{p^3 C_3(u,u,u) + 3p^2 q C_3(u,u,d) + 3pq^2 C_3(u,d,d) + q^3 C_3(d,d,d)}{(1+r)^3}. \qquad (9)$$

9. 5. Многопериодная модель

Теорема. *Премия за опцион «колл» может быть вычислена по следующей формуле:*

$$C_0 = \frac{1}{(1+r)^n} \sum_{u^{n-k} d^k S_0 > X} \binom{n}{k} p^{n-k} q^k (u^{n-k} d^k S_0 - X), \qquad (10)$$

где

S_0 – *текущая цена актива* A,

u и *d* – *соответственно коэффициенты повышения и понижения цены актива за один период*;

r – *безрисковая процентная ставка*;

X – *цена исполнения опциона «колл» на актив A*;

n – *число периодов до момента исполнения опциона*.

Доказательство. Сначала заметим, что формула (10) может быть записана следующим образом:

$$C_0 = \frac{1}{(1+r)^n}\sum_{k=0}^{n}\binom{n}{k}p^{n-k}q^k \max(u^{n-k}d^k S_0 - X) =$$

$$= \frac{1}{(1+r)^n}\sum_{k=0}^{n}\binom{n}{k}p^{n-k}q^k C_n(\underbrace{u,\ldots,u}_{n-k \text{ раз}},\underbrace{d,\ldots,d}_{k \text{ раз}}).$$

Теперь воспользуемся методом математической индукции. Как было установлено ранее, формула (10) верна для одного и для двух периодов (формулы (7), (8)), т. е. при $n = 1$ и $n = 2$. Покажем, что если формула (10) верна для n и меньшего числа периодов, то она верна и для $n + 1$ периода.

Итак, предположим, что срок исполнения опциона наступает через $n + 1$ период.

Как было установлено при анализе однопериодной модели, справедливо следующее равенство:

$$C_0 = \frac{1}{1+r}\left(pC_1(u) + qC_1(d)\right). \qquad (11)$$

От момента $t = 1$ до момента $t = n+1$ проходит n периодов. В случае, если за первый период цена актива повысилась, можно воспользоваться формулой (10) для n периодов, считая S_0u текущей ценой актива. Получаем:

$$C_1(u) = \frac{1}{(1+r)^n} \sum_{k=0}^{n} \binom{n}{k} p^{n-k} q^k \max(u^{n-k+1} d^k S_0 - X).$$

Аналогично:

$$C_1(d) = \frac{1}{(1+r)^n} \sum_{k=0}^{n} \binom{n}{k} p^{n-k} q^k \max(u^{n-k} d^{k+1} S_0 - X).$$

После подстановки в (11) находим:

$$C_0 = \frac{1}{(1+r)^{n+1}} \sum_{k=0}^{n} \binom{n}{k} p^{n-k+1} q^k \max(u^{n-k+1} d^k S_0 - X) +$$

$$+ \frac{1}{(1+r)^{n+1}} \sum_{k=0}^{n} \binom{n}{k} p^{n-k} q^{k+1} \max(u^{n-k} d^{k+1} S_0 - X) =$$

$$= \frac{1}{(1+r)^{n+1}} p^{n+1} \max(u^{n+1} S_0 - X) +$$

$$+ \frac{1}{(1+r)^{n+1}} \sum_{k=1}^{n} \binom{n}{k} p^{n-k+1} q^k \max(u^{n-k+1} d^k S_0 - X) +$$

$$+ \frac{1}{(1+r)^{n+1}} \sum_{k=1}^{n} \binom{n}{k-1} p^{n-k+1} q^k \max(u^{n-k+1} d^k S_0 - X) +$$

$$+ \frac{1}{(1+r)^{n+1}} q^{n+1} \max(d^{n+1} S_0 - X).$$

Воспользовавшись тождеством

$$\binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k},$$

приходим к следующему равенству:

$$C_0 = \frac{1}{(1+r)^{n+1}} p^{n+1} \max(u^{n+1} S_0 - X) +$$

$$+ \frac{1}{(1+r)^{n+1}} \sum_{k=1}^{n} \binom{n+1}{k} p^{n-k+1} q^k \max(u^{n-k+1} d^k S_0 - X) +$$

$$+ \frac{1}{(1+r)^{n+1}} q^{n+1} \max(d^{n+1} S_0 - X),$$

откуда и следует, что

$$C_0 = \frac{1}{(1+r)^{n+1}} \sum_{k=0}^{n+1} \binom{n+1}{k} p^{n-k+1} q^k \max(u^{n-k+1} d^k S_0 - X),$$

чем и завершается доказательство теоремы.

Преобразуем формулу (10). Пусть m – наибольшее целое число, для которого $u^{n-m} d^m S_0 > X$. Тогда

$$C_0 = \frac{1}{(1+r)^n} \sum_{k=0}^{m} \binom{n}{k} p^{n-k} q^k (u^{n-k} d^k S_0 - X) =$$

$$= \frac{S_0}{(1+r)^n} \sum_{k=0}^{m} \binom{n}{k} p^{n-k} u^{n-k} q^k d^k - \frac{X}{(1+r)^n} \sum_{k=0}^{m} \binom{n}{k} p^{n-k} q^k .$$

В теории вероятностей устанавливается следующий фундаментальный факт (*формула Лапласа*).

Если p и q – неотрицательные числа такие, что $p + q = 1$, то при больших значениях n имеет место приближенное равенство:

$$\sum_{k=0}^{m}\binom{n}{k}p^{n-k}q^{k} \approx N\left(\frac{m-nq}{\sqrt{npq}}\right), \qquad (12)$$

где

$$N(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-t^2/2}dt \ -$$

так называемая *функция нормального распределения*.

Часто вместо функции $N(x)$ используется функция

$$\Phi(x) = N(x) - 0{,}5 = \frac{1}{\sqrt{2\pi}}\int_{0}^{x} e^{-t^2/2}dt,$$

таблица значений которой приводится, как правило, в любом учебнике по теории вероятностей. Функция $\Phi(x)$ нечетна, монотонно возрастает и $\lim\limits_{x\to\infty}\Phi(x) = 0{,}5$. Значения, близкие к 0,5, функция $\Phi(x)$ принимает уже при сравнительно небольших значениях аргумента. Например, $\Phi(3) \approx 0{,}49865$.

В соответствии с (12) имеем:

$$\frac{X}{(1+r)^n}\sum_{k=0}^{m}\binom{n}{k}p^{n-k}q^{k} \approx \frac{X}{(1+r)^n}\, N\left(\frac{m-nq}{\sqrt{npq}}\right).$$

Заметим, что $pu + qd = 1 + r$. В самом деле:

$$pu + qd = \frac{1+r-d}{u-d}u + \frac{u-(1+r)}{u-d}d$$

$$= \frac{(1+r)u-(1+r)d}{u-d} = 1+r.$$

Положим

$$p' = \frac{pu}{1+r},\ q' = \frac{qd}{1+r}.$$

Так как $p' + q' = 1$, то

$$\frac{S_0}{(1+r)^n}\sum_{k=0}^{m}\binom{n}{k}p^{n-k}u^{n-k}q^k d^k = S_0\sum_{k=0}^{m}\binom{n}{k}p'^{n-k}q'^k \approx$$

$$\approx S_0 N\left(\frac{m-nq'}{\sqrt{np'q'}}\right).$$

Окончательно получаем следующее приближенное равенство:

$$C_0 \approx S_0 N\left(\frac{m-nq'}{\sqrt{np'q'}}\right) - \frac{X}{(1+r)^n}N\left(\frac{m-nq}{\sqrt{npq}}\right). \qquad (13)$$

Модификация (13) приводит к широко применяемой *формуле Блэка–Шоулза*.

Пример. Текущая цена актива составляет 100 руб. За один день цена актива может увеличиться на 0,3% или уменьшиться на 0,3%. Безрисковая ставка равна 10%. Требуется определить премию за опцион «колл» на этот актив со сроком исполнения через 360 дней и ценой исполнения 105 руб. Имеем следующие исходные данные:

$S_0 = 100$; $X = 105$;

$1 + r = 1{,}1^{1/360} = 1{,}00026$;

$u = 1{,}003$; $d = 0{,}997$; $n = 360$.

Найдем наибольшее m, при котором

$$S_0 u^{360-m} d^m > X.$$

Логарифмируя, получаем неравенство

$$(360 - m)\ln u + m \ln d > 1{,}05,$$

откуда

$$m < (360 \ln u - 1{,}05)/(\ln u/d) \approx 171{,}6,$$

так что $m = 171$. Вычисляем:

$$p = \frac{1{,}00026 - 0{,}997}{1{,}003 - 0{,}997} = 0{,}54413; \; q = \frac{1{,}003 - 1{,}00026}{1{,}003 - 0{,}997} = 0{,}45587;$$

$$p' = \frac{0{,}54413 \cdot 1{,}003}{1{,}00026} = 0{,}54562; \; q' = \frac{0{,}45587 \cdot 0{,}997}{1{,}00026} = 0{,}45438;$$

$$\frac{171 - 360q'}{\sqrt{360 p' q'}} = 0{,}78571; \qquad \frac{171 - 360q}{\sqrt{360 pq}} = 0{,}72881.$$

Далее,

$$N(0{,}78571) = 0{,}78398; \qquad N(0{,}72881) = 0{,}766948.$$

Наконец,

$$C_0 = 100 \cdot 0{,}78398 - 105 \cdot 0{,}766948/1{,}1 = 5{,}19 \text{ руб.}$$

10. Биномиальный ряд. Производящие функции

10. 1. Степенные ряды

Хотя название «бином Ньютона» и закрепилось за формулой

$$(x+y)^n = \sum_{k=0}^{n} C_n^k x^{n-k} y^k , \qquad (1)$$

истинная заслуга Ньютона состоит в том, что ему удалось обобщить формулу (1) на случай произвольных показателей степени. Изучением этого обобщения мы займемся в п. 10.2. Ниже излагаются необходимые сведения о степенных рядах.

Формальным степенным рядом (от переменной z) называется выражение вида

$$\sum_{k=0}^{\infty} a_k z^k , \qquad (2)$$

где a_k – члены числовой последовательности, называемые коэффициентами степенного ряда (2). Символ суммирования в (2), вообще говоря, формальный. Тем не менее (2) записывают также в виде

$$a_0 + a_1 z + a_2 z^2 + a_3 z^3 + \ldots .$$

Пусть даны формальные ряды

$$\sum_{k=0}^{\infty} a_k z^k \text{ и } \sum_{k=0}^{\infty} b_k z^k . \qquad (3)$$

Их *суммой* по определению является ряд

$$\sum_{k=0}^{\infty}(a_k + b_k)z^k .$$

Произведением рядов (3) называется ряд

$$\sum_{k=0}^{\infty} c_k z^k ,$$

такой, что

$$c_k = a_0 b_k + a_1 b_{k-1} + \ldots + a_k b_0$$

для всех $k = 0, 1, 2, \ldots$. Заметим, что коэффициенты c_k получаются так, как если бы в произведении рядов были раскрыты скобки и приведены подобные:

$$(a_0 + a_1 z + a_2 z^2 + \ldots)(b_0 + b_1 z + b_2 z^2 + \ldots) =$$
$$= a_0 b_0 + (a_0 b_1 + a_1 b_0)z + (a_0 b_2 + a_1 b_1 + a_2 b_0)z^2 + \ldots$$

Производная ряда (2) определяется как ряд

$$a_1 + 2a_2 z + 3a_3 z^2 + \ldots .$$

Несложно проверить, что так определенные сложение, умножение и дифференцирование рядов обладают привычными свойствами. Сложение и умножение ассоциативны и коммутативны, сложение дистрибутивно относительно умножения. Производная суммы равна сумме производных. Справедливо также соотношение $(uv)' = u'v + uv'$.

Говорят, что ряд (2) *сходится* при $z = t$, где t – некоторое числовое значение, если существует предел

$$\lim_{n\to\infty} \sum_{k=0}^{n} a_k t^k . \qquad (4)$$

Предел (4) в случае, когда он существует, называют *суммой ряда* при $z = t$ и обозначают $\sum_{k=0}^{\infty} a_k t^k$.

Любой степенной ряд сходится при $z = 0$. Суммой ряда (2) в нуле является число a_0. Если ряд сходится при $z = t_1$, то он сходится и при любом $z = t_2$, для которого $|t_2| < |t_1|$. Таким образом, если степенной ряд сходится в некоторой точке, отличной от нуля, то он сходится во всех точках некоторого открытого промежутка вида $(-r, r)$, где $r \in (0; +\infty]$. На промежутке, в точках которого ряд сходится, его сумма задает функцию. Пишут

$$f(z) = \sum_{k=0}^{\infty} a_k z^k,$$

имея в виду, что ряд в правой части сходится в некоторой окрестности нуля, и в каждой точке t из этой окрестности выполняется равенство $f(t) = \sum_{k=0}^{\infty} a_k t^k$.

Пример. Ряд

$$1 + z + z^2 + z^3 + \ldots$$

сходится во всех точках промежутка $(-1, 1)$. В каждой точке $t \in (-1, 1)$ этот ряд представляет собой сумму бесконечно убывающей геометрической прогрессии со знаменателем t. В соответствии с известной формулой

$$1 + z + z^2 + z^3 + \ldots = \frac{1}{1-z}. \qquad (5)$$

Если степенные ряды $\sum_{k=0}^{\infty} a_k z^k$ и $\sum_{k=0}^{\infty} b_k z^k$ сходятся в некоторой общей точке, то в этой точке сходятся их сумма и произведение; если степенной ряд $\sum_{k=0}^{\infty} a_k z^k$ сходится в некоторой точке, отличной от нуля, то в этой точке сходится и его производная. Пусть,

$$f(z) = \sum_{k=0}^{\infty} a_k z^k \, ; \; g(z) = \sum_{k=0}^{\infty} b_k z^k \, .$$

Тогда:

$$\left(\sum_{k=0}^{\infty} a_k z^k \right) + \left(\sum_{k=0}^{\infty} b_k z^k \right) = f(z) + g(z);$$

$$\left(\sum_{k=0}^{\infty} a_k z^k \right) \left(\sum_{k=0}^{\infty} b_k z^k \right) = f(z) g(z);$$

$$\sum_{k=1}^{\infty} k a_k z^{k-1} = f'(z).$$

10. 2. Биномиальный ряд

Начнем с того, что запишем формулу (1) в виде

$$(1+z)^n = \sum_{k \geq 0} \binom{n}{k} z^k \qquad (6)$$

(когда числа C_n^k трактуются как биномиальные коэффициенты, обычно используется обозначение $\binom{n}{k}$).

Заменяя в формуле (5) z на $-z$, получаем:

$$1 - z + z^2 - z^3 + \ldots = \frac{1}{1+z}. \qquad (7)$$

По аналогии с (7) запишем последнее соотношение в виде

$$(1+z)^{-1} = \sum_{k \ge 0} \binom{-1}{k} z^k. \qquad (8)$$

Сравнивая (7) и (8), получаем:

$$\binom{-1}{0} = 1; \ \binom{-1}{1} = -1, \ \ldots .$$

Вообще,

$$\binom{-1}{k} = (-1)^k.$$

Возьмем производную от обеих частей равенства (8):

$$-(1+z)^{-2} = \sum_{k \ge 1} k \binom{-1}{k} z^{k-1}.$$

Полагая

$$\binom{-2}{0} = -\binom{-1}{1}; \ \ldots; \ \binom{-2}{k} = -(k+1) \cdot \binom{-1}{k+1}; \ \ldots ,$$

получаем равенство

$$(1+z)^{-2} = \sum_{k \ge 0} \binom{-2}{k} z^k.$$

Продолжая подобным образом, мы будем получать равенства вида

$$(1+z)^{-n} = \sum_{k \ge 0} \binom{-n}{k} z^k. \qquad (9)$$

Чтобы получить рекуррентные соотношения, связывающие биномиальные коэффициенты, продифференцируем обе части (9):

$$-n(1+z)^{-(n+1)} = \sum_{k\geq 1} k\binom{-n}{k} z^{k-1}.$$

Следовательно,

$$(1+z)^{-(n+1)} = \sum_{k\geq 1} -\frac{k}{n}\binom{-n}{k} z^{k-1} = \sum_{k\geq 0} -\frac{k+1}{n}\binom{-n}{k+1} z^{k}.$$

Сравнивая это равенство с равенством вида (9) для показателя $n + 1$, приходим к следующему соотношению:

$$\binom{-(n+1)}{k} = -\frac{k+1}{n}\binom{-n}{k+1}. \qquad (10)$$

Используя (8) и (10) и применяя метод математической индукции, нетрудно придти к следующему заключению:

$$\binom{-n}{k} = (-1)^k \frac{n(n+1)\cdot\ldots\cdot(n+k-1)}{k!} = (-1)^k C_{n+k-1}^{k}. \qquad (11)$$

Равенство (11) можно записать в следующем виде:

$$\binom{-n}{k} = \frac{-n(-n-1)\cdot\ldots\cdot(-n-k+1)}{k!}. \qquad (12)$$

Формулу для числа сочетаний и (12) можно свести в одну общую формулу:

$$\binom{\text{б}}{k} = \frac{\text{б}(\text{б}-1)\cdot\ldots\cdot(\text{б}-k+1)}{k!}, \qquad (13)$$

где α – положительное или отрицательное целое число.

На самом деле в качестве α можно взять любое действительное число. При этом справедливо соотношение

$$(1+z)^{б} = 1 + \binom{б}{1} z + \binom{б}{2} z^2 + \binom{б}{3} z^3 + \ldots \,. \qquad (14)$$

Чтобы из (14) получить (13) достаточно продифференцировать k раз обе части равенства (14) и подставить z =0:

$$б(б-1)\ldots(б-k+1)(1+z)^{б-k} = k!\binom{б}{k} + (k+1)!\binom{б}{k+1} z + \ldots;$$

$$б(б-1)\ldots(б-k+1) = k!\binom{б}{k}.$$

Пример. Покажем, что

$$\binom{1/2}{k+1} = (-1)^k \binom{2k}{k} \Big/ (k+1)2^{2k+1}. \qquad (15)$$

В соответствии с (13) имеем:

$$\binom{1/2}{k+1} = \frac{1/2\,(1/2-1)\cdot\ldots\cdot(1/2-k)}{(k+1)!} = (-1)^k \, \frac{1\cdot 1\cdot 3\cdot\ldots\cdot(2k-1)}{2^{k+1}(k+1)!}.$$

Умножив числитель и знаменатель последней дроби на

$$2 \cdot 4 \cdot\ldots\cdot 2k = 2^k \cdot k!,$$

получаем доказываемое соотношение:

$$\binom{1/2}{k+1} = (-1)^k \, \frac{(2k)!}{2^{2k+1} k!(k+1)!} = (-1)^k \, \frac{1}{2^{2k+1}(k+1)} \, \frac{(2k)!}{k!k!}.$$

Используя предыдущую формулу, вычислим начальные коэффициенты разложения в ряд функции $(1 + z)^{1/2}$. Получаем:

$$\binom{1/2}{0} = 1;\ \binom{1/2}{1} = \frac{1}{2};\ \binom{1/2}{2} = -\frac{1}{8};\ \binom{1/2}{3} = \frac{1}{16};\ \binom{1/2}{4} = -\frac{5}{128}.$$

Таким образом,

$$\sqrt{1+z} = 1 + \tfrac{1}{2}z - \tfrac{1}{8}z^2 + \tfrac{1}{16}z^3 - \tfrac{5}{128}z^4 + \ldots\ .$$

10. 3. Производящие функции и примеры их применения при решении комбинаторных задач

Если

$$f(z) = \sum_{k=0}^{\infty} a_k z^k\ , \qquad (16)$$

то $f(z)$ называют *производящей функцией* числовой последовательности $a_0, a_1, \ldots, a_k, \ldots$. Если последовательность конечна или, начиная с некоторого места, все ее члены равны нулю, то наряду с (16) пишут также

$$f(z) = \sum_{k=0}^{n} a_k z^k\ ,$$

предполагая, что члены последовательности с номерами больше чем n отсутствуют или равны нулю. Иногда производящей функцией последовательности $a_0, a_1, \ldots, a_k, \ldots$ называют ряд в правой части (16) и говорят, что производящая функция представлена *в замкнутом виде*, если для $f(z)$ указано конечное аналитическое выражение.

Примеры. В приводимой ниже таблице даны некоторые простые последовательности и их производящие функции. В основе большинства примеров лежат частные случаи биномиального ряда: производящей функцией конечной последовательности биномиальных коэффициентов $C_n^0,\ C_n^1,\ \ldots,\ C_n^n$ служит функция $f(z) = (1 + z)^n$; для бесконечной последовательности биномиальных коэффициентов $\binom{\text{б}}{k}$, k = 0, 1, …, производящей функцией служит $f(z) = (1 + z)^{\alpha}$. Обоснование для последних двух примеров дается в курсе анализа.

Последовательность	Производящая функция	Замкнутый вид
1,2,1,0,0,…	$1 + z + z^2$	$1 + z + z^2$
1,1,1,1,…	$\sum z^k$	$\frac{1}{1-z}$
1,–1,1,–1,…	$\sum (-1)^k z^k$	$\frac{1}{1+z}$
1,0,1,0,…	$\sum z^{2k}$	$\frac{1}{1-z^2}$
1,2,4,8,16,…	$\sum 2^k z^k$	$\frac{1}{1-2z}$
1,2,3,4,…	$\sum (k+1) z^k$	$\frac{1}{(1-z)^2}$
$1,1,\frac{1}{2},\frac{1}{3!},\frac{1}{4!}\ldots$	$\sum \frac{z^k}{k!}$	e^z
$0,1,\frac{1}{2},\frac{1}{3},\frac{1}{4}\ldots$	$\sum \frac{z^k}{k}$	$\ln\frac{1}{1-z}$

Рассмотрим несколько примеров комбинаторного применения производящих функций.

Задача о «размене». Требуется найти число способов разменять заданную сумму n рублей монетами достоинством 1, 2 и 5 рублей.

Обозначим искомое число способов размена через P_n, $n = 0, 1, 2, \ldots$. Для малых значений n числа P_n легко вычисляются непосредственно:

$P_0 = 1$; $P_1 = 1$; $P_2 = 2$ (1 + 1, 2); $P_3 = 2$ (1 + 1 + 1, 1 + 2);

$P_4 = 3$ (1 + 1 + 1 + 1, 1 + 1 + 2, 2 + 2).

Составим производящую функцию последовательности P_n:

$$P(z) = P_0 + P_1 z + P_2 z^2 + \ldots .$$

Рассмотрим произведение

$$(1 + z + z^2 + \ldots)(1 + z^2 + z^4 + \ldots)(1 + z^5 + z^{10} + \ldots). \qquad (17)$$

После раскрытия скобок и приведения подобных коэффициент при z^n окажется равным P_n. В самом деле, коэффициент при z^n в (17) – это число способов представить z^n в виде произведения степеней z^u, z^{2v}, z^{5w}, равное числу решений в целых неотрицательных числах уравнения

$$u + 2v + 5w = n.$$

Но это как раз и есть число «разменов» числа n единицами, двойками и пятерками, т. е. число P_n. Следовательно,

$$P(z) = (1 + z + z^2 + \ldots)(1 + z^2 + z^4 + \ldots)(1 + z^5 + z^{10} + \ldots).$$

Суммируя геометрические прогрессии в скобках, получаем:

$$P(z) = \frac{1}{1-z} \cdot \frac{1}{1-z^2} \cdot \frac{1}{1-z^5}. \qquad (18)$$

Используя производную, можно найти явное выражение для P_n:

$$P_n = \frac{P^{(n)}(0)}{n!}.$$

Впрочем, представление (18) позволяет получить рекуррентные соотношения для вычисления чисел P_n. Вместе с

$$P(z) = \frac{1}{1-z} \cdot \frac{1}{1-z^2} \cdot \frac{1}{1-z^5} = P_0 + P_1 z + P_2 z^2 + \ldots$$

пусть

$$Q(z) = \frac{1}{1-z} \cdot \frac{1}{1-z^2} = Q_0 + Q_1 z + Q_2 z^2 + \ldots,$$

$$R(z) = \frac{1}{1-z} = R_0 + R_1 z + R_2 z^2 + \ldots.$$

Тогда:

$$(1 - z)R(z) = 1;$$

$$(1 - z^2)Q(z) = R(z);$$

$$(1 - z^5)P(z) = Q(z).$$

Приравнивая коэффициенты при одинаковых степенях, получаем уравнения:

$$R_0 = 1,\ R_n - R_{n-1} = 0\ (n \ge 1);$$

$$Q_n - Q_{n-2} = R_n\ (n \ge 2);$$

$$P_n - P_{n-5} = Q_n\ (n \ge 5).$$

Отсюда выводятся следующие рекуррентные соотношения:

$$R_0 = 1, \qquad R_n = R_{n-1} \ (n \geq 1);$$

$$Q_n = R_n \ (n < 2), \ Q_n = Q_{n-2} + R_n \ (n \geq 2);$$

$$P_n = Q_n \ (n < 5), \ P_n = P_{n-5} + Q_n \ (n \geq 5).$$

Последовательные вычисления можно свести в таблицу.

n	0	1	2	3	4	5	6	7	8	9	10
R	1	1	1	1	1	1	1	1	1	1	1
Q	1	1	2	2	3	3	4	4	5	5	6
P	1	1	2	2	3	4	5	6	7	8	10

Задача о «счастливых билетах». Автобусный билет с шестизначным номером считается «счастливым», если сумма первых трех цифр равна сумме последних трех цифр. Например, билет 054 342 «счастливый». Требуется найти число счастливых билетов. Считается, что допустимы любые шестизначные номера.

Пусть $x_1x_2x_3\ y_1y_2y_3$ – номер билета. Билет счастливый, если

$$x_1 + x_2 + x_3 = y_1 + y_2 + y_3.$$

Заменив последние три цифры их дополнениями до 9, т. е. полагая $x_4 = 9 - y_1$; $x_5 = 9 - y_2$; $x_6 = 9 - y_3$, предыдущее уравнение можно переписать так:

$$x_1 + x_2 + x_3 + x_4 + x_5 + x_6 = 27. \qquad (19)$$

Задача о счастливых билетах может быть переформулирована теперь следующим образом: требуется найти число целых решений уравнения (19), удовлетворяющих следующим условиям:

$$0 \le x_i < 10,\ i = 1, 2,\ldots, 6.$$

Рассмотрим более общую задачу. Требуется найти число целых решений следующей системы:

$$x_1 + x_2 +\ldots+ x_k = n;\ 0 \le x_i < m,\ i = 1, 2,\ldots, k. \qquad (20)$$

Обозначим число решений (20) через $C_n(k, m)$ (число счастливых билетов в этих обозначениях равно $C_{27}(6, 10)$). Если в выражении

$$(1 + z +\ldots+z^{m-1})^k$$

раскрыть скобки и привести подобные, то коэффициент при z^n окажется равен числу способов представить n в виде суммы k слагаемых, которые могут принимать значения $1, 2,\ldots, m-1$. Но это как раз и есть число $C_n(k, m)$. Таким образом,

$$(1 + z +\ldots+z^{m-1})^k = \textstyle\sum_n C_n(k, m)z^n.$$

Суммируя геометрическую прогрессию в левой части равенства, получаем представление производящей функции чисел $C_n(k, m)$ в свернутом виде:

$$(1 - z^m)^k(1 - z)^{-k} = \textstyle\sum_n C_n(k, m)z^n.$$

По формулам бинома Ньютона имеем:

$$(1-z)^{-k} = \sum_s (-1)^s \binom{-k}{s} z^s; \qquad (1-z^m)^k = \sum_t (-1)^t \binom{k}{t} z^{tm}.$$

После перемножения этих рядов получится ряд с коэффициентами $C_n(k, m)$. Следовательно,

$$C_n(k, m) = \sum_{s+tm=n} (-1)^{s+t} \binom{-k}{s} \binom{k}{t}. \qquad (21)$$

Используя формулу (21), заменяя s на $n - tm$ и воспользовавшись соотношением $C_{k+n-tm-1}^{k-1} = C_{k+n-tm-1}^{n-tm}$, получаем

$$C_n(k, m) = \sum_{s+tm=n} (-1)^{2s+t} C_{k+s-1}^{s} C_k^t = \sum_{tm \le n} (-1)^t C_{k+n-tm-1}^{k-1} C_k^t .$$

В частности, если $m = 10$, $n = 27$, то $tm \le n$ при $t = 0, 1, 2$. Следовательно,

$$C_{27}(6, 10) = C_{32}^5 C_6^0 - C_{22}^5 C_6^1 + C_{12}^5 C_6^2 = 55252.$$

11. Рекуррентные последовательности

11. 1. Рекуррентные соотношения

Последовательность $u_0, u_1, u_2,\ldots$ называют *рекуррентной*, если указана зависимость общего члена последовательности от предыдущих и заданы значения необходимого числа начальных членов. Саму зависимость иногда называют *рекуррентностью*.

Примерами рекуррентных последовательностей могут служить арифметические и геометрические прогрессии.

Члены геометрической прогрессии $u_0, u_1, u_2,\ldots$ со знаменателем q связаны рекуррентным соотношением

$$u_{n+1} = qu_n.$$

Члены произвольной арифметической прогрессии $u_0, u_1, u_2,\ldots$ связаны рекуррентным соотношением

$$u_{n+2} = 2u_{n+1} - u_n.$$

Последовательность факториалов $1, 1, 2, 6, \ldots, n!,\ldots$ определяется рекуррентным соотношением

$$u_{n+1} = (n + 1)u_n$$

(при $u_0 = 1$).

Как рекуррентость может трактоваться формула

$$C_{n+1}^{k} = C_{n}^{k-1} + C_{n}^{k},$$

связывающая биномиальные коэффициенты.

В качестве еще одного примера рассмотрим *задачу о разбиении плоскости прямыми*.

Пусть D_n – число областей, на которые разбивают плоскость n прямых общего положения (таких, что никакие три из них не пересекаются в общей точке и никакие две прямые не параллельны). Ясно, что $D_0 = 1$, $D_2 = 2$. Предположим, что на плоскости уже проведено n прямых, и посмотрим, сколько новых областей добавляется при проведении «новой» $n + 1$-й прямой (рис. 1).

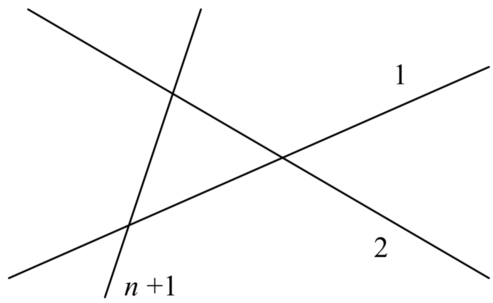


Рис. 1

Каждую область, по которой проходит эта прямая, она рассекает на две, поскольку все области выпуклы. Таким образом, общее число областей увеличится на число областей, через которые проходит $n + 1$-я прямая. Двигаясь по $n + 1$-й прямой в одном направлении, мы пересечем границы областей n раз по числу «старых» прямых. Значит, $n + 1$-я прямая пройдет через $n + 1$ область (в последовательности область – граница – … – область – граница – область число областей на единицу больше, чем число границ).

В результате получаем рекуррентное соотношение

$$D_{n+1} = D_n + (n + 1).$$

Чтобы найти замкнутое выражение для членов последовательности D_n, просуммируем следующие равенства:

$$D_1 = D_0 + 1;$$

$$D_2 = D_1 + 2;$$

$$\dots\dots\dots\dots\dots\dots$$

$$D_n = D_{n-1} + n\ .$$

После сокращений получаем

$$D_n = D_0 + 1 + 2 + \ldots + n\ .$$

Следовательно,

$$D_n = 1 + \frac{n(n+1)}{2}.$$

11. 2. Линейные рекуррентные соотношения

Рекуррентная последовательность $u_0, u_1, u_2,\ldots$ называется *линейной*, если ее члены связаны соотношением вида

$$u_{n+r} = a_1 u_{n+r-1} + a_2 u_{n+r-2} + \ldots + a_r u_n\ , \qquad (1)$$

где a_i, $i = 1, 2,\ldots, r$, – постоянные, не зависящие от n. Соотношение (1) называется *линейным рекуррентным уравнением порядка* r.

Например, члены геометрической прогрессии связаны линейным рекуррентным уравнением первого порядка:

$$u_{n+1} = q u_n.$$

Члены арифметической прогрессии связаны линейным рекуррентным уравнением второго порядка:

$$u_{n+2} = 2u_{n+1} - u_n.$$

Члены последовательности факториалов связаны нелинейным соотношением:

$$u_{n+1} = (n + 1)u_n.$$

Еще один важный пример – рекуррентность, заданная линейным уравнением второго порядка

$$u_{n+2} = u_{n+1} + u_n.$$

Этому уравнению удовлетворяет последовательность *чисел Фибоначчи* F_0, F_1, F_2, F_3, которая получается, если положить $F_0 =0$ и $F_1 =1$:

$$0, 1, 1, 2, 3, 5, 8, 13, 21, \ldots$$

Числа Фибоначчи имеют много важных приложений, в том числе в финансовом анализе. Изучению чисел Фибоначчи посвящена следующая глава.

Последовательность сумм. Пусть $u_0, u_1, u_2,\ldots$ – рекуррентная последовательность, удовлетворяющая соотношению (1). Покажем, что последовательность сумм

$$s_0 = u_0, \qquad s_1 = u_0 + u_1, \qquad \ldots, \quad s_n = u_0 + u_1 +\ldots+ u_n, \ldots$$

также удовлетворяет некоторому линейному рекуррентному соотношению. Заметим сначала, что

$$u_{n+1} = s_{n+1} - s_n , n = 0, 1, \ldots .$$

С учетом этого рекуррентное соотношение, связывающее члены последовательности $u_0, u_1, u_2,\ldots,$ перепишется следующим образом:

$$s_{n+r+1} - s_{n+r} =$$

$$= a_1(s_{n+r} - s_{n+r-1}) + a_2(s_{n+r-1} - s_{n+r-2}) + \ldots+ a_r(s_{n+1} - s_n).$$

Отсюда

$$s_{n+r+1} =$$

$$= (1 + a_1)s_{n+r} + (a_2 - a_1)s_{n+r-1} + \ldots + (a_r - a_{r-1})s_{n+1} - a_r s_n \qquad (2)$$

Таким образом, последовательные суммы связаны рекуррентным соотношением порядка $r + 1$.

Примеры. 1. Для сумм членов геометрической прогрессии имеем:

$$s_{n+2} - s_{n+1} = q(s_{n+1} - s_n),$$

откуда

$$s_{n+2} = (1 + q)s_{n+1} - qs_n.$$

2. Для сумм членов арифметической прогрессии имеем:

$$s_{n+3} - s_{n+2} = 2(s_{n+2} - s_{n+1}) - (s_{n+1} - s_n),$$

и, значит,

$$s_{n+3} = 3s_{n+2} - 3s_{n+1} + s_n.$$

Последовательность степеней натуральных чисел. Начнем с последовательности квадратов:

$$u_0 = 0,\ u_1 = 1,\ u_2 = 4, \ldots,\ u_n = n^2,\ \ldots\ .$$

Члены этой последовательности удовлетворяют следующим соотношениям:

$$u_{n+1} = u_n + 2n + 1;$$

$$u_{n+2} = u_{n+1} + 2n + 3;$$

$$u_{n+3} = u_{n+2} + 2n + 5.$$

Вычитая из второго уравнения первое, а из третьего второе, находим:

$$u_{n+2} - u_{n+1} = u_{n+1} - u_n + 2;$$

$$u_{n+3} - u_{n+2} = u_{n+2} - u_{n+1} + 2.$$

Теперь, вычитая первое из полученных уравнений из второго, получаем рекуррентное соотношение

$$u_{n+3} = 3u_{n+2} - 3u_{n+1} + u_n .$$

Подобно тому как это сделано для последовательности квадратов, несложно поверить, что последовательность кубов

$$u_0 = 0,\ u_1 = 1,\ u_2 = 8, \dots,\ u_n = n^3,\ \dots$$

удовлетворяет соотношению

$$u_{n+4} = 4u_{n+3} - 6u_{n+2} + 4u_{n+1} - u_n.$$

В общем случае последовательность

$$u_0 = 0,\ u_1 = 1,\ u_2 = 2^r, \dots,\ u_n = n^r,\ \dots$$

удовлетворяет соотношению порядка $r + 1$:

$$u_{n+r+1} = \sum_{k=1}^{r+1} (-1)^{k-1} C_{r+1}^{k} u_{n+r+1-k} . \qquad (3)$$

Используя рекуррентные соотношения для степеней натуральных чисел, можно получить рекуррентные соотношения для сумм степеней.

Например, последовательность квадратов натуральных чисел удовлетворяет соотношению

$$u_{n+3} = 3u_{n+2} - 3u_{n+1} + u_n .$$

Следовательно, последовательность сумм квадратов удовлетворяет соотношению

$$s_{n+4} = 4s_{n+3} - 6s_{n+2} + 4s_{n+1} - s_n.$$

11.3. Производящие функции линейных рекуррентных последовательностей

Наша ближайшая цель – определить, какой вид имеет производящая функция произвольной линейной рекуррентной последовательности.

Рассмотрим в качестве примера производящую функцию последовательности чисел Фибоначчи:

$$F(z) = F_0 + F_1 z + F_2 z^2 + \dots \quad . \tag{4}$$

Умножим $F(z)$ на z и на z^2:

$$zF(z) \quad = \quad F_0 z + F_1 z^2 + F_2 z^3 + \dots \ ;$$

$$z^2 F(z) \quad = \quad F_0 z^2 + F_1 z^3 + F_2 z^2 + \dots \ .$$

Так как $F_2 = F_1 + F_0$; $F_3 = F_2 + F_1$;… , то

$$zF(z) + z^2 F(z) = F_0 z + F_2 z^2 + F_3 z^3 + \dots \ .$$

Учитывая, что $F_0 = 0$ и $F_1 = 1$, получаем:

$$z + zF(z) + z^2 F(z) = F(z).$$

Отсюда получается компактная формула:

$$F(z) = \frac{z}{1 - z - z^2}. \tag{5}$$

Используя (5), можно найти явные выражения для чисел Фибоначчи. Начнем с того, что представим дробь из (5) в виде

$$\frac{z}{1 - z - z^2} = \frac{A}{1 - бz} + \frac{B}{1 - вz}. \tag{6}$$

Тогда

$$\frac{A}{1 - бz} = A \sum_{n \ge 0} (бz)^n \ ; \ \frac{B}{1 - вz} = B \sum_{n \ge 0} (вz)^n \ ,$$

и, значит,

$$\frac{z}{1-z-z^2} = A\sum_{n\geq 0}(\text{б}z)^n + B\sum_{n\geq 0}(\text{в}z)^n = \sum_{n\geq 0}(A\text{б}^n + B\text{в}^n)z^n \,.$$

Таким образом,

$$F_n = A\alpha^n + B\beta^n.$$

Чтобы найти неизвестные коэффициенты A, B, α и β, приведем дроби из (6) к общему знаменателю и приравняем коэффициенты при одинаковых степенях z. Получаются следующие уравнения:

$$\alpha + \beta = 1; \qquad \alpha\beta = -1; \quad A + B = 0; \qquad A\beta + B\alpha = -1.$$

Решая полученную систему, находим:

$$\text{б} = \frac{1+\sqrt{5}}{2}\,;\ \text{в} = \frac{1-\sqrt{5}}{2}\,;\ A = \frac{1}{\sqrt{5}}\,;\ B = -\frac{1}{\sqrt{5}}\,.$$

Окончательно получаем:

$$F_n = \frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right). \qquad (7)$$

Пусть теперь

$$u_0, u_1, u_2, \ldots - \qquad (8)$$

произвольная рекуррентная последовательность, удовлетворяющая соотношению

$$u_{n+r} = a_1 u_{n+r-1} + a_2 u_{n+r-2} + \ldots + a_r u_n\,, \qquad (9)$$

и

$$f(z) = u_0 + u_1 z + u_2 z^2 + \ldots -$$

ее производящая функция. Рассмотрим произведение $f(z)$ на многочлен

$$h(z) = 1 - a_1 z - a_2 z^2 - \ldots - a_r z^r .$$

Коэффициент при z^{n+r} равен

$$u_{n+r} - a_1 u_{n+r-1} - a_2 u_{n+r-2} - \ldots - a_r u_n ,$$

и, значит, обращается в ноль в силу (9). Таким образом, в произведении $h(z)f(z)$ все коэффициенты при степенях z, больших или равных r, обращаются в ноль, т. е. $g(z) = h(z)f(z)$ – это *многочлен степени не выше чем* $r - 1$. Производящая функция последовательности (8) оказывается равной дробно-рациональной функции $f(z) = \dfrac{g(z)}{h(z)}$.

Примеры. 1. Для геометрической прогрессии со знаменателем q имеем

$$f(z) = u_0 + u_0 qz + u_0 qz^2 + \ldots; \qquad h(z) = 1 - qz; \qquad h(z)f(z) = u_0 ,$$

откуда получается хорошо известная формула:

$$f(z) = \frac{u_0}{1 - qz}.$$

2. Пусть теперь $u_0, u_1, u_2,\ldots$ – арифметическая прогрессия. Так как $u_{n+2} = 2u_{n+1} - u_n$, то

$$h(z) = 1 - 2z + z^2$$

и

$$h(z)f(z) = u_0 + (u_1 - 2u_0)z^2.$$

Следовательно,

$$f(z) = \frac{u_0 + (u_1 - 2u_0)z}{(1-z)^2}. \qquad (10)$$

Рассмотрим, в частности, последовательность положительных целых чисел 1, 2, 3, … . Запишем ее производящую функцию:

$$f(z) = 1 + 2z + 3z^2 + \ldots . \qquad (11)$$

По формуле (10) получаем:

$$f(z) = \frac{1}{(1-z)^2}. \qquad (12)$$

К тому же результату можно придти по-другому. Ряд (11) получается как производная ряда

$$z + z^2 + z^3 + \ldots ,$$

сумма которого (как сумма геометрической прогрессии) равна

$$\frac{z}{1-z}. \qquad (13)$$

Дифференцируя (13), получаем (12).

Наряду с многочленом $h(z)$ рассмотрим многочлен

$$k(z) = z^r - a_1 z^{r-1} - a_2 z^{r-2} - \ldots - a_r,$$

называемый *характеристическим* (для последовательности, члены которой удовлетворяют рекуррентному соотношению (1)).

Многочлены $h(z)$ и $k(z)$ связаны простым соотношением:

$$h(z) = z^r k(1/z).$$

Пусть $\alpha_1, \alpha_2, \ldots, \alpha_s$ – корни характеристического многочлена (возможно, комплексные) с кратностями $p_1, p_2, \ldots, p_s$. Тогда

$$k(z)=(z-\text{б}_1)^{p_1}(z-\text{б}_2)^{p_2}\ldots(z-\text{б}_s)^{p_s},$$

и, соответственно,

$$h(z)=(1-\text{б}_1 z)^{p_1}(1-\text{б}_2 z)^{p_2}\ldots(1-\text{б}_s z)^{p_s}.$$

Рациональная функция $f(z)=\dfrac{g(z)}{h(z)}$ может быть представлена в виде суммы простых дробей вида

$$\frac{\text{в}}{(1-\text{б}z)^p}, \tag{14}$$

где $\alpha=\alpha_i$ – один из корней характеристического многочлена, p лежит в промежутке от 1 до p_i, а β – некоторое число. Разлагая (14) по формуле бинома, получаем:

$$\frac{\text{в}}{(1-\text{б}z)^p}=\sum_{n\geq 0}\text{в}\binom{-p}{n}(-1)^n\text{б}^n z^n.$$

Относительно n выражение

$$\binom{-p}{n}(-1)^n=(-1)^n C_{p+n-1}^n\cdot(-1)^n=C_{p+n-1}^n=C_{p+n-1}^{p-1}$$

является многочленом степени не выше $p-1$ и, тем более, не выше p_i-1. Следовательно, сумма всех дробей вида (14), относящихся к одному и тому же корню $\alpha=\alpha_i$, может быть представлена в виде $\sum\limits_{n\geq 0}P_i(n)\text{б}_i^n z^n$, причем степень многочлена $P_i(n)$ не превышает p_i-1. Суммируя по всем $i=1, 2, \ldots, s$, получаем:

$$f(z)=\sum_{n\geq 0}\left(\sum_{i=1}^{s}P_i(n)\alpha_i^n\right)z^n.$$

Сравнивая полученное выражение с

$$f(z) = u_0 + u_1 z + u_2 z^2 + \ldots ,$$

приходим к заключению, что

$$u_n = \sum_{i=1}^{s} P_i(n) б_i^n . \qquad (15)$$

Пример. Рассмотрим последовательность 6, 0, 12, … , удовлетворяющую рекуррентному соотношению

$$u_{n+3} = 3u_{n+1} + 2u_n.$$

Многочлен

$$k(z) = z^3 - 3z - 2$$

является характеристическим для этой последовательности. Раскладывая его на множители получаем:

$$k(z) = (z + 1)^2(z - 2).$$

Следовательно, члены рассматриваемой последовательности имеют вид

$$u_n = P(n)\cdot(-1)^n + Q\cdot 2^n ,$$

где $P(n)$ – многочлен степени не выше первой, а Q – степени не выше нулевой (т. е. ноль или некоторое число).

Будем искать u_n в виде

$$u_n = (An + B)\cdot(-1)^n + Q\cdot 2^n .$$

Поскольку $u_0 = 6$, $u_1 = 0$, $u_2 = 12$, получаем следующие уравнения относительно A, B и Q:

$$B + Q = 6; \; -(A + B) + 2Q = 0; \quad 2A + B + 4Q = 12.$$

Решая систему, находим:

$$A = 0;\ B = 4;\ Q = 2.$$

Следовательно,

$$u_n = (-1)^n \cdot 4 + 2^{n+1}.$$

Пример. Используя (15), найдем формулу для суммы квадратов первых n натуральных чисел.

Суммы квадратов связаны рекуррентным соотношением

$$s_{n+4} = 4s_{n+3} - 6s_{n+2} + 4s_{n+1} - s_n.$$

Характеристический многочлен последовательности сумм квадратов

$$k(z) = z^4 - 4z^3 + 6z^2 - 4z + 1 = (z - 1)^4$$

имеет единственный корень 1 кратности четыре. Следовательно, s_n представимо в виде

$$s_n = An^3 + Bn^2 + Cn + D.$$

Так как

$$s_0 = 0,\ s_1 = 1,\ s_2 = 5,\ s_3 = 14,$$

то

$$D = 0; \qquad A + B + C = 1;$$

$$8A + 4B + 2C = 5; \qquad 27A + 9B + 3C = 14.$$

Вычисляем определитель системы уравнений, содержащих неизвестные A, B, C, и определители, соответствующие переменным:

$$\Delta = \begin{pmatrix} 1 & 1 & 1 \\ 8 & 4 & 2 \\ 27 & 9 & 3 \end{pmatrix} = -12;\ \Delta_A = \begin{pmatrix} 1 & 1 & 1 \\ 5 & 4 & 2 \\ 14 & 9 & 3 \end{pmatrix} = -4;$$

$$\Delta_B = \begin{pmatrix} 1 & 1 & 1 \\ 8 & 5 & 2 \\ 27 & 14 & 3 \end{pmatrix} = -6;\ \Delta_C = \begin{pmatrix} 1 & 1 & 1 \\ 8 & 4 & 5 \\ 27 & 9 & 14 \end{pmatrix} = -2.$$

Находим: $A = 1/3$; $B = 1/2$; $C = 1/6$. Следовательно,

$$s_n = \tfrac{1}{3}n^3 + \tfrac{1}{2}n^2 + \tfrac{1}{6}n = \frac{n(n+1)(2n+1)}{6}.$$

11. 4. Числа Каталана. Случайное блуждание

В этом параграфе мы исследуем рекуррентное соотношение, которое встречается во многих задачах (см., например, далее подсчет числа бинарных деревьев), в том числе в задачах финансового анализа.

Рассмотрим биномиальную решетку цен актива. Пусть, как и ранее, $u > 1$ – повышающий коэффициент, а $d < 1$ – понижающий. Если стоимость актива повышалась p раз и понижалась q раз, то она станет равной $S_0u^pd^q$.

Сопоставим каждому пути на решетке, последовательность вида

$$(a_1, a_2, \ldots, a_m), \tag{16}$$

состоящую из –1 и 1, в которой –1 соответствует понижению стоимости, а 1 – повышению. Пусть $a = a_1 + a_2 + \ldots + a_m$. Число a показывает на сколько число повышений больше числа понижений. Число «минус единиц» q в последовательности (16) равно $(n - a)/2$, число «единиц» p равно $(n + a)/2$. Стоимость актива с последовательностью повышений-понижений (16) к моменту m равна

$$S_0u^pd^q = S_0u^{(n+a)/2}d^{(n-a)/2}.$$

Предположим, что $p + q = 2n$ – четное число. Наибольшее число путей C_{2n}^n ведет в узел $S_0u^nd^n$. На этих путях число понижений равно числу повышений. Если считать, что $ud = 1$, то стоимость актива в результате развития событий по пути, приводящему в узел $S_0u^nd^n$, остается неизменной: $S_0u^nd^n = S_0$. Мы займемся подсчетом числа путей, на которых стоимость актива ни разу не опускается ниже S_0. Иными словами, нас интересует число последовательностей

$$(a_1, a_2, \ldots, a_{2n}), \qquad (17)$$

для которых сумма всех элементов равна нулю:

$$a_1 + a_2 + \ldots + a_{2n} = 0,$$

а все частичные суммы неотрицательны:

$$s_1 = a_1 \geq 0, \quad s_2 = a_1 + a_2 \geq 0, \quad \ldots ,$$

$$s_{2n-1} = a_1 + a_2 + \ldots + a_{2n-1} \geq 0.$$

Про такие последовательности будем говорить, что они *сбалансированы*. Графически последовательность вида (17) может быть представлена диаграммой частичных сумм. На рис. 2 представлены две сбалансированные последовательности длины 4 (при $n = 2$); на рис. 3 – пять сбалансированных последовательностей длины 6 (при $n = 3$). Сбалансированность означает, что диаграмма не опускается ниже нулевого уровня.

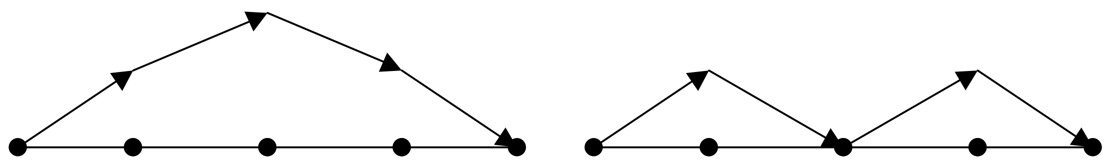


Рис. 2

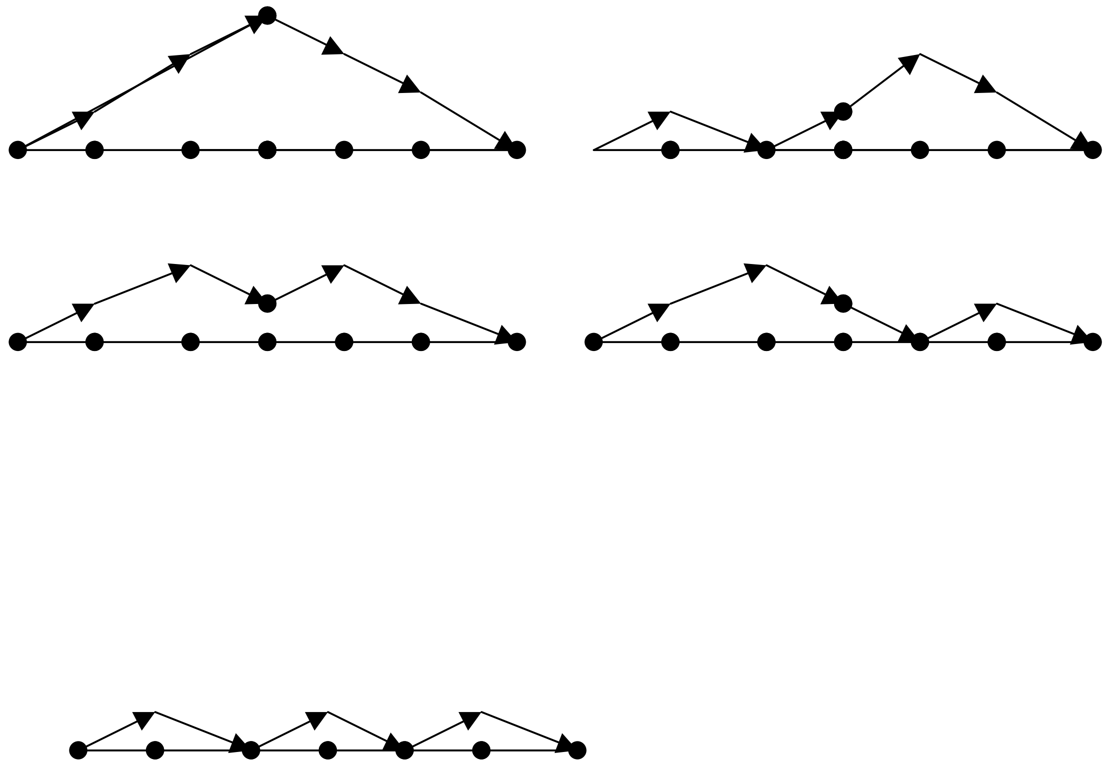


Рис. 3

Обозначим через C_n – число различных сбалансированных последовательностей длины $2n$. Ясно, что $C_0 = 1$ (имеется ровно одна пустая последовательность) и $C_1 = 1$. Мы видели также, что $C_2 = 2$ и $C_3 = 5$.

При $n > 0$ всякая сбалансированная последовательность начинается с единицы и заканчивается минус единицей:

$$a_1 = 1,\ a_{2n} = -1.$$

Если все частичные суммы сбалансированной последовательности (17) положительны, то последовательность (a_2, …, a_{2n-1}), имеющая длину $2n - 2$, также оказывается сбалансированной. Обратно, любую сбалансированную последовательность длины $2n - 2$ можно дополнить до сбалансированной последовательности длины $2n$, приписав к ней слева единицу, а справа – минус единицу. Тем самым между множеством сбалансированных последовательностей длины $2n$ с положительными частичными суммами и множеством всех сбалансированных последовательностей длины $2n - 2$ устанавливается взаимно однозначное соответствие. Следовательно, число сбалансированных последовательностей длины $2n$ с положительными частичными суммами равно C_{n-1} – общему числу сбалансированных последовательностей длины $2n - 2$.

Рассмотрим теперь произвольную сбалансированную последовательность (17), имеющую нулевые частичные суммы. Пусть $2k < 2n$ – номер последней частичной суммы, равной нулю. Последовательность (17) можно разбить на две сбалансированные последовательности

$$(a_1, a_2, \ldots, a_{2k}),\ (a_{2k+1}, a_2, \ldots, a_{2n}), \qquad (18)$$

длины которых суть $2k$ и $2n$, при этом частичные суммы второй последовательности положительны. Обратно, имея две сбалансированные последовательности (18) длин $2k$ и $2n$, можно составить из них сбалансированную последовательность (17) длины n. Если вторая из последовательностей (18) имеет положительные частичные суммы, то последней нулевой частичной суммой последовательности (17) окажется частичная сумма с номером $2k$. По правилу произведения число последовательностей (17), для которых последняя нулевая частичная сумма имеет номер $2k$, составляет $C_k \cdot C_{n-k-1}$.

Таким образом, суммируя по всем $k < n$ и учитывая, что $C_0 = 1$, приходим к равенству

$$C_n = C_0C_{n-1} + C_1C_{n-2} + \ldots + C_{n-1}C_0. \qquad (19)$$

В частности,

$$C_1 = C_0^2 = 1;\ C_2 = C_0C_1 + C_1C_0 = 2;$$

$$C_3 = C_0C_2 + C_1C_1 + C_2C_0 = 5;\ \ldots .$$

Составим производящую функцию:

$$C(z) = C_0 + C_1z + C_2z^2 + \ldots .$$

Используя (19), получаем:

$$C(z)^2 = C_0^2 + (C_0C_1 + C_1C_0)z + (C_0C_2 + C_1C_1 + C_2C_0)z^2\ \ldots =$$
$$= C_1 + C_2z + C_3z^2 + \ldots .$$

Умножая обе части последнего равенства на z и добавляя 1, приходим к соотношению

$$z \cdot C(z)^2 + 1 = C(z).$$

Таким образом, $y = C(z)$ удовлетворяет квадратному уравнению

$$z \cdot y^2 - y + 1 = 0.$$

Решая его, находим:

$$y = \frac{1}{2z}\left(1 \pm \sqrt{1-4z}\right)$$

Для функции $y = C(z)$ должно выполняться соотношение $\lim\limits_{z\to 0} C(z) = 1$. Для $y = \frac{1}{2z}\left(1 + \sqrt{1-4z}\right)$ предел при $z\to 0$ не существует. Следовательно,

$$C(z) = \frac{1}{2z}\left(1 - \sqrt{1-4z}\right).$$

Воспользовавшись биномом Ньютона, находим:

$$C(z) = \frac{1}{2z}\left(1 - \sum_{k\geq 0}\binom{1/2}{k}(-4z)^k\right) = \sum_{k\geq 1}(-1)^{k-1}\binom{1/2}{k}\cdot 2^{2k-1}z^{k-1},$$

т.е.

$$C(z) = \sum_{n\geq 0}(-1)^k\binom{1/2}{n+1}\cdot 2^{2n+1}z^n.$$

Таким образом,

$$C_n = (-1)^k\binom{1/2}{n+1}\cdot 2^{2n+1}.$$

Несложно проверить, что

$$(-1)^k\binom{1/2}{n+1}\cdot 2^{2n+1} = \frac{1}{n+1}\binom{2n}{n}.$$

Окончательно получаем:

$$C_n = \frac{1}{n+1}\binom{2n}{n}. \qquad (20)$$

Числа C_n называются *числами Каталана*. Числа Каталана часто встречаются при решении задач перечисления (см. далее перечисление бинарных деревьев).

Формула (20) позволяет установить и некоторые другие факты о числе путей.

Для начала напомним, что при выводе (20) мы пришли к заключению, что число сбалансированных последовательностей длины $2n$ с положительными частичными суммами равно C_{n-1}.

Любая последовательность (17) с положительными частичными суммами после умножения на –1 превращается в последовательность с отрицательными частичными суммами и обратно. Следовательно, число последовательностей (17) с отрицательными частичными суммами и нулевой суммой s_{2n} столько же, сколько и сбалансированных последовательностей с положительными частичными суммами, т. е. C_{n-1}. Таким образом, число последовательностей (17), для которых $s_{2n} = 0$ и ни одна частичная сумма не обращается в ноль, равно $2C_{n-1}$.

Если пути на биномиальной решетке соответствует последовательность (17) такая, что $s_{2n} = 0$, а все частичные суммы отличны от нуля, говорят , что *первое возвращение в ноль* происходит в момент $2n$.

Сумма

$$\sum_{k=1}^{t} 2C_{k-1} = 2C_0 + 2C_1 + \ldots + 2C_{t-1}$$

дает число путей, на которых первое возвращение в ноль происходит в один из моментов 2, 4, …, $2t$. Соответственно

$$2^{2t} - \sum_{k=1}^{t} 2C_{k-1} -$$

это число путей длины $2t$, для которых возвращения в ноль до момента $2t$ (включая его) не происходит. Из них половина путей расположена выше нулевого уровня, а половина – ниже. Следовательно, количество путей, для которых число повышений превосходит число понижений в любой момент времени от 1 до $2t$, равно

$$2^{2t-1} - \sum_{k=1}^{t} C_{k-1}.$$

Анализ (как правило, довольно сложный) подобного рода величин позволяет оценить вероятность того или иного варианта развития событий, связанных с изменением цены актива.

12. Числа Фибоначчи

12. 1. Простейшие свойства

Напомним, что числа Фибоначчи образуют последовательность, в которой первые два члена равны 0 и 1, а каждый следующий равен сумме двух предыдущих. В соответствии с определением последовательность чисел Фибоначчи

$$F_0 = 0, F_1 = 1, F_2 = 1, F_3 = 2, F_4 = 3, F_5 = 5, F_6 = 8,\ldots$$

удовлетворяет рекуррентному соотношению

$$F_{n+2} = F_{n+1} + F_n. \qquad (1)$$

Числа Фибоначчи возникают естественным образом во многих задачах. Исторически одной из первых является *задача о кроликах*, восходящая к Леонардо Пизанскому, которого иногда называют Леонардо Фибоначчи (публикация 1202 г.). В этой задаче требуется определить число пар зрелых кроликов, образовавшихся от одной пары в течение года, если известно, что каждая зрелая пара кроликов ежемесячно рождает новую пару, причем новорожденные кролики достигают зрелости через два месяца.

Обозначим через u_n, u_{n+1}, u_{n+2} число пар зрелых кроликов соответственно через n месяцев, через $n + 1$ месяц и через $n + 2$ месяца. Нетрудно убедиться, что $u_{n+2} = u_{n+1} + u_n$: к моменту $n + 2$ зрелости достигают u_n пар кроликов, родившихся в момент n, которые добавляются к u_{n+1} паре кроликов, зрелых на момент $n + 1$.

Рассмотрим еще одну задачу, так называемую *задачу о прыгуне*. Некоторая величина x увеличивается за единицу времени на 1 или на 2. Требуется определить, сколькими способами может произойти увеличение рассматриваемой величины на n единиц. Обозначим искомое число способов через u_n. Пусть $x(t)$ – значение величины x в момент времени t. Ясно, что $x(1) = x(0) + 1$ или $x(1) = x(0) + 2$ в зависимости от того, на 1 или на 2 изменилась величина x за первый временной промежуток. Имеются два пути достичь значения $x(0) + n$: вырасти на $n - 1$ единицу со значения $x(0) + 1$ или на $n - 2$ единицы со значения $x(0) + 2$. Первое может произойти u_{n-1} способами, второе – u_{n-2} способами. Следовательно,

$$u_n = u_{n-1} + u_{n-2}.$$

Многие свойства чисел Фибоначчи нетрудно получить по индукции. Например, для любого $n \geq 0$ справедливо равенство

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^{n+1} = \begin{pmatrix} F_{n+2} & F_{n+1} \\ F_{n+1} & F_n \end{pmatrix}. \tag{2}$$

В самом деле, (2) выполняется при $n = 0$. Следующая выкладка позволяет сделать индуктивный шаг:

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^{n+2} = \begin{pmatrix} F_{n+2} & F_{n+1} \\ F_{n+1} & F_n \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} =$$

$$= \begin{pmatrix} F_{n+2} + F_{n+1} & F_{n+2} \\ F_{n+1} + F_n & F_{n+1} \end{pmatrix} = \begin{pmatrix} F_{n+3} & F_{n+2} \\ F_{n+2} & F_{n+1} \end{pmatrix}.$$

Если взять определители матриц, стоящих в правой и левой частях (2), получится следующее соотношение:

$$F_{n+2}F_n - F_{n+1}^2 = (-1)^{n+1}.$$

Приведем одно из важных свойств чисел Фибоначчи, доказательство которого значительно сложнее и здесь не приводится.

Каждое целое положительное число имеет единственное представление вида

$$n = F_{k_1} + F_{k_2} + \ldots + F_{k_r}, \tag{3}$$

где $k_1 \geq k_2 + 2$; $k_2 \geq k_3 + 2$; $\ldots$, $k_{r-1} \geq k_r + 2$; $k_r \geq 2$.

Чтобы получить представление (3) нужно в качестве F_{k_1} взять наибольшее число Фибоначчи, не превосходящее n, в качестве F_{k_2} – наибольшее число Фибоначчи, не превосходящее $n - F_{k_1}$, и т. д., пока очередной «остаток» не станет равным нулю. Например:

$1 = F_1$; $2 = F_2$; $3 = F_3$;$4 = F_3 + F_1$; $5 = F_4$; $8 = F_5$; $13 = F_6$;

$20 = F_6 + F_4 + F_2$.

12.2. Формула Бине и некоторые ее применения

Напомним (см. 11.3, формула (5)), что производящая функция для последовательности чисел Фибоначчи имеет вид

$$F(z) = F_0 + F_1 z + F_2 z^2 + \ldots = \frac{z}{1 - z - z^2}.$$

Корнями характеристического многочлена

$$k(z) = z^2 - z - 1$$

являются числа

$$б = \frac{1+\sqrt{5}}{2} \text{ и } в = \frac{1-\sqrt{5}}{2}.$$

Число Фибоначчи как функция своего номера представляется в виде:

$$F_n = \frac{б^n - в^n}{\sqrt{5}}, \tag{4}$$

или

$$F_n = \frac{\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n}{\sqrt{5}},$$

Формула (4) называется *формулой Бине*.

Так как $\alpha^2 = \alpha + 1$, то любую степень числа α можно представить в виде целочисленной комбинации $a\alpha + b$. Оказывается, коэффициентами служат числа Фибоначчи:

$$\alpha^{k+2} = F_{k+2}\alpha + F_{k+1}.$$

Формула Бине позволяет убедиться в этом прямой проверкой.

Приведем несколько оценок чисел Фибоначчи.

*Число Фибоначчи F_n есть ближайшее целое к чис*лу $\frac{б^n}{\sqrt{5}}$.

Для доказательства заметим, что $|\beta| < 1$, и, значит, $|\beta^n| < 1$. Следовательно,

$$\left| F_n - \frac{\text{б}^n}{\sqrt{5}} \right| = \left| \frac{\text{б}^n - \text{в}^n}{\sqrt{5}} - \frac{\text{б}^n}{\sqrt{5}} \right| = \frac{|\text{в}|^n}{\sqrt{5}} < \frac{1}{2}. \qquad (5)$$

Из (4) вытекает также следующее свойство.

С ростом n числа Фибоначчи неограниченно сближаются с членами геометрической прогрессии с начальным членом $\frac{1}{\sqrt{5}}$ *и знаменателем* $\alpha = \frac{\sqrt{5}+1}{2}$:

$$\lim_{n\to\infty}\left(F_n - \frac{\text{б}^n}{\sqrt{5}} \right) = 0.$$

*Отношение соседних чисел Фибоначчи (следующего к предыдущему) с ростом n стремится к чис*лу $\frac{2}{\sqrt{5}+1} \approx 0{,}618$.

Действительно,

$$\frac{F_n}{F_{n+1}} = \frac{\text{б}^n - \text{в}^n}{\text{б}^{n+1} - \text{в}^{n+1}} = \frac{1 - (\text{в}/\text{б})^n}{\text{б} - \text{в}(\text{в}/\text{б})^n}.$$

Так как $\beta/\alpha < 1$, то $(\beta/\alpha)^n$ с ростом n стремится к нулю. Следовательно,

$$\lim_{n\to\infty} \frac{F_n}{F_{n+1}} = \frac{1}{\text{б}} \approx 0{,}618.$$

Из предыдущего равенства следует, что

$$\lim_{n\to\infty}\frac{F_n}{F_{n+2}}=\frac{1}{б^2}\approx 0{,}382\,.$$

Установим справедливость формулы, которая дает представление чисел Фибоначчи в виде суммы биномиальных коэффициентов, и как ее следствие получим одно тождество для биномиальных коэффициентов.

При любом n справедливо следующее равенство:

$$F_{n+1}=\sum_{2k\le n}C_{n-k}^{k}\,. \qquad (6)$$

Заметим сначала, что при четном *n* равенство (6) имеет вид

$$F_{n+1}=C_n^0+C_{n-1}^1+\ldots+C_{n/2}^{n/2},$$

а при нечетном –

$$F_{n+1}=C_n^0+C_{n-1}^1+\ldots+C_{(n+1)/2}^{(n-1)/2}\,.$$

Например:

$$F_5=C_4^0+C_3^1+C_2^2\,;\ F_6=C_5^0+C_4^1+C_3^2\,.$$

Для доказательства (6) воспользуемся производящей функцией последовательности чисел Фибоначчи:

$$F(z)=F_0+F_1z+F_2z^2+\ldots=\frac{z}{1-z-z^2}\,. \qquad (7)$$

Используя формулу для суммы членов бесконечно убывающей геометрической прогрессии, преобразуем правую часть равенства (7):

$$\frac{z}{1-z-z^2} = z\cdot\frac{1}{1-(z+z^2)} = z(1 + (z + z^2) + (z + z^2)^2 + \ldots) =$$

$$= z + z^2(1 + z) + z^3(1 + z)^2 + \ldots . \qquad (8)$$

Коэффициент при z^{n+1}, получающийся после раскрытия скобок и приведения подобных в (8), должен в соответствии с (7) равняться F_{n+1}. Слагаемые в (8), следующие за $z^{n+1}(1 + z)^n$, содержат z в степенях более высоких, чем $n + 1$. Слагаемые $z^{k+1}(1 + z)^k$ при $2k + 1 < n + 1$ содержат z в степенях более низких, чем $n + 1$. Так как

$$z^{p+1}(1 + z)^p = z^{p+1}\sum_{k\le p} C_p^k z^k = \sum_{k\le p} C_p^k z^{k+p+1},$$

то коэффициент при z^{n+1} получается суммированием по p всех таких коэффициентов C_p^k, что $k+p+1 = n + 1$ и $k \le p$. Последние условия означают, что $p = n - k$ и $2k \le n$. Таким образом, мы суммируем коэффициенты C_{n-k}^k, $2k \le n$, и получаем сумму, стоящую в правой части (6), что и доказывает это равенство.

С учетом формулы Бине равенство (7) может быть записано следующим образом:

$$\sum_{2k\le n} C_{n-k}^k = \frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^{n+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{n+1}\right).$$

12. 3. Золотое сечение

Разделим отрезок AB на части так, чтобы большая часть была бы средним пропорциональным между всем отрезком и меньшей его частью. Иными словами, мы ищем точку C такую, что $AC:BC = BC:AB$ (рис. 1).

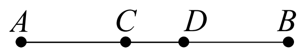


Рис. 1

Пусть $BC = x \cdot AB$. Тогда $AC = x \cdot BC = x^2 \cdot AB$. Так как $AC + BC = AB$, то

$$x^2 + x = 1. \tag{9}$$

Уравнение (9) имеет два решения:

$$\frac{-1+\sqrt{5}}{2} \approx 0{,}618 \qquad \text{и} \qquad \frac{-1-\sqrt{5}}{2} \approx -1{,}618.$$

Точка C соответствует положительному решению

$$x = \frac{-1+\sqrt{5}}{2}.$$

Разбиение отрезка AB на две части точкой C называют *золотым сечением* этого отрезка.

Точно так же можно получить золотое сечение отрезка AB точкой D такой, что $BD:AD = AD:AB$.

Оказывается, точка D дает золотое сечение отрезка BC.

Действительно:

$$BC = x \cdot AB,\ AD = x \cdot AB,\ BD = (1 - x) \cdot AB,$$

и, значит, $CD = BC - BD = (2x - 1) \cdot AB$. Таким образом,

$$\frac{CD}{BD} = \frac{2x-1}{1-x};\ \frac{BD}{BC} = \frac{1-x}{x}.$$

Так как x служит корнем уравнения $x^2 + x - 1 = 0$, то, как легко проверить, $\frac{2x-1}{1-x} = \frac{1-x}{x}$. Следовательно,

$$\frac{CD}{BD} = \frac{BD}{BC}.$$

Заметим, что $x = \frac{-1+\sqrt{5}}{2}$ – это предел, к которому стремится отношение соседних чисел Фибоначчи F_n/F_{n+1}. Приближение золотого сечения, изображенного на рис. 1, можно получить, взяв точки C' и D' такие, что

$$AC' = \left(1 - \frac{F_{n+1}}{F_{n+2}}\right)AB = \frac{F_n}{F_{n+2}}AB;\ AD' = \frac{F_{n+1}}{F_{n+2}}AB.$$

К «золотому» пределу $\frac{-1+\sqrt{5}}{2} = \frac{1}{\text{б}}$, где $\text{б} = \frac{1+\sqrt{5}}{2}$, отношение u_n/u_{n+1} стремится для любой последовательности с положительными членами, удовлетворяющей рекуррентному соотношению $u_{n+2} = u_{n+1} + u_n$. В самом деле, члены любой такой последовательности имеют вид

$$u_n = a\alpha^n + b\beta^n,$$

где $\text{б} = \frac{1+\sqrt{5}}{2}$ и $\text{в} = \frac{1-\sqrt{5}}{2}$ – корни характеристического уравнения $x^2 - x - 1 = 0$. Поскольку все члены последовательности (u_n) положительны, $a \neq 0$. Так как $|\beta/\alpha| < 1$, то $(\beta/\alpha)^n \to 0$ при $n \to \infty$. Следовательно,

$$\frac{u_n}{u_{n+1}} = \frac{a\text{б}^n + b\text{в}^n}{a\text{б}^{n+1} + b\text{в}^{n+1}} = \frac{1 + \frac{b}{a}\left(\frac{\text{в}}{\text{б}}\right)^n}{\text{б} + \frac{b}{a}\left(\frac{\text{в}}{\text{б}}\right)^n \text{в}} \to \frac{1}{\text{б}} \text{ при } n \to \infty.$$

Таким образом, аппроксимировать золотое сечение можно с помощью любой положительной последовательности (u_n), члены которой удовлетворяют рекуррентному соотношению $u_{n+2} = u_{n+1} + u_n$.

Золотое сечение используется в техническом анализе при торговли ценными бумагами на инвестиционных рынках.

Центральную роль в так называемом анализе Фибоначчи играют *Фиб-узлы* (или *уровни Ди Наполи*). На ценовой волне берутся две крайние точки A и B и отрезок AB делится точками C и D (рис. 2) подобно тому, как это сделано на рис. 1. Полученные точки называются Фиб-узлами. Считается, что в этих точках происходит значительное сопротивление изменению цен при обратном движении.

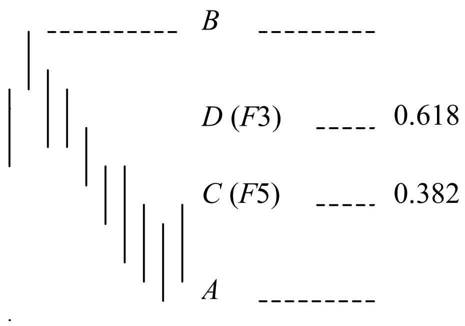


Рис. 2

В качестве Фиб-узлов используются также близкие к ним точки – $F3$ и $F5$ – Фиб-узлы, расположенные на уровнях 3/8 и 5/8:

$$F3 = A + \frac{F_4}{F_6}(B - A);\ F5 = A + \frac{F_5}{F_6}(B - A).$$

Некоторое теоретическое объяснение эмпирически обоснованному выбору Фиб-узлов дается в следующем параграфе.

12. 4. Числа Фибоначчи и поиск экстремума

Рассмотрим следующую задачу. Имеются две величины x и y, связанные функциональной зависимостью. Известно, что когда x пробегает промежуток $[a, b]$, величина y сначала убывает, затем возрастает и в некоторой точке принимает свое наименьшее значение. Более формально: на промежутке $[a, \xi]$ величина y убывает, на промежутке $[\xi, b]$ возрастает, а в точке ξ принимает наименьшее значение. При этом не исключается, что $a = \xi$ (величина y возрастает на всем промежутке) или $\xi = b$ (величина y убывает на всем промежутке). Требуется указать оптимальный алгоритм определения минимума, т. е. приближенного нахождения точки минимума и значения функции в этой точке при условии, что разрешается произвести n наблюдений: можно произвольным образом выбрать n точек $x_1, x_2, \ldots, x_n$ на промежутке $[a, b]$ и получить соответствующие им значения $y_1, y_2, \ldots, y_n$.

Уточним понятие оптимальности. Пусть Ω – некоторый алгоритм определения точки минимума. Параметрами, определяющими его работу, служат величины n, a, b. Алгоритм Ω может быть применен к любой функции $y = f(x)$, определенной на промежутке $[a, b]$, которая убывает на $[a, \xi]$ и возрастает на $[\xi, b]$, где ξ – некоторая (неизвестная) точка из $[a, b]$. Будем для краткости называть такие функции *функциями с одним минимумом* (рис. 3).

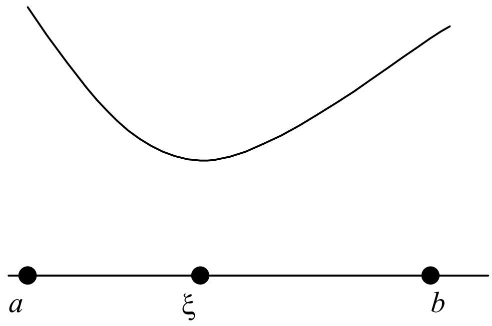


Рис. 3

Множество всех функций с одним минимумом на промежутке $[a, b]$ обозначим через $V(a, b)$ (или просто V, если из контекста ясно, о каком промежутке идет речь). Результатом применения алгоритма Ω к функции $f \in V$ является некоторая пара чисел (ω, η) такая, что $\omega \in [a, b]$ – приближенное значение точки минимума, а $\eta = f(\omega)$. Мы будем писать $\omega = \Omega(f; n, a, b)$ или просто $\omega = \Omega(f)$, когда ясно, каковы значения параметров n, a, b. Качество алгоритма будем оценивать абсолютной величиной погрешности определения точки минимума в наихудшем случае:

$$\tau(\Omega) = \sup \{|\omega - \xi| : f \in V\}.$$

Будем считать алгоритм Ω *оптимальным*, если он дает наименьшую ошибку, т. е. $\tau(\Omega) \le \tau(\Omega')$ для любого алгоритма определения минимума Ω'.

Таким образом, если положить

$$ф_0 = \min_{\Omega} \sup_{f} \left| щ - о \right|,$$

то $\tau_0 = \tau(\Omega')$ для оптимального алгоритма Ω.

Вообще говоря, значение τ_0 зависит от числа допустимых наблюдений n и промежутка $[a, b]$. Легко понять, что при фиксированном числе наблюдений существенным является не сам отрезок $[a, b]$, а его длина $l = b - a$, так что $\tau_0 = \tau_0(n, l)$. Далее, относительно переменной l функция $\tau_0 = \tau_0(n, l)$ является однородной первой степени: $\tau_0(n, \lambda l) = \lambda\tau_0(n, l)$ для $\lambda > 0$ (оптимальность алгоритма не зависит от выбора единицы измерения длины). В оставшейся части параграфа будет установлено, что

$$\tau_0(n, l) = \frac{l}{F_{n+2}},$$

где F_{n+2} – число Фибоначчи. При этом в оптимальном алгоритме в качестве первых точек наблюдения берутся «фибоначчиевы узлы»:

$$c = a + \frac{F_n}{F_{n+2}}(b - a); \; d = a + \frac{F_{n+1}}{F_{n+2}}(b - a).$$

Начнем с анализа алгоритмов поиска минимума при малом числе наблюдений.

1. При $n = 0$ алгоритма поиска минимума не существует: чтобы указать (пусть и произвольным образом) точку и значение функции в ней, нужно произвести хотя бы одно

наблюдение – вычисление значения функции в этой точке. При $n = 1$ оптимальный алгоритм состоит в том, чтобы в качестве ω указать середину отрезка $[a, b]$. В этом случае $|\omega - \xi| \le l/2$, где бы ни располагалась точка минимума ξ. При этом одно наблюдение не дает никакой информации о расположении точки минимума. В частности, возможно $\xi = a$ или $\xi = b$. Какова бы ни была точка ω, имеем $|\omega - \xi| \ge l/2$ для функции с минимумом в точке a, или $|\omega - \xi| \ge l/2$ для функции с минимумом в точке b. Значит,

$$\sup \{|\omega - \xi| : f \in V\} \ge l/2.$$

Таким образом,

$$\tau_0(1, l) = l/2.$$

2. Рассмотрим случай, когда допустимы два наблюдения, $n = 2$. Пусть x_1 и x_2 – точки, в которых производились наблюдения, а y_1, y_2 – значения функции в точках x_1 и x_2 соответственно. Для определенности будем считать, что $x_1 < x_2$. Если $y_1 \le y_2$, то можно утверждать, что точка минимума расположена на отрезке $[a, x_2]$ (рис. 4).

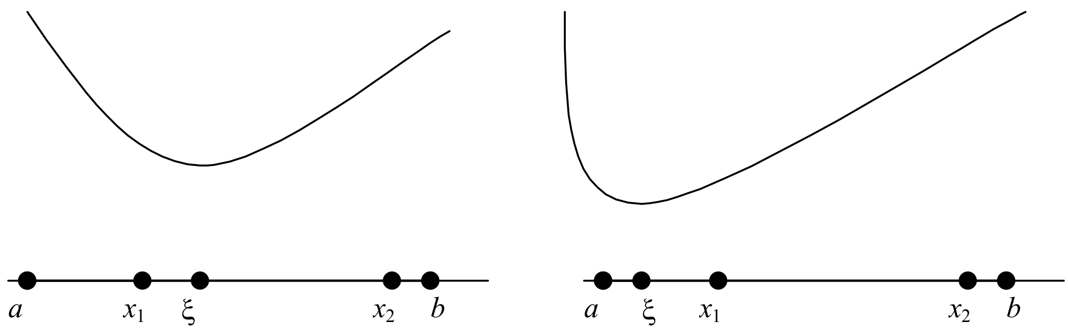


Рис. 4

Так как лимит наблюдений (вычислений значения функции) исчерпан, лучшим претендентом на роль точки минимума в этом случае служит x_1. Наибольшее отклонение τ_1 от настоящей точки минимума получится в том случае, когда минимум достигается в концевой точке отрезка $[a, x_2]$, наиболее удаленной от x_1, то есть

$$\tau_1 = \max\,\{x_1 - a, x_2 - x_1\}.$$

Аналогично при $y_2 \le y_1$ лучшим претендентом на роль точки минимума служит x_2. Отклонение от настоящей точки минимума в «наихудшем» случае составляет

$$\tau_2 = \max\,\{b - x_2, x_2 - x_1\}.$$

Положим

$$\tau = \max\,\{\tau_1, \tau_2\} = \max\,\{x_1 - a, b - x_2, x_2 - x_1\}.$$

Так как

$$(x_1 - a) + (b - x_2) + (x_2 - x_1) = b - a,$$

то τ принимает наименьшее значение τ_0 в случае, когда

$$x_1 - a = b - x_2 = x_2 - x_1 = (b - a)/3,$$

и при этом $\tau_0 = (b - a)/3$. Следовательно,

$$\tau_0(2, l) = l/3.$$

3. Случай трех наблюдений, $n = 3$, особенно поучителен. Покажем, что в этом случае $\tau_0 = l/5$, а в оптимальном алгоритме две точки наблюдения суть

$$x_{10} = \tfrac{3}{5}a + \tfrac{2}{5}b, \quad x_{20} = \tfrac{2}{5}a + \tfrac{3}{5}b,$$

а третья точка определяется как

$$x_{30} = \frac{4}{5}a + \frac{1}{5}b$$

при $f(x_{10}) \le f(x_{20})$, и

$$x_{30} = \frac{1}{5}a + \frac{4}{5}b$$

при $f(x_{10}) > f(x_{20})$.

Если $f(x_{30}) \le f(x_{10}) \le f(x_{20})$ (в этом случае $\xi \in [a, x_{10}]$), то полагаем $\omega = x_{30}$. Если $f(x_{10}) \le f(x_{20})$, но $f(x_{30}) > f(x_{10})$ (в этом случае $\xi \in [x_{30}, x_{20}]$), то полагаем $\omega = x_{10}$.

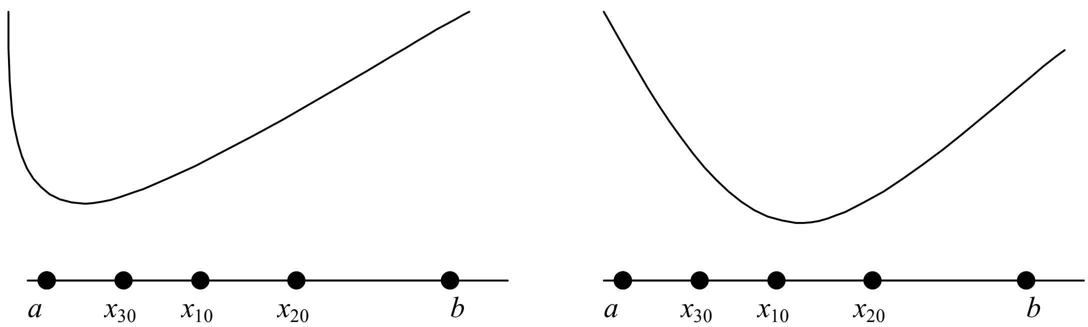


Рис. 5

Погрешность в худшем случае составит $x_{30} - a = x_{20} – x_{10} = l/5$ (рис. 5).

Аналогично, если $f(x_{10}) > f(x_{20})$, полагаем $\omega = x_{20}$ при $f(x_{20}) \le f(x_{30})$, или $\omega = x_{30}$ при $f(x_{20}) < f(x_{30})$. Здесь погрешность в худшем случае также равна $l/5$.

Покажем теперь, что при ином выборе точек наблюдения погрешность (в неблагоприятном случае) окажется больше чем $l/5$.

Предположим, что произведены два наблюдения в точках x_1 и x_2, для определенности пусть $x_1 < x_2$. Если $f(x_1) \le f(x_2)$, то можно утверждать, что точка минимума расположена на отрезке $[a, x_2]$. Погрешность, с которой можно, произведя два наблюдения, найти точку минимума на отрезке длины $x_2 - a$, оценивается величиной

$$\tau_0(2, x_2 - a) = \frac{x_2 - a}{l} \phi_0(2, l) = \frac{x_2 - a}{l} \cdot \frac{1}{3} = \frac{x_2 - a}{3l}.$$

Если $x_2 > x_{20} = \frac{2}{5}a + \frac{3}{5}b$, то $\tau_0(2, x_2 - a) > (b - a)/5 = l/5$. Аналогично можно показать, что и при $x_2 < x_{20}$ погрешность определения точки минимума в худшем случае превосходит $l/5$. Таким образом, если $x_2 \neq x_{20}$, погрешность в определении минимума превосходит $l/5$. Точно так же доказывается, что и при $x_1 \neq x_{10}$ погрешность в определении минимума превосходит $l/5$.

Тем самым описанный в начале этого пункта алгоритм оказывается оптимальным алгоритмом определения минимума при трех наблюдениях, и

$$\tau_0(3, l) = l/5.$$

Приведем теперь общее рекурсивное описание оптимального алгоритма поиска минимума.

Обозначим описанные выше оптимальные алгоритмы поиска минимума при одном, двух или трех допустимых наблюдениях через Φ_1, Φ_2, Φ_3 соответственно. Мы дадим рекурсивное описание последовательности алгоритмов $\Phi_1, \Phi_2, \Phi_3, \ldots$, в которой Φ_k – алгоритм поиска минимума при k допустимых наблюдениях. Параметрами алгоритма Φ_k служат концы отрезка $[a, b]$, на котором ищется минимум, аргументом – функция $y = f(x)$, определенная на отрезке $[a, b]$. Результатом работы алгоритма является пара чисел ω и $\eta = f(\omega)$.

Предположим, что уже имеются описания алгоритмов Φ_1, Φ_2, …, Φ_n. При этом алгоритмы удовлетворяют следующим требованиям:

1) при исполнении алгоритма Φ_k значения функции f вычисляются k раз;
2) если $k \geq 2$, первые два вычисления производятся в точках
$$\frac{F_{k+1}}{F_{k+2}}a + \frac{F_k}{F_{k+2}}b,\ \frac{F_k}{F_{k+2}}a + \frac{F_{k+1}}{F_{k+2}}b;$$
3) погрешность определения точки минимума в худшем случае составляет $\varepsilon_k(l) = l/F_{k+2}$, где $l = b - a$ – длина отрезка;
4) алгоритм Φ_k является оптимальным алгоритмом поиска минимума с числом наблюдений, не превосходящим k.

Дадим теперь *описание алгоритма* Φ_{n+1}. Положим
$$x_1 = \frac{F_{n+2}}{F_{n+3}}a + \frac{F_{n+1}}{F_{n+3}}b,\ x_2 = \frac{F_{n+1}}{F_{n+3}}a + \frac{F_{n+2}}{F_{n+3}}b.$$
Вычисляем $f(x_1)$ и $f(x_2)$. Пусть сначала $f(x_1) \leq f(x_2)$.

Применим алгоритм Φ_n для поиска минимума на отрезке $[a, x_2]$. Пусть ω и $\eta = f(\omega)$ – результат этого применения. Примем пару чисел ω и $\eta = f(\omega)$ в качестве результата применения алгоритма Φ_{n+1} к отрезку $[a, b]$. При исполнении алгоритма Φ_n применительно к отрезку $[a, x_2]$ первые два вычисления должны производиться в точках

$$\frac{F_{n+1}}{F_{n+2}}a+\frac{F_n}{F_{n+2}}x_2\,,\ \frac{F_n}{F_{n+2}}a+\frac{F_{n+1}}{F_{n+2}}x_2\,.$$

Используя тождество

$$F_nF_{n+3}+F_{n+1}^2=(F_{n+2}-F_{n+1})(F_{n+2}+F_{n+1})+F_{n+1}^2=F_{n+2}^2\,,$$

получаем

$$\frac{F_n}{F_{n+2}}a+\frac{F_{n+1}}{F_{n+2}}x_2=\frac{F_n}{F_{n+2}}a+\frac{F_{n+1}}{F_{n+2}}\left(\frac{F_{n+1}}{F_{n+3}}a+\frac{F_{n+2}}{F_{n+3}}b\right)=$$

$$=\frac{F_nF_{n+3}+F_{n+1}^2}{F_{n+2}F_{n+3}}a+\frac{F_{n+1}}{F_{n+2}}\frac{F_{n+2}}{F_{n+3}}b=\frac{F_{n+2}^2}{F_{n+2}F_{n+3}}a+\frac{F_{n+1}}{F_{n+2}}\frac{F_{n+2}}{F_{n+3}}b=$$

$$=\frac{F_{n+2}}{F_{n+3}}a+\frac{F_{n+1}}{F_{n+3}}b=x_1\,,$$

так что одно из двух вычислений уже выполнено, и для последующего исполнения алгоритма Φ_n потребуется $n-1$ вычисление. В общей сложности (с учетом вычислений $f(x_1)$ и $f(x_2)$) выполнено $n+1$ вычисление значений функции f. Таким образом, алгоритм Φ_{n+1} удовлетворяет первым двум требованиям, предшествующим его описанию. Оценим теперь возникающую погрешность $\varepsilon=\varepsilon_n(x_2-a)$. Так как

$$x_2-a=\frac{F_{n+1}}{F_{n+3}}a+\frac{F_{n+2}}{F_{n+3}}b-a=\frac{F_{n+1}-F_{n+3}}{F_{n+3}}a+\frac{F_{n+2}}{F_{n+3}}b=$$

$$=\frac{-F_{n+2}}{F_{n+3}}a+\frac{F_{n+2}}{F_{n+3}}b=\frac{F_{n+2}}{F_{n+3}}(b-a)\,,$$

то

$$\varepsilon=\varepsilon_n(x_2-a)=(x_2-a)/F_{n+2}=(b-a)/F_{n+3}.$$

Следовательно, третье условие для Φ_{n+1} также выполняется.

Пусть теперь $f(x_1) > f(x_2)$. Применим алгоритм Φ_n для поиска минимума на отрезке $[x_1, b]$. Пусть ω и $\eta = f(\omega)$ – результат этого применения. Примем пару чисел ω и $\eta = f(\omega)$ в качестве результата применения алгоритма Φ_{n+1} к отрезку $[a, b]$. Так же, как и в предыдущем случае, имеем:

$$\frac{F_{n+1}}{F_{n+2}} x_1 + \frac{F_n}{F_{n+2}} b = \frac{F_{n+1}}{F_{n+2}} \left(\frac{F_{n+2}}{F_{n+3}} a + \frac{F_{n+1}}{F_{n+3}} b \right) + \frac{F_n}{F_{n+2}} b =$$

$$= \frac{F_{n+1}}{F_{n+3}} a + \frac{F_{n+1}^2 + F_{n+3} F_n}{F_{n+2} F_{n+3}} b = \frac{F_{n+1}}{F_{n+3}} a + \frac{F_{n+1}^2 + (F_{n+2} + F_{n+1})(F_{n+2} - F_{n+1})}{F_{n+2} F_{n+3}} b =$$

$$= \frac{F_{n+1}}{F_{n+3}} a + \frac{F_{n+2}^2}{F_{n+2} F_{n+3}} b = \frac{F_{n+1}}{F_{n+3}} a + \frac{F_{n+2}}{F_{n+3}} b = x_2,$$

так что для исполнения алгоритма Φ_n потребуется $n - 1$ вычисление. Как и в предыдущем случае, алгоритм Φ_{n+1} удовлетворяет обоим требованиям, предшествующим его описанию. Третье требование тоже выполняется. Так как

$$b - x_1 = b - \frac{F_{n+2}}{F_{n+3}} a - \frac{F_{n+1}}{F_{n+3}} b = -\frac{F_{n+2}}{F_{n+3}} a + \frac{F_{n+3} - F_{n+1}}{F_{n+3}} b =$$

$$-\frac{F_{n+2}}{F_{n+3}} a + \frac{F_{n+2}}{F_{n+3}} b = \frac{F_{n+2}}{F_{n+3}} (b - a),$$

то

$$\varepsilon = \varepsilon_n (b - x_1) = (b - x_1)/F_{n+2} = (b - a)/F_{n+3}.$$

Осталось доказать, что алгоритм Φ_{n+1} является оптимальным при условии, что используется не более $n + 1$ наблюдения.

Пусть t_1 и t_2, $t_1 < t_2$, – точки первых двух наблюдений. Если $f(t_1) \le f(t_2)$, точку минимума нужно искать на отрезке $[a, t_2]$, в противном случае – на отрезке $[t_1, b]$. Поскольку минимум может оказаться в одной из концевых точек этих отрезков, сузить область поиска нельзя. Предположим, что

$$t_1 < \frac{F_{n+2}}{F_{n+3}} a + \frac{F_{n+1}}{F_{n+3}} b .$$

При $f(t_1) > f(t_2)$ требуется найти точку минимума на отрезке $[t_1, b]$, произведя не более $n - 1$ наблюдения. Можно, правда воспользоваться тем, что значение в точке $t_2 \in [t_1, b]$ уже вычислено. По индуктивному предположению оптимальным алгоритмом определения минимума за n наблюдений служит Φ_n. Если точка t_2 выбрана в соответствии с алгоритмом Φ_n, то применение этого алгоритма обеспечит погрешность

$$\varepsilon = \varepsilon_n(b - t_1) = (b - t_1)/F_{n+2}.$$

Применение любого другого алгоритма даст большую погрешность. Если

$$t_1 < \frac{F_{n+2}}{F_{n+3}} a + \frac{F_{n+1}}{F_{n+3}} b ,$$

то

$$b - t_1 > \frac{F_{n+2}}{F_{n+3}} (b - a)$$

и, значит,

$$\varepsilon > (b - a)/F_{n+3}.$$

Следовательно, в оптимальном алгоритме должно быть

$$t_1 \geq \frac{F_{n+2}}{F_{n+3}} a + \frac{F_{n+1}}{F_{n+3}} b .$$

Если точка минимума ξ находится на отрезке $[a, t_1]$, то на ее поиск остается $n - 1$ наблюдение (знание $f(t_1)$ не дает в этом случае никакой дополнительной информации о расположении точки минимума). По индуктивному предположению оптимальным алгоритмом определения минимума за $n - 1$ наблюдение является Φ_{n-1}. Этот алгоритм обеспечивает определение минимума с погрешностью

$$\varepsilon = \varepsilon_{n-1}(t_1 - a) = (t_1 - a)/F_{n+1}.$$

Если

$$t_1 > \frac{F_{n+2}}{F_{n+3}} a + \frac{F_{n+1}}{F_{n+3}} b,$$

то

$$t_1 - a > \frac{F_{n+1}}{F_{n+3}}(b - a).$$

Значит, и в этом случае

$$\varepsilon > (b - a)/F_{n+3}.$$

Следовательно, в оптимальном алгоритме

$$t_1 = \frac{F_{n+2}}{F_{n+3}} a + \frac{F_{n+1}}{F_{n+3}} b.$$

По симметрии в оптимальном алгоритме должно также выполняться и равенство

$$t_2 = \frac{F_{n+1}}{F_{n+3}} a + \frac{F_{n+2}}{F_{n+3}} b.$$

Но справедливость последних двух равенств и означает, что алгоритм Φ_{n+1} является оптимальным.

13. Графы. Основные понятия

13. 1. Понятие графа

Линейность является характерной чертой большинства современных естественных и искусственных языков. Линейное представление информации (в виде последовательности символов) не является естественным с точки зрения человеческого восприятия. Использование нелинейных форм во многих случаях существенно облегчает понимание. В математике главным средством нелинейного представления информации служат чертежи. Один из видов чертежей – графы, которые, сохранив присущую чертежам наглядность, допускают точное теоретико-множественное описание и тем самым становятся объектом математического исследования.

В разных задачах удобно использовать чертежи разных типов. Соответственно определенные вариации допускает и определение графа. Неотъемлемыми атрибутами графов (при всем разнообразии определений) являются вершины и соединяющие их ребра или дуги.

Граф $G = (V, E)$ состоит из конечного множества *вершин* (или *узлов*) V и конечного множества *ребер* E. Каждое ребро связывает (соединяет) пару вершин. Если ребро a соединяет вершины x и y, то говорят, что ребро a и вершины x, y *инцидентны*.

Например, на рис. 1

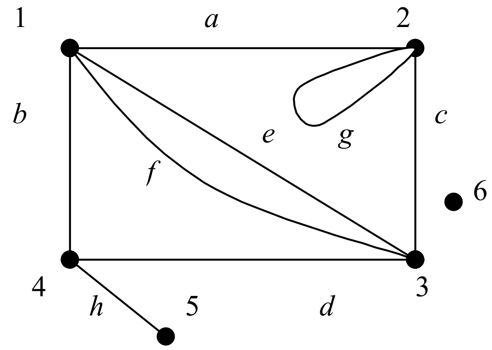


Рис. 1

изображен граф с шестью вершинами, обозначенными цифрами 1, 2, 3, 4, 5, 6, и восемью ребрами, обозначенными буквами a, b, c, d, e, f, g, h. Ребро a связывает вершины 1 и 2; ребра e и f связывают вершины 1 и 4; ребро g связывает вершину 2 саму с собой; вершина 1 инцидентна ребрам a, b, e, f; ребро c инцидентно вершинам 2 и 3.

Два ребра, связывающие одну и ту же пару вершин (как e и f), называют *параллельными* (или *кратными*); ребро, связывающее вершину саму с собой (как g), называют *петлей*. Иногда в определении графа запрещают наличие параллельных ребер и/или петель, иногда нет. Мы не будем жестко фиксировать определение, оговаривая специально, если это оказывается существенным, какого типа граф рассматривается.

Пусть $G = (V, E)$ – некоторый граф. Граф $G' = (V', E')$, вершины и ребра которого являются вершинами и ребрами графа G, т.е. $V' \subset V$, $E' \subset E$ называется *подграфом* графа G.

Степенью вершины графа называется число ребер графа, инцидентных этой вершине (петли считаются дважды). Степень вершины v обозначается $\delta(v)$. Вершина степени 0 называется *изолированной*, вершина степени 1 – *висячей*. Так, для графа на рис. 1 имеем: $\delta(1) = \delta(2) = \delta(3) = 4$, $\delta(4) = 3$, $\delta(5) = 1$, $\delta(6) = 0$; вершина 5 – висячая, вершина 6 – изолированная.

Несложно убедиться в справедливости следующего соотношения:

$$\sum_{v\in V} д(v) = 2m,$$

где m – число ребер графа $G = (V, E)$. В самом деле, ребро, соединяющее вершины x и y, вносит вклад по единице в слагаемые : $\delta(x)$ и $\delta(y)$ (при $x = y$ ребро является петлей и в соответствии с определением вносит вклад 2 в одно слагаемое $\delta(x)$).

В некоторых случаях рассматриваются направленные ребра, которые называют *дугами*. Для дуги, соединяющей две вершины, указывают, из какой вершины она выходит (начало дуги), и в какую входит (конец дуги). На рисунке направление дуги указывают стрелкой. Если все ребра графа направлены, его называют *ориентированным* графом, или *орграфом*. В орграфе параллельными считаются дуги, соединяющие одинаковые вершины и имеющие одинаковое направление, то есть дуги, имеющие общее начало и общий конец. Когда мы, имея в виду ориентированный граф, говорим, что дуга a соединяет вершины x и y, предполагается, что дуга a направлена от x к y.

На рис. 2 изображен орграф. Из вершины 1 выходят дуги a и b, в нее входит дуга e.

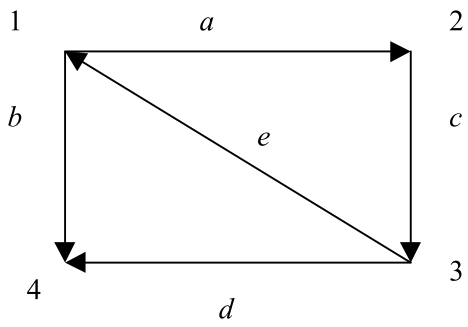


Рис. 2

Полустепенью исхода вершины орграфа называется число дуг графа, начинающихся в этой вершине; *полустепенью захода* – число дуг графа, заканчивающихся в ней. Полустепени исхода и захода вершины v обозначаются соответственно через $\delta^+(v)$ и $\delta^-(v)$. Так, для графа на рис. 2 имеем $\delta^+(1) = 2$, $\delta^-(1) = 1$.

Для ориентированного графа $G = (V, E)$, содержащего m дуг, выполняется следующее соотношение:

$$\sum_{v\in V} \text{д}^+(v) = \sum_{v\in V} \text{д}^-(v) = m.$$

Вершины и дуги графа могут быть дополнительно помечены. В этом случае говорят о *нагруженном*, или *взвешенном*, графе.

Подграфом орграфа G называют любой орграф, вершины которого составляют часть множества вершин графа G, а дуги – часть множества его дуг.

13. 2. Маршруты, цепи и циклы

Последовательность вершин v_0, v_1, v_2, …, v_k графа G представляет собой *маршрут* в этом графе от вершины v_0 к вершине v_k, если для любого $i = 0, 1, 2, \ldots, k–1$ вершины v_i и v_{i+1} соединены дугой. В случае, когда допускаются параллельные дуги, нужно дополнительно указать, по какой дуге из v_i в v_{i+1} проходит маршрут. В этом случае маршрут от вершины v_0 к вершине v_k, задается последовательностью вида

$$v_0, a_1, v_1, a_2, v_2, \ldots, v_{k-1}, a_k, v_k,$$

где v_0, v_1, …, v_k – последовательность вершин, a_1, a_2, …, a_k – последовательность дуг, причем дуга a_i соединяет вершину v_{i-1} с вершиной v_i. На самом деле, поскольку концы дуг определены однозначно, маршрут можно представить последовательностью дуг a_1, a_2, …, a_k. *Длиной маршрута* считается число дуг, которые он содержит. Все вершины маршрута, кроме начальной и конечной, называют внутренними или промежуточными. Вообще говоря, и начальная, и конечная вершины могут встретиться на маршруте как промежуточные вершины. Для любой вершины имеется маршрут из этой вершины в нее же, не содержащий ни одной дуги (длины 0).

Маршрут называется *цепью*, если каждая дуга встречается в нем не более одного раза, и простой цепью, если любая вершина графа инцидентна не более, чем двум дугам маршрута. *Путем* называют маршрут, в котором все вершины различны.

Впрочем, довольно часто «путь» используют как синоним «маршрута».

Если начальная вершина маршрута совпадает с конечной, его называют *замкнутым*. Замкнутый маршрут называется *циклом*, если он является цепью; если эта цепь к тому же простая, то и цикл называется простым. Таким образом, цикл – это замкнутый маршрут, у которого все вершины различны, кроме первой и последней. Например, в графе на рис.2 маршрут $1a2c3e1$, или, короче, *ace*, является простым циклом. Поскольку параллельных дуг на графе нет, этот цикл можно указать и по вершинам: 1231. Ясно, что маршруты 2312 и 3123 представляют тот же цикл. Граф, не содержащий циклов, называется *ациклическим*.

Будем говорить, что вершина y *достижима* из вершины x, если в графе G имеется путь из x в y.

Пусть G – произвольный орграф. Пополним его новыми дугами. Новая дуга из вершины x в вершину y проводится в том случае, если y достижима из x, а граф G не содержит дуги из x в y. Обозначим пополненный граф через $G^{\wedge}$. Минимальный подграф графа G, индуцирующий на множестве вершин то же отношение достижимости и, соответственно, порождающий тот же пополненный граф $G^{\wedge}$, что и граф G, обозначается через G_B и называется *базисным* графом для графа G. Для ациклического графа его базисный граф определен однозначно. На рис. 3

представлен ациклический граф; «жирными» наконечниками отмечены дуги, входящие в базисный граф.

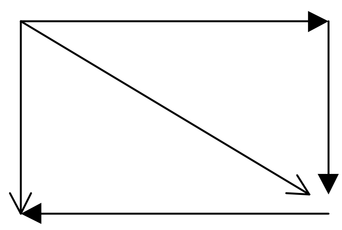


Рис. 3

На множестве вершин неориентированного графа G отношение достижимости является отношением эквивалентности. Класс эквивалентности составляют все вершины, которые могут быть связаны друг с другом некоторым путем. Эти классы эквивалентности называются *компонентами связности*. Неориентированный граф G называется связным, если в нем любые две вершины можно соединить путем. Связный граф имеет всего одну компоненту связности. На рис. 4 изображен граф с четырьмя компонентами связности.

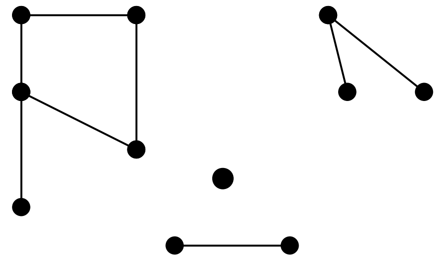


Рис. 4

13. 3. Эйлеровы цепи и циклы

Считается, что теория графов началась с *задачи о кенигсбергских мостах*, поставленной и решенной Эйлером. На рис. 5 приведена схема мостов в г. Кенигсберге времен Эйлера.

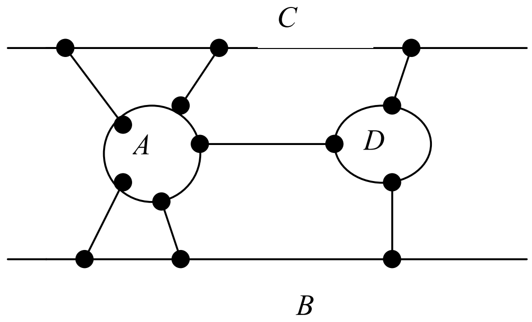


Рис. 5

На этой схеме B и C – берега реки Преголь, а A и D – острова. Острова соединены с берегами и друг с другом семью мостами. Требуется так составить маршрут, чтобы пройти каждый мост ровно по одному разу и вернуться в исходную точку. Можно построить граф задачи, в котором каждой части города соответствует вершина, а каждому мосту – ребро (рис. 6).

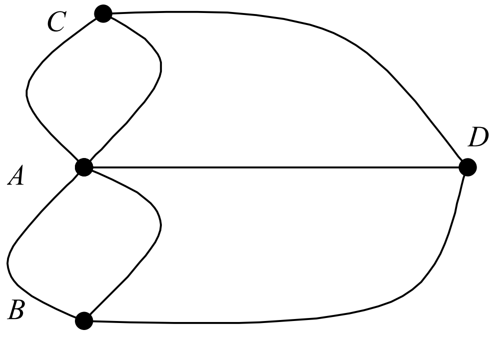


Рис. 6

Решение задачи о кенигсбергских мостах сводится теперь к поиску цикла на построенном графе, в который все ребра графа входят по одному разу. В общем случае цикл, обладающий таким свойством, называется *эйлеровым*. Аналогично цепь называется эйлеровой, если она проходит по одному разу через каждое ребро.

Задача о кенигсбергских мостах решается с помощью следующего несложного рассуждения. Предположим, что на графе с рис. 6 существует эйлеров цикл. Рассмотрим последовательность «выходов» – «заходов» для вершины из этого цикла. Если вершина промежуточная, последовательность имеет вид: «зашел» – «вышел» – «зашел» – «вышел» –… – «зашел» – «вышел»; если вершина начальная: «вышел» – «зашел» – «вышел» – «зашел» – … – «вышел» – «зашел». В любом случае вершина должна быть инцидентна четному числу ребер, по которым только и можно «зайти» и «выйти». Таким образом, если на графе имеется эйлеров цикл, степени всех вершин должны быть четными. Граф на рис. 6 этим свойством не обладает, а значит, составит соответствующий маршрут невозможно.

Используя аналогичное рассуждение нетрудно показать, что если на графе существует эйлерова цепь, то на этом графе ровно две вершины имеют нечетную степень. В самом деле, для промежуточных вершин эйлеровой цепи последовательность «выходов» – «заходов» имеет вид: «зашел» – «вышел» –… – «зашел» – «вышел»; для начальной вершины: «вышел» –… – «зашел» – «вышел»; для конечной: зашел» – «вышел» –… – «зашел». Таким образом, все промежуточные вершины имеют четные степени, а начальная и конечная – нечетные. Менее тривиальным является обратное утверждение.

Теорема. *Связный граф обладает эйлеровым циклом тогда и только тогда, когда степени всех его вершин четны.*

13. 4. Матрицы смежности и инцидентности

Любой ориентированный граф с вершинами $v_1, v_2, \ldots, v_n$ и дугами $a_1, a_2, \ldots, a_m$ можно задать его *матрицей инцидентности* $B=(b_{ij})$, $i = 1,2,\ldots, n$, $j = 1,2,\ldots, m$, размера $n\times m$, в которой $b_{ij} = 1$, если дуга a_j исходит из вершины v_i; $b_{ij} = -1$, если дуга a_j заходит в вершину v_i; $b_{ij} = 0$, если дуга a_j не инцидентна вершине v_i.

Матрицу инцидентности можно использовать и для задания неориентированного графа. В этом случае $b_{ij} = 1$, если дуга a_j инцидентна вершине v_i, и $b_{ij} = 0$, если дуга a_j не инцидентна вершине v_i. Например, граф на рис. 2 (с. 107) можно задать следующей матрицей инцидентности (дуги упорядочены в алфавитном порядке):

$$B = \begin{pmatrix} 1 & 1 & 0 & 0 & -1 \\ -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & -1 & 0 & -1 & 0 \end{pmatrix}.$$

Графы без параллельных дуг удобно представлять при помощи *матриц смежности*. Для графа с n вершинами матрица смежности – это квадратная матрица A=(a_{ij}) порядка n, состоящая из нулей и единиц. Элемент a_{ij} равен 1, если имеется дуга, соединяющая вершины i и j, и равен 0 в противном случае. Впрочем, если в графе имеются параллельные дуги, то и тогда можно использовать матрицу смежности: можно полагать a_{ij} равным числу дуг, соединяющих вершины i и j.

Матрица смежности неориентированного графа симметрична. Например, матрицей смежности графа, представленного на рис. 7.

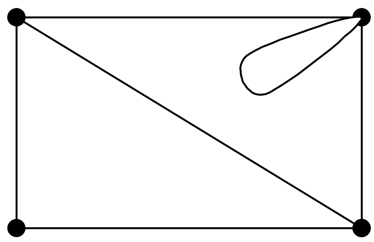


Рис. 7

служит матрица

$$A = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$$

при том, что вершины занумерованы, начиная с левой верхней, по часовой стрелке. Если изменить порядок нумерации вершин, то изменится и матрица смежности. Например, нумеруя вершины того же графа по часовой стрелке, начав с правой верхней вершины, мы получим матрицу смежности

$$A' = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}.$$

Обе матрицы представляют один и тот же граф и получаются одна из другой перестановкой строк и столбцов. Вообще, любая перестановка, применяемая одновременно и к строкам и к столбцам матрицы смежности некоторого графа, приводит снова к матрице смежности того же графа. В случае, когда вершины графа упорядочены, матрица смежности определена однозначно.

Матрица смежности ориентированного графа, вообще говоря, несимметрична. Например, следующая матрица является матрицей смежности ориентированного графа на рис. 2 (с. 107):

$$A = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Теорема. *Пусть G – ориентированный граф с вершинами 1, 2, ..., n, и A = (a_{ij}) – его матрица смежности. Тогда элемент $a_{ij}^{(k)}$ матрицы A^k равен числу путей длины k из вершины i в вершину j.*

Доказательство. Докажем теорему индукцией по k. При $k = 1$ имеем $A^1 = A$, так что заключение теоремы верно, поскольку элемент $a_{ij}^{(1)} = (a_{ij})$ равен числу дуг из i в j, то есть числу путей длины 1. Предположим теперь, что каждый элемент $a_{ij}^{(k)}$ матрицы $A^k = (a_{ij}^{(k)})$ равен числу путей длины k из вершины i в вершину j и покажем, что каждый элемент $a_{ij}^{(k+1)}$ матрицы $A^{k+1} = (a_{ij}^{(k+1)})$ равен числу путей длины k+1 из вершины i в вершину j. Так как $A^{k+1} = A^k \times A$, то

$$a_{ij}^{(k+1)} = a_{i1}^{(k)} \cdot a_{1j} + a_{i2}^{(k)} \cdot a_{2j} + \ldots + a_{in}^{(k)} \cdot a_{nj}\,. \qquad (1)$$

Слагаемое $a_{i1}^{(k)} \cdot a_{1j}$ равно числу путей из i в j длины k+1, в которых вершина 1 предпоследняя: пути длины k из вершины i в вершину 1 комбинируются с дугами из вершины 1 в вершину

j. Аналогично слагаемое $a_{i2}^{(k)} \cdot a_{2j}$ равно числу путей из i в j длины k+1, в которых вершина 2 предпоследняя, и т. д. Каждый путь из i в j длины k+1 получается из некоторого пути длины k, начинающегося в вершине i, присоединением к нему некоторой дуги, заканчивающейся в вершине j. Поэтому сумма в правой части равенства (1) равна числу путей из i в j длины k+1.

Пример. Рассмотрим граф на рис. 8. Пути длины 1 представлены дугами. Все пути длины 2 и более выходят из вершины 2. Путь длины k из вершины 2 в вершину 2 представляет собой петлю, повторенную k раз. Остальные пути получаются как комбинации путей длины 1 и 2 с соответствующим силом повторений петли.

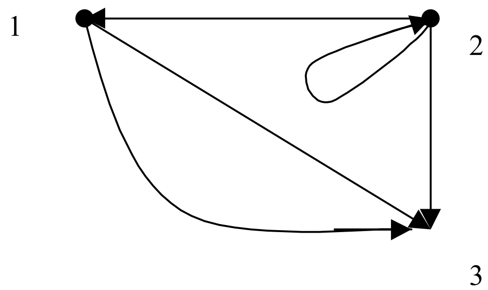


Рис. 8

Матрица смежности:

$$A = \begin{pmatrix} 0 & 0 & 2 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

дает число путей длины 1. Ее квадрат:

$$A^2 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix}$$

– число путей длины 2. Легко видеть, что $A^k = A^2$ при $k \geq 2$.

Замечание. Утверждение предыдущей теоремы можно распространить и на случай $k = 0$. Для каждой вершины графа имеется ровно один путь длины 0 из этой вершины в нее саму. Других путей длины 0 на графе нет. Следовательно, элемент e_{ij} единичной матрицы $E = A^0$ равен числу путей длины 0 из вершины i в вершину j ($e_{ij} = 0$ при $i \neq j$, $e_{ij} = 1$ при $i = j$).

Пусть G – ориентированный граф и A – его матрица смежности. Рассмотрим последовательность матриц

$$A^0, A^1, A^2, \ldots, A^{n-1}. \qquad (2)$$

Зафиксируем пару вершин i и j. Если существует какой-нибудь путь из i в j, то существует и путь длины меньше n. В самом деле, если длина пути превосходит $n - 1$, то такой путь проходит через более чем n вершин, и, значит, на таком пути хотя бы одна вершина, скажем, v, встретится более одного раза. Отбросив часть пути, ведущую из вершины v в нее саму, получаем более короткий путь из i в j. Повторив подобную операцию несколько раз, можно получить путь из i в j, длина которого не превосходит $n - 1$. Таким образом, если из i в j имеется некоторый путь, то в одной из матриц (2) на месте (i,j)

встретится элемент, отличный от нуля. Если в матрице A^k на месте (i,j) находится элемент, отличный от нуля, а во всех предшествующих матрицах A^0, A^1, A^2, …, A^{k-1} на месте (i,j) стоят нули, то k – это длина кратчайшего пути из i в j.

13. 5. Булевы матрицы и операции над ними

Если G – граф (ориентированный или нет) без кратных дуг, то его матрица смежности A является *булевой*, то есть состоит из нулей и единиц. Для произвольной матрицы $X = (x_{ij})$ с неотрицательными элементами будем обозначать через $\mathrm{sign}(X)$ булеву матрицу, полученную из X заменой всех ее положительных элементов единицами. Например,

$$\mathrm{sign}\begin{pmatrix} 1 & 2 & 0 \\ 0 & 2 & 1 \\ 3 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}.$$

Равенство $\mathrm{sign}(X) = X$ означает, что матрица X булева. Легко видеть, что

$$\mathrm{sign}(X+Y) = \mathrm{sign}(\mathrm{sign}(X) + \mathrm{sign}(Y)),$$

$$\mathrm{sign}(X\cdot Y) = \mathrm{sign}(\mathrm{sign}(X)\cdot\mathrm{sign}(Y))$$

в случае, когда X и Y неотрицательные матрицы, для которых определены соответствующие матричные операции. Далее, если матрицы X и Y имеют одинаковую размерность, то

$$\mathrm{sign}(X+Y) = \mathrm{sign}(X) \vee \mathrm{sign}(Y)$$

(дизъюнкция булевых матриц вычисляется поэлементно).

Пусть A и B – булевы матрицы. Матрицу $\operatorname{sign}(A \cdot B)$ будем называть их *булевым произведением* и обозначать через $A * B$. В соответствии с определением $\operatorname{sign}(A \cdot B) = A * B$. Если $A = (a_{ij})$ и $B = (b_{ij})$, то элементы булева произведения $A * B = (c_{ij})$ определяются формулами

$$c_{ij} = a_{i1}b_{1j} \vee = a_{i2}b_{2j} \vee \ldots \vee a_{in}b_{nj} \,,$$

где n – число столбцов матрицы A и число строк матрицы B. Положим $A^{[k]} = \operatorname{sign}(A^k)$ и будем называть матрицу $A^{[k]}$ булевой степенью матрицы A.

Если A – матрица смежности графа G, то на месте (i,j) матрицы $A^{[k]}$ находится 1, если на графе G существует путь длины k из i в j, и 0 в противном случае. Пусть n – число вершин графа G. Тогда матрица

$$E \vee A \vee A^{[k]} \vee \ldots \vee A^{[n-1]}$$

содержит 1 на месте (i,j) в том и только том случае, когда на графе G имеется хотя бы один путь из вершины i в вершину j.

13. 6. Бинарные отношения и графы

Бинарное отношение R на конечном множестве V может быть представлено ориентированным графом $G(R)$, называемым *графом отношения* R. Вершинами графа служат элементы множества V; вершины x и y соединены направленной дугой с началом x и концом y, если $(x,y) \in R$. Обратно, всякий ориентированный граф без параллельных дуг G задает бинарное отношение $R(G)$ на множестве своих вершин, чьим графом он и является: вершины x и y связаны отношением $R(G)$, если они соединены направленной дугой с началом x и концом y. Если R – бинарное отношение на конечном множестве $V = \{1, 2, \ldots, n\}$, а G – граф с вершинами $V = \{1, 2, \ldots, n\}$, то матрица смежности графа G совпадает с характеристической матрицей отношения R в том и только том случае, когда $G = G(R)$ или, что равносильно, $R = R(G)$.

Рассмотрим, как связаны свойства отношения R и соответствующего ему графа $G = G(R)$.

Отношение R *симметрично*, если для любых x, $y \in V$ из xRy следует yRx. Иными словами, если на ориентированном графе G имеется дуга из x в y, то имеется также и дуга из y в x. В этом случае матрица смежности графа G симметрична. По существу, граф G оказывается неориентированным. Можно считать, что симметричным отношениям отвечают неориентированные графы. *Антисимметричность* отношения R означает, что xRy и yRx влечет $x = y$ и равносильна тому, что две различные вершины графа G могут быть связаны дугой лишь в одном направлении. Если отношение R *асимметрично*, то есть xRy влечет $\neg yRx$, то, кроме того, граф G не должен иметь петель.

Если R – *рефлексивное* отношение, то есть xRx для любого $x \in V$, то граф G имеет петлю в каждой вершине, а диагональ матрицы смежности состоит из одних единиц. Соответственно отношение R *антирефлексивно* тогда и только тогда, когда граф G не имеет петель.

Отношение R *транзитивно*, если из xRy и yRz следует xRz. Для графа G это означает, что если G содержит дуги из x в y и из y в z, то он содержит и дугу из x в z. Более того, если существует путь из вершины x в вершину y, то имеется и дуга из x в y.

В терминах графа $G = G(R)$ простую интерпретацию получает понятие *транзитивного замыкания* отношения R. На множестве вершин графа G определим отношение достижимости $R^{\wedge}$, полагая $xR^{\wedge}y$ тогда и только тогда, когда вершина y достижима из вершины x (в графе G имеется путь из x в y). Отношение достижимости $R^{\wedge}$ является транзитивным замыканием отношения R, то есть $R^{\wedge}$ – это наименьшее транзитивное отношение, содержащее R. Матрицей отношения $R^{\wedge}$ служит матрица $E \vee A \vee A^{[k]} \vee \ldots \vee A^{[n-1]}$, где A – матрица смежности графа G.

Отношение R называется *ацикличным*, если граф $G(R)$ не содержит нетривиальных циклов. Если вершины x и y на графе ацикличного отношения R соединены некоторым путем, то в этом графе нет дуги из y в x.

Теорема. *Отношение R на конечном множестве V ациклично тогда и только тогда, когда существует отображение $\varphi: V \to N$ такое, что xRy влечет $\varphi(x) < \varphi(y)$ для любых $x, y \in V$.*

Доказательство. *Достаточность.* Предположим, что отображение φ обладает указанным свойством. Тогда для любого цикла $v_0, v_1, \ldots, v_n, v_0 = v_n$, в графе $G(R)$ имеем:

$$\varphi(v_0) < \varphi(v_1) < \ldots < \varphi(v_n) = \varphi(v_0),$$

что невозможно при $n > 0$.

Необходимость. Пусть $n = |V|$ – число элементов множества V. Для произвольного $x \in V$ рассмотрим множество всевозможных простых цепей в графе $G(R)$, заканчивающихся в вершине x. Ни одна из этих цепей не содержит более чем n вершин, так что их длины ограничены в совокупности. Положим $\varphi(x) = k$, где k – максимальная из длин простых цепей, заканчивающихся в вершине x. Если нет ни одной цепи, ведущей в x, то есть ни одна дуга не заканчивается в x, полагаем $\varphi(x)=0$. Покажем, что так определенное отображение φ обладает нужным свойством. Предположим, что xRy. Пусть $\varphi(x)=k$. Тогда в графе $G(R)$ имеется простая цепь $v_0, v_1, \ldots, v_k = x$. Добавив к этой цепи дугу из x в y, мы получим маршрут $v_0, v_1, \ldots, v_k = x, y$ (длины k+1). Этот маршрут является простой цепью. В самом деле, так как цепь $v_0, v_1, \ldots, v_k$ простая, все вершины в ней различны. Если y совпадает с одной из вершин этой цепи, скажем $y = v_i$, получается цикл $y = v_i, v_1, \ldots, v_k = x, y$, что противоречит ацикличности отношения R. Следовательно, имеется простая цепь длины k+1, заканчивающаяся в y. Значит, $\varphi(y) \geq k+1$ и потому $\varphi(y) > \varphi(x)$.

Пример. Рассмотрим бинарное отношение R, граф которого $G(R)$ представлен на рис 9.

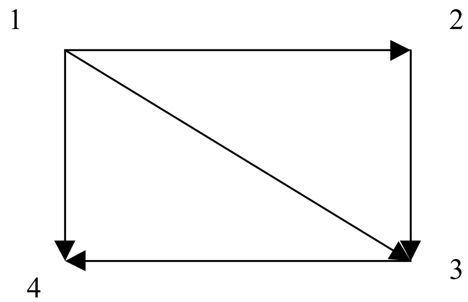


Рис. 9

В соответствии с определением отображения φ находим:

$$\varphi(1) = 0;\ \varphi(2) = 1;\ \varphi(3) = 2;\ \varphi(4) = 3.$$

Из предыдущей теоремы следует, что функция выбора C^R (блокировка), построенная по ацикличному отношению R, является расширением некоторой функции выбора по скалярному критерию. В самом деле, пусть $\Omega = \{1, 2, \ldots, n\}$ – множество альтернатив, а R – ацикличное бинарное отношение на Ω. Тогда, по предыдущей теореме, на множестве Ω существует числовая функция $\varphi : \{1, 2, \ldots, n\} \to N$ такая, что iRj влечет $\varphi(i) < \varphi(j)$. Зададим функцию f (скалярный критерий), полагая $f(i) = n - \varphi(i)$. Ясно, что iRj влечет $f(i) > f(j)$. Следовательно, $f(i) \le f(j)$ влечет $\neg iRj$. Обозначим через C функцию выбора по скалярному критерию f. Тогда

$$x \in C(X) \Leftrightarrow \forall y \in X\ f(x) \ge f(y).$$

С другой стороны, в соответствии с определением функции выбора C^R имеем:

$$x \in C^R(X) \Leftrightarrow \forall y \in X \neg(yRx)\}.$$

Следовательно,

$$x \in C(X) \Rightarrow x \in C^R(X),$$

то есть $C(X) \subset C^R(X)$.

14. Деревья

14.1. Общее понятие дерева

К числу наиболее важных нелинейных структур, связанных с вычислительными алгоритмами, относятся графы специального вида, называемые деревьями. Структура дерева появляется в тех ситуациях, где в той или иной форме имеется «ветвление».

Граф (неориентированный) называется *деревом*, если он связен и не имеет циклов. Например, граф на рис.1 является деревом.

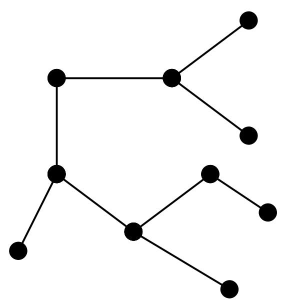


Рис. 1

Укажем некоторые свойства деревьев.

1. *Любые две вершины дерева можно соединить ровно одной простой цепью*.

В силу связности между любыми двумя вершинами дерева имеется соединяющая их цепь. Так как дерево не содержит циклов, эта цепь является простой и единственной.

2. *Если дерево G содержит хотя бы одно ребро, на нем найдется висячая вершина.*

Граф G не содержит циклов, поэтому ни одно его ребро не является петлей, т. е. концевые вершины любого ребра различны. Предположим, что степень любой вершины графа G превосходит 1. Пусть v_1, v_2 – пара вершин, соединенных ребром a_1. Так как степень вершины v_2 больше 1, имеется ребро a_2, отличное от a_1, соединяющее вершину v_2 с некоторой вершиной v_3. Вершины v_1 и v_3 должны различаться, иначе путь v_1, a_1, v_2, a_2 v_3 являлся бы циклом. Пользуясь тем, что степень вершины v_3 больше 1, можно продолжить путь v_1, a_1, v_2, a_2 v_3 ребром a_3, отличным от a_2, соединяющим вершину v_3 с некоторой вершиной v_4. Вершина v_4 должна отличаться от всех предыдущих вершин построенного пути (иначе получается цикл). Подобным образом можно построить сколь угодно длинный путь, на котором все вершины различны. С другой стороны, любой путь, длина которого превосходит число вершин графа, должен содержать повторяющиеся вершины. Таким образом, предположение об отсутствии висячих вершин на графе G приводит к противоречию.

3. *Число ребер дерева G на единицу меньше числа его вершин.*

Доказательство проведем индукцией по числу вершин. Дерево с одной вершиной не содержит ребер, так что при $n = 1$ имеем $m = 0$ – доказываемое соотношение выполняется.

Предположим теперь, что доказываемое равенство справедливо для любого дерева с $n - 1$ вершиной и докажем его справедливость для любого дерева с n вершинами. Пусть $n \geq 2$ – число вершин, а m число ребер дерева G. Покажем, что $m = n - 1$. В соответствии с предыдущим пунктом дерево G содержит висячую вершину v. Удалим эту вершину вместе с инцидентным ей ребром. Легко видеть, что оставшийся граф является деревом, которое содержит $n - 1$ вершину и $m - 1$ ребро. По индуктивному предположению имеем $m - 1 = n - 2$. Отсюда $m = n - 1$.

Отметим, что последнее свойство является характеристическим.

4. *Связный граф, в котором число ребер на единицу меньше числа вершин, является деревом.*

Доказательство проведем индукцией по числу вершин графа n. При $n = 1$ граф состоит из одной вершины и не содержит ребер и, значит, является деревом. Предположим теперь, что доказываемое утверждение справедливо для всех графов, содержащих $n - 1$ вершину, и рассмотрим граф G с n вершинами и $n - 1$ ребром. Сначала заметим, что граф G содержит висячую вершину. В самом деле, как показано в 13.1, сумма степеней вершин графа равна удвоенному числу ребер:

$$\delta(v_1) + \delta(v_2) + \ldots + \delta(v_n) = 2(n - 1).$$

Хотя бы одно слагаемое в левой части этого равенства меньше двух (в противном случае сумма превосходила бы $2n$). Пусть для определенности, $\delta(v_n) < 2$. Так как граф связный, то степень любой вершины не меньше 1. Следовательно, $\delta(v_n) = 1$. Это означает, что вершина v_n висячая. Граф G', получающийся удалением этой вершины вместе с инцидентным ей ребром, связен, содержит $n - 1$ вершину и $n - 2$ ребра. По индуктивному предположению G' является деревом. Но тогда и исходный граф также является деревом (ни один цикл не может содержать висячую вершину v_n, а циклов, не содержащих эту вершину, нет, поскольку они должны были бы целиком находиться в G').

14. 2. Остовное дерево связного графа

Пусть G – некоторый связный граф. Дерево, множество вершин которого совпадает с множеством вершин графа G, а ребра являются ребрами графа G, называется *остовным деревом* графа G. Иными словами, остовное дерево графа G – это его подграф, содержащий все вершины и являющийся деревом.

Пример. На рис. 2 представлен граф с пятью вершинами и восемью ребрами. Четыре выделенные ребра (вместе с пятью вершинами) составляют остовное дерево этого графа. Остовное дерево определено неоднозначно.

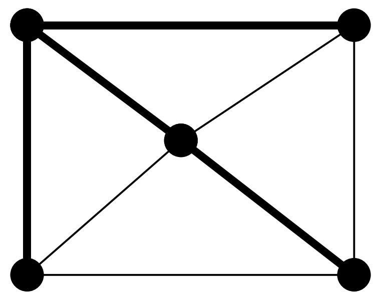


Рис. 2

На рис. 3 выделены ребра, составляющие еще одно остовное дерево того же графа (заметим, что на этом рисунке невыделенные ребра также составляют остовное дерево).

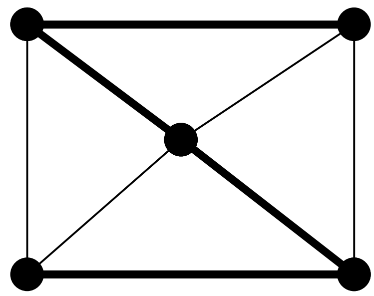


Рис. 3

В прикладном плане задачу нахождения остовного дерева можно понимать как поиск системы связей между заданными точками без ненужного дублирования.

Построение остовного дерева. Остовное дерево T графа G можно сформировать следующим образом. Сначала берем в качестве T произвольную вершину графа G. Далее последовательно наращиваем дерево T в соответствии со следующим правилом: если на графе G имеется ребро, соединяющее вершины u и v такие, что u содержится в T, а v –

нет, добавляем к T это ребро вместе с вершиной v. Покажем, что при этом граф T остается деревом. Во-первых, каждая добавляемая вершина связана с одной из уже входящих в T, так что добавление новых вершин и ребер оставляет граф T связным. Далее, в начальный момент граф T содержит одну вершину и не имеет ни одного ребра. Затем на каждом шаге построения число его вершин и число ребер увеличивается на единицу. Таким образом, число вершин графа T на единицу превосходит число его ребер. Следовательно, граф T является деревом. На некотором шаге дальнейшее расширение T окажется невозможным. Покажем, что в этом случае T представляет собой остовное дерево графа G, т.е. все вершины графа G попали в T. Обозначим через T' граф, который содержит все вершины графа G, не попавшие в T, и все соединяющие их ребра. На графе G нет ребер, которые имели бы одну концевую вершину в T, а другую в T'. Так как граф G связный, T' не может быть непустым.

Пусть n – число вершин, а m – число ребер графа G. Любое его остовное дерево T имеет n вершин и $n - 1$ ребро. Таким образом, остовное дерево получается отбрасыванием $m - n + 1$ ребер графа G. Число $\nu(G) = m - n + 1$ называют *цикломатическим числом* графа G.

Остовное дерево графа G может быть получено последовательным удалением $\nu(G)$ «лишних» ребер: на каждом шаге можно отбросить любое ребро, если это не ведет к нарушению связности графа.

Предположим, что каждому ребру графа приписано некоторое положительное число, называемое его весом или стоимостью (это может быть, например, в случае информационных сетей, стоимость установления связи между концевыми точками). В этом случае граф называется нагруженным. Стоимостью нагруженного графа будем считать суммарную стоимость всех его ребер. Многие задачи, связанные с построением экономичных систем сообщения или информационных систем, приводят к задаче поиска *остовного дерева минимальной стоимости*.

Пусть G – связный нагруженный граф. Обозначим через $c(T)$ стоимость произвольного подграфа T графа G. Если граф T связный, содержит все вершины графа G и имеет минимальную стоимость среди всех подграфов графа G, обладающих этими двумя свойствами, то T является остовным деревом. В самом деле, если бы граф T содержал цикл, то можно было бы отбросить, по крайней мере, одно ребро и тем самым уменьшить стоимость T.

Опишем *алгоритм построения остовного дерева минимальной стоимости*. Он получается небольшим видоизменением алгоритма построения остовного дерева связного графа. Сначала берем в качестве T вершину графа G, из которой выходит ребро минимальной стоимости. Далее, пока это возможно, последовательно наращиваем дерево T: из всех ребер графа G, соединяющих вершины $u \in T$ и $v \notin T$, выбираем ребро минимальной стоимости и добавляем к T это ребро вместе с вершиной v. Полученное в результате исполнения этого алгоритма остовное дерево будет иметь минимальную стоимость.

Пример. На рис. 4 выделено остовное дерево минимальной стоимости. Его стоимость равна 6. Остовное дерево, представленное на рис. 2, при тех же оценках ребер имеет стоимость 8, остовное дерево на рис. 3 – стоимость 8, невыделенные ребра на рис. 3 составляют остовное дерево стоимости 10.

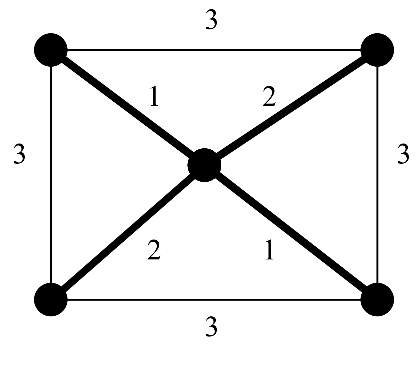


Рис. 4

14. 3. Ориентированные и упорядоченные деревья

Пусть G –дерево с $n > 1$ вершинами. Зафиксируем одну из вершин v_0 и назовем ее *корнем*. Выделение корня определяет на множестве вершин некоторое естественное отношение частичного порядка (ориентацию): вершина u предшествует вершине v (а вершина v следует за вершиной u), если u

встречается на пути из корневой вершины v_0 в вершину v. Если к тому же вершины u и v связаны дугой, то u называется непосредственно предшествующей вершине v (а v – непосредственно следующей за u). Говорят, что вершина v, удаленная на расстояние k от корневой вершины, расположена на уровне k (или является вершиной уровня k). Значение уровня согласуется с отношением порядка: если вершина u предшествует вершине v, то уровень u меньше уровня v. Обратное, вообще говоря, неверно. Представляя дерево графически, вершины обычно располагают так, чтобы значение уровня увеличивалось в направлении сверху вниз.

На рис. 5 представлено дерево, в котором вершина 0 является корнем; вершины 1 и 2 находятся на первом уровне, вершина 3 – на третьем; вершина 1 предшествует вершине 3; вершины 1 и 3 с вершиной 2 отношением предшествования не связаны.

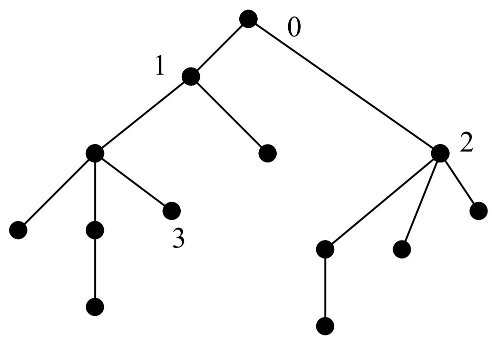


Рис. 5

Вершины, за которыми не следует ни одна вершина, называют *концевыми*, или, в «древесном» стиле, – *листьями*. На рис. 5 вершина 3 является листом, всего же дерево на этом рисунке имеет 7 листьев. У деревьев с выделенным корнем вершины часто называют узлами. Неконцевые узлы называют *узлами ветвления*. Число вершин, непосредственно следующих за узлом ветвления, называют *степенью ветвления* этого узла. На рис 5 узлы 0 и 1 имеют степень ветвления 2, узел 2 – степень 3.

Каждую вершину дерева можно рассматривать как корневую вершину дерева, состоящего из всех следующих за ней вершин. Например, взяв в качестве корневых вершины 1 и 2 (см. рис 5), получим пару деревьев, изображенных на рис. 6.

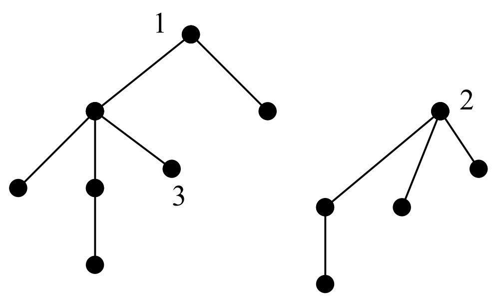


Рис. 6

С учетом этого можно дать следующее рекурсивное определение дерева с корнем (или ориентированного дерева). *Дерево с корнем T* – это конечное множество, состоящее из одного или более узлов таких, что:

1) имеется один выделенный узел, называемый корнем данного дерева;

2) остальные узлы разбиты на $m \geq 0$ попарно непересекающихся множеств, каждое из которых в свою очередь является деревом.

Например, дерево на рис. 7 можно описать следующим образом:

$$T = \{1, T_2\},\ T_2 = \{2, T_3, T_4\},\ T_3 = \{3, T_5, T_6, T_7\},$$

$$T_5 = \{5\},\ T_6 = \{6\},\ T_7 = \{7\},$$

$$T_4 = \{4, T_8\},\ T_8 = \{8, T_9, T_{10}\},\ T_9 = \{9\},\ T_{10} = \{10\}.$$

В явном виде это описание выглядит так:

$$T = \{1; \{2; \{3; \{5\}, \{6\}, \{7\}\}, \{4; \{8; \{9\}, \{10\}\}\}\}\}.$$

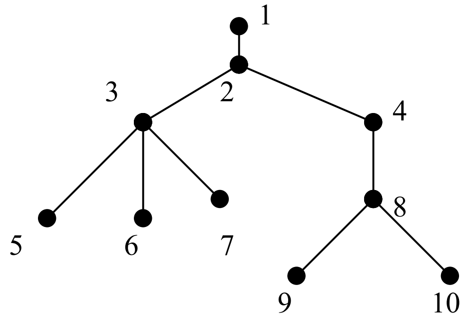


Рис. 7

Если в рекурсивном определении ориентированного дерева считать существенным порядок, в котором перечисляются поддеревья, растущие из заданного корня, получается определение упорядоченного дерева. Например, дерево с рис. 7 можно представить как упорядоченное дерево

$$T = (1; (2; (3; (5, 6, 7)), (4; (8; (9, 10))))).$$

Возможно и иное представление. Например,

$$T' = (1; (2; (4; (8; (9, 10))), (3; (5, 6, 7)))).$$

Более формально:

1) всякий узел a является упорядоченным деревом с корнем a;

2) если a – некоторый узел, а T_1, T_1, …, T_m, $m \geq 0$, – упорядоченные деревья, то $T = (a; (T_1, T_1, \ldots, T_m))$ – упорядоченное дерево с корнем a.

Упорядоченные деревья – одна из наиболее распространенных информационных и алгоритмических структур.

Например, дерево на рис. 8 задает алгоритм вычисления значения арифметического выражения $(1+2) - (3/4)^2$.

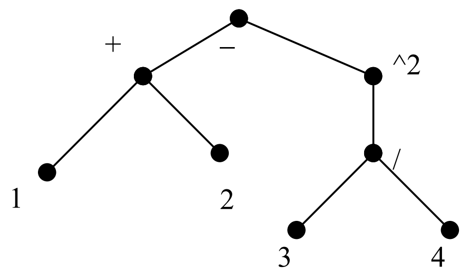


Рис. 8

14. 4. Бинарные деревья

Бинарными называют упорядоченные деревья, в которых каждый узел имеет не более двух, непосредственно следующих за ним узлов (рис. 9).

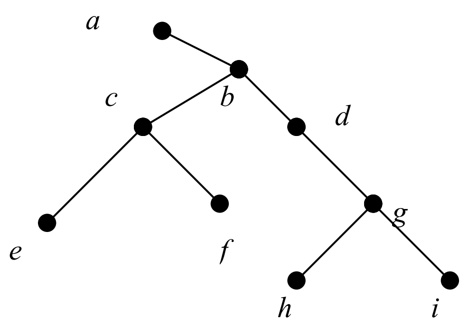


Рис. 9

Бинарные деревья наиболее часто используются для представления информации, обрабатываемой с помощью вычислительных машин. Рекурсивно бинарное дерево можно определить как конечное множество узлов, которое или пусто, или состоит из корня и двух не имеющих общих узлов бинарных деревьев – левого и правого.

Приписывая дуге, ведущей из произвольного узла к левому дереву, символ «0», а дуге, ведущей к правому дереву, символ «1», можно любой путь из корневой вершины (а вместе с ним и конечную вершину этого пути) закодировать с помощью нулей и единиц.

Например, на рис. 9 корневая вершина a получает в качестве кода пустую последовательность $\varnothing$, а остальные вершины следующие коды:

$$b-1; \quad c-10; \quad e-100; \quad f-101;$$
$$d-11; \quad g-111; \quad h-1110; \quad i-1111.$$

Найдем *число различных бинарных деревьев с заданным числом вершин*.

Пусть b_n – число различных бинарных деревьев с n вершинами. Ясно, что $b_0 = 1$ (имеется ровно одно пустое бинарное дерево) и $b_1 = 1$.

При $n > 0$ всякое бинарное дерево с n вершинами имеет вид $(a;\ (T_0,\ T_1))$, где T_0 – бинарное дерево с $k < n$ вершинами, а T_1 – бинарное дерево с $n-k-1$ вершиной. Число способов расположить одно бинарное дерево с $k < n$ вершинами слева от корня, а другое бинарное дерево с $n - k - 1$ вершиной справа от корня равно $b_k b_{n-k-1}$. Следовательно, суммируя по всем $k < n$, получаем:

$$b_n = b_0 b_{n-1} + b_1 b_{n-2} + \ldots + b_{n-1} b_0. \tag{1}$$

В частности:

$$b_1 = b_0^2 = 1;\ b_2 = b_0 b_1 + b_1 b_0 = 2;$$

$$b_3 = b_0 b_2 + b_1 b_1 + b_2 b_0 = 5;\ \ldots .$$

Последние равенства показывают, что последовательность чисел (b_n) – это последовательность чисел Каталана (см. 11.4).

Для полноты изложения напомним коротко, как могут быть получены явные выражения в факториалах для чисел b_n.

Запишем производящую функцию:

$$B(z) = b_0 + b_1 z + b_2 z^2 + \ldots . \tag{2}$$

Основываясь на (1), получаем:

$$B(z)^2 = b_0^2 + (b_0 b_1 + b_1 b_0)z + (b_0 b_2 + b_1 b_1 + b_2 b_0)z^2\ \ldots =$$

$$= b_1 + b_2 z + b_3 z^2 + \ldots .$$

Умножая обе части последнего равенства на z и добавляя 1, приходим к соотношению

$$z \cdot B(z)^2 + 1 = B(z).$$

Таким образом, $y = B(z)$ удовлетворяет квадратному уравнению

$$z \cdot y^2 - y + 1 = 0.$$

Решая его, находим:

$$B(z) = \frac{1}{2z}\left(1 - \sqrt{1 - 4z}\right).$$

Теперь разложим правую часть полученного равенства, воспользовавшись биномом Ньютона:

$$B(z) = \sum_{n\geq 0} (-1)^k \binom{1/2}{n+1} \cdot 2^{2n+1} z^n .$$

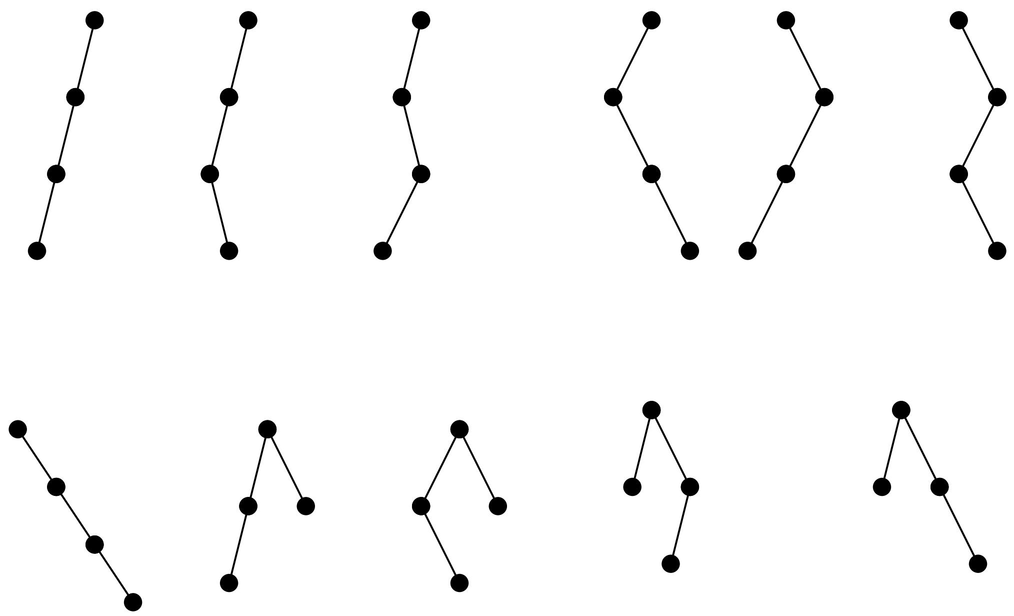
Сравнивая (1) и (2), приходим к заключению, что для всех n справедливо следующее равенство:

$$b_n = (-1)^k \binom{1/2}{n+1} \cdot 2^{2n+1} = \frac{1}{n+1} \binom{2n}{n}. \qquad (3)$$

Пример. Перечислим все бинарные деревья с 4 вершинами. В соответствии с формулой (3) имеем:

$$b_4 = \frac{1}{5} \binom{8}{4} = \frac{8 \cdot 7 \cdot 6 \cdot 5}{5 \cdot 4 \cdot 3 \cdot 2} = 14 .$$

На рис. 10, представлены четырнадцать бинарных деревьев с четырьмя вершинами.



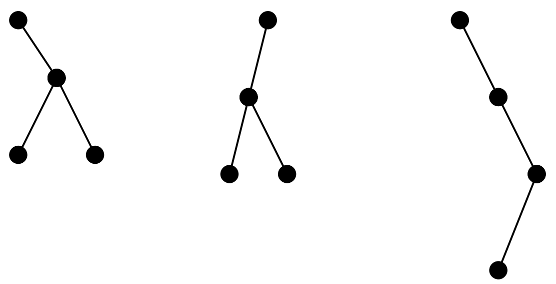


Рис. 10

Применим к (3) формулу Стирлинга. Получаем:

$$b_n \approx \frac{1}{n+1}\frac{\sqrt{4\pi n}\cdot(2n)^{2n}e^{-2n}}{\left(\sqrt{2\pi n}\cdot n^n e^{-n}\right)^2} = \frac{2^{2n}}{(n+1)\sqrt{\pi n}}. \qquad (4)$$

Формула (4) дает хорошее приближение даже при сравнительно малых значениях n. Например, при $n = 4$ имеем:

$$b_4 \approx \frac{2^8}{5\sqrt{4\pi}} \approx 14{,}4.$$

15. Доминирование. Внутренняя и внешняя устойчивость в графах

15. 1. Порядковая функция графа

Пусть $G = (V, E)$ – ориентированный граф. Наличие дуги, идущей из вершины u в вершину v, можно интерпретировать как доминирование u над v (в содержательных моделях – превосходство или предпочтительность в некотором смысле). Для вершины u обозначим через $G(u)$ множество всех вершин, над которыми u доминирует, т.е. множество концевых вершин дуг, начинающихся в вершине u. Если $U \subset V$ – некоторое множество вершин, обозначим через $G(U)$ объединение всех множеств $G(u)$, где $u \in U$. Таким образом, $v \in G(U)$ тогда и только тогда, когда над вершиной v доминирует хотя бы одна вершина из U.

Теорема. *В непустом ацикличном орграфе G имеется вершина, из которой не исходит ни одна дуга.*

Доказательство. Предположим противное. Это означает, что для любой вершины v множество $G(v)$ не пусто. Тогда, выбрав произвольно вершину v_0, можно построить сколь угодно длинную цепочку вершин v_0, $v_1 \in G(v_0)$, $v_2 \in G(v_1)$, … . Поскольку число вершин графа G конечно, в этой цепочке некоторые вершины будут повторяться. Но цепочка вида v_s, v_{s+1}, …, $v_t = v_s$ является циклом, что противоречит ацикличности графа G.

Условимся обозначать через $L_0(G)$ множество всех вершин графа G, из которых не исходит ни одна дуга. В соответствии с предыдущей теоремой $L_0(G)$ не пусто, если G – непустой ацикличный граф.

Обозначим через G^{-1} граф, полученный из графа G изменением направлений всех дуг на противоположные. Для любой вершины v множество $G^{-1}(v)$ содержит все те вершины, из которых в вершину v ведет некоторая дуга графа G. Множество $L_0(G^{-1})$ состоит из всех тех вершин, в которые не заходит ни одна дуга графа G.

В 13.6 было доказано, что отношение R на конечном множестве V ациклично тогда и только тогда, когда каждому элементу v множества V можно приписать натуральное число $\varphi(v)$ (оценку элемента v) так, что xRy влечет $\varphi(x) < \varphi(y)$ для любых $x,y \in V$. В этом пункте мы рассмотрим один из алгоритмов «оценивания» вершин ацикличного орграфа, при котором оценки согласуются с отношением доминирования: если вершина u доминирует над вершиной v, то оценка u будет выше, чем оценка v.

Предполагая, что граф G не содержит циклов, разобьем множество его вершин на так называемые *уровневые множества* (или просто уровни). Уровневые множества строятся рекурсивно:

$L_0 = \{ v \in V \mid G(v) = \varnothing \} = L_0(G);$

$L_1 = \{ v \in V \setminus L_0 \mid G(v) \subset L_0 \};$

$L_2 = \{ v \in V \setminus (L_0 \cup L_1) | \; G(v) \subset L_0 \cup L_1 \}$;

$L_3 = \{ v \in V \setminus (L_0 \cup L_1 \cup L_2) | \; G(v) \subset L_0 \cup L_1 \cup L_2 \}$;

…………………………………………………………….

Так как граф G ацикличен, множество L_0 не пусто. Обозначим через G' граф, полученный из графа G отбрасыванием вершин из L_0 и входящих в них дуг. Заметим, что $L_1 = L_0(G')$. Если $V_0 \neq V$, граф G' не пуст и не содержит циклов. Следовательно, множество $L_1 = L_0(G')$ не пусто.

Продолжаем аналогичным образом. Если

$$L_0 \cup L_1 \cup \ldots \cup L_{k-1} \neq V,$$

обозначим через $G^{(k)}$ граф, полученный из графа G отбрасыванием всех вершин из $L_0 \cup L_1 \cup \ldots \cup L_{k-1}$ и водящих в них дуг. Тогда $L_k = L_0(G^{(k)}) \neq \varnothing$. Так как множество вершин графа конечно, в ряду уровневых множеств L_0, L_1, …, найдется такое множество L_r, что $L_r \neq \varnothing$, а $L_{r+1} = \varnothing$. Все множества $L_0, L_1, \ldots, L_r$ не пустые, не пересекаются, а их объединение равно V.

Каждая вершина графа попадает в свое уровневое множество. При этом, если вершина u доминирует над вершиной v, то u окажется на уровне, имеющем больший номер, чем уровень, на котором находится v.

Пример. На рис. 1 изображен граф, вершины которого расположены на соответствующих уровнях.

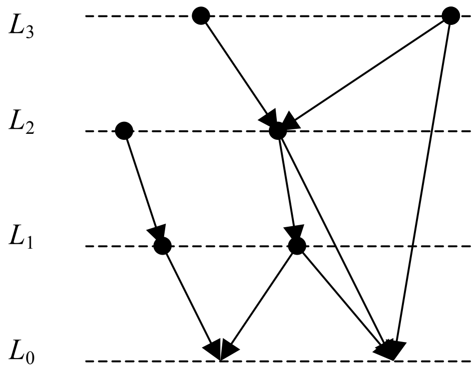


Рис. 1

Заметим, что предыдущий алгоритм построения уровневых множеств можно использовать для проверки ацикличности графа. Если граф G содержит циклы, то на некотором шаге возникнет ситуация, когда

$$L_0 \cup L_1 \cup \ldots \cup L_{k-1} \neq V, \text{ и } L_0(G^{(k)}) = \varnothing.$$

15. 2. Внешняя устойчивость

Пусть $G = (V, E)$ – ориентированный граф. Множество вершин $D \subset V$ называется *доминирующим*, если в каждую вершину, не принадлежащую D, ведет некоторая дуга из вершины, находящейся в D. Иными словами, множество D доминирующее, если $D \cup G(D) = V$. *Минимальным доминирующим* множеством называется такое доминирующее множество вершин D, никакое собственное подмножество которого не является доминирующим.

Доминирующее множество D является минимальным тогда и только тогда, когда для каждой вершины $d \in D$ выполняется одно из следующих двух условий:

1) $d \notin G(D)$;

2) имеется дуга, идущая из вершины d в некоторую вершину $d' \notin D$, причем никакая другая дуга с началом в одной из вершин множества D не ведет в d'.

В самом деле, если для вершины $d \in D$ выполняется одно из этих двух условий, множество $D \setminus \{d\}$ не является доминирующим.

Множество вершин $U \subset V$ называется *внешне устойчивым* в графе G, если U является доминирующим в графе G^{-1}. Таким образом, множество вершин U внешне устойчиво, если любая вершина v, не входящая в U, служит началом хотя бы одной дуги, конец которой находится в U.

Пример. Пусть G – граф, изображенный на рис. 2. Любое доминирующее множество графа G должно содержать вершину 2, поскольку в нее не заходит ни одна дуга. Вершина 2 вместе с любой другой вершиной образует минимальное доминирующее множество.

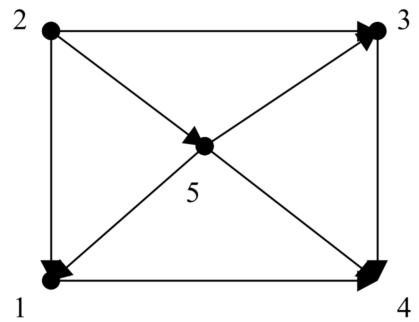


Рис. 2

Любое внешне устойчивое множество содержит вершину 4. Эта вершина вместе с любой другой составляет минимальное внешне устойчивое множество вершин графа *G*.

Пусть *G* – ориентированный граф. Перенумеровав его вершины, будем считать, что вершинами графа *G* служат числа 1, 2, …, *n*. Пусть $A = (a_{ij})$ – матрица смежности графа *G*. Тогда равенство $a_{ij} = 1$ означает, что на графе *G* имеется дуга из вершины *i* в вершину *j*. Каждому подмножеству *U* множества вершин графа *G* сопоставим его характеристический вектор (u_i) так, что $u_i = 1$ тогда и только тогда, когда $i \in U$.

Множество вершин *U* внешне устойчиво тогда и только тогда, когда для каждого $i = 1, 2, \ldots, n$ выполняется условие:

если $u_i = 0$, найдется $j \in \{1, 2,\ldots, n\}$ такое, что $u_j = 1$ и $a_{ij} = 1$.

Иными словами, множество вершин *U* внешне устойчиво, если булево выражение

$$\neg u_i \to (a_{i1}u_1 \vee a_{i2}u_2 \vee \ldots \vee a_{in}u_n) \qquad (1)$$

принимает значение 1 для всех $i = 1, 2, \ldots, n$. В равносильной форме (1) можно переписать так:

$$u_i \vee (\bigvee_{a_{ij}=1} u_j).$$

Положим

$$f(U) = \bigwedge_{i=1}^{n} (u_i \vee (\bigvee_{a_{ij}=1} u_j)) \qquad (2)$$

Множество *U* внешне устойчиво в том и только том случае, когда $f(U) = 1$. Таким образом, булева функция (2) является характеристической функцией семейства внешне устойчивых множеств в совокупности всех подмножеств множества вершин графа G.

Представим функцию (2) в виде ДНФ. Каждый дизъюнктивный член соответствует внешне устойчивому множеству. Если в ДНФ входит элементарная конъюнкция $u_i u_j \ldots u_k$, то на множестве вершин $U = \{i, j, \ldots, k\}$ функция f принимает значение 1 и, значит, множество U внешне устойчиво. Раскрыв в (2) скобки и проведя все сокращения в соответствии с тождествами

$$xx = x, x \vee x = x, x \vee (xy) = x, x(x \vee y) = x,$$

мы получим ДНФ, в которой дизъюнктивные члены соответствуют минимальным внешне устойчивым множествам.

Пример. Составим характеристическую функцию внешне устойчивых множеств (2) для графа G, представленного на рис.2:

$$f = (u_1 \vee u_4)(u_2 \vee u_1 \vee u_3 \vee u_5)(u_3 \vee u_4)u_4(u_5 \vee u_1 \vee u_3 \vee u_4).$$

После раскрытия скобок и упрощений получаем:

$$f = u_1u_4 \vee u_2u_4 \vee u_3u_4 \vee u_4u_5.$$

Следовательно, граф G обладает следующими минимальными внешне устойчивыми множествами вершин:

$$\{1, 4\}, \{2, 4\}, \{3, 4\}, \{4, 5\}.$$

15. 3. Внутренняя устойчивость

Пусть $G = (V, E)$ – ориентированный граф. Множество вершин $U \subset V$ называется *внутренне устойчивым*, если $U \cap D(U) = \varnothing$, т. е. в графе G не существует дуги, связывающей пару вершин из U. Внутренне устойчивое множество вершин $U \subset V$ является *максимальным*, если при добавлении к нему хотя бы одной вершины оно перестает быть внутренне устойчивым.

Пример. Рассмотрим граф G, изображенный на рис. 2. Поскольку граф не содержит петель, любое множество, содержащее одну вершину, является внутренне устойчивым. Максимальными внутренне устойчивыми являются множества вершин $\{2, 4\}$, $\{1, 3\}$ и $\{5\}$.

Пусть G – ориентированный граф. Так же, как и в предыдущем пункте, перенумеровав его вершины, будем считать, что вершинами графа G служат числа $1, 2, \ldots, n$. Пусть $A = (a_{ij})$ – матрица смежности графа G. Для множества вершин U пусть (u_i) – его характеристический вектор.

Множество вершин U внутренне устойчиво тогда и только тогда, когда для любых $i, j = 1, 2, \ldots, n$ выполняется условие:

$$\text{если } u_i = 1 \text{ и } a_{ij} = 1, \text{ то } u_j = 0.$$

Иными словами, множество вершин U внутренне устойчиво, если булево выражение

$$(u_i \wedge a_{ij}) \to \neg u_j \qquad (3)$$

принимает значение 1 для всех $i, j = 1, 2, \ldots, n$. В равносильной форме предыдущее условие можно представить так:

$$\neg u_i \vee \neg a_{ij} \vee \neg u_j = 1$$

для всех $i, j = 1, 2, \ldots, n$. Положим

$$f(U) = \bigwedge_{i,j=1}^{n} (\neg u_i \vee \neg a_{ij} \vee \neg u_j). \qquad (4)$$

Множество U внутренне устойчиво в том и только том случае, когда $f(U) = 1$. Таким образом, булева функция (4) является характеристической функцией семейства внутренне устойчивых множеств в совокупности всех подмножеств множества вершин графа G. Те множители (конъюнктивные члены), в которых $a_{i,j} = 0$, обращаются в единицу, поэтому их можно опустить.

С учетом этого получаем:

$$f(U) = \bigwedge_{a_{ij}=1}^{n} (\neg u_i \vee \neg u_j).$$

Представим функцию $f(U)$ в виде ДНФ. Каждый дизъюнктивный член соответствует внутренне устойчивому множеству. Если в ДНФ входит элементарная конъюнкция $u_i u_j \ldots u_k$, то на множестве вершин $\{i, j, \ldots, k\}$ функция f принимает значение 0 и, значит, дополнение этого множества внутренне устойчиво. Проведя все сокращения в соответствии с тождествами

$$xx = x,\ x \vee x = x,\ x \vee (xy) = x,\ x(x \vee y) = x,$$

мы получим ДНФ, в которой дизъюнктивные члены соответствуют максимальным внутренне устойчивым множествам.

Пример. Составим характеристическую функцию внутренне устойчивых множеств (2) для графа G, представленного на рис. 2:

$$f = (\neg u_1 \vee \neg u_4)(\neg u_2 \vee \neg u_1)(\neg u_2 \vee \neg u_3)(\neg u_2 \vee \neg u_5)\wedge$$

$$\wedge(\neg u_3 \vee \neg u_4)(\neg u_5 \vee \neg u_1)(\neg u_5 \vee \neg u_3)(\neg u_5 \vee \neg u_4).$$

После раскрытия скобок и упрощений получаем

$$f = \neg u_1 \neg u_2 \neg u_3 \neg u_4 \vee \neg u_1 \neg u_3 \neg u_5 \vee \neg u_2 \neg u_4 \neg u_5.$$

Следовательно, граф G обладает следующими максимальными внутренне устойчивыми множествами вершин:

$$\{5\},\ \{2, 4\},\ \{1, 3\}.$$

15. 4. Ядро графа

Пусть $G = (V, E)$ – ориентированный граф. Множество вершин $N \subset V$ называется *ядром*, если оно одновременно внешне и внутренне устойчиво.

Примеры. Граф на рис. 2 обладает единственным ядром $\{2, 4\}$. Граф на рис. 3 ядер не имеет.

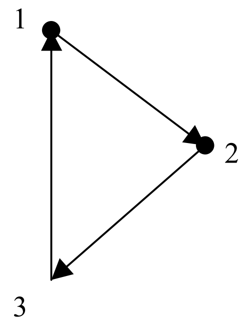


Рис. 3

Граф на рис. 4 имеет два ядра: {1, 3} и {2, 4}.

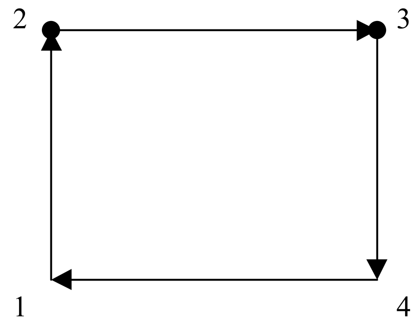


Рис. 4

Теорема. *Множество вершин ориентированного графа является ядром тогда и только тогда, когда оно максимальное внутренне устойчивое и минимальное внешне устойчивое множество.*

Предположим, что ориентированный граф $G = (V, E)$ представляет некоторое отношение предпочтения. Более точно: множество вершин V – это множество вариантов (альтернатив); наличие дуги из x в y означает, что y «лучше», чем x. Для ситуаций выбора типично, когда известно, что значит «лучше», но не известно, что значит «хорошо». Ядро, которое называют *решением задачи выбора по Нейману – Моргенштерну*, это в некотором смысле множество «хороших» вариантов. В соответствии с определением для каждого невыбранного варианта в ядре содержится вариант, который лучше него. Кроме того, если некоторый вариант попал в ядро, худшие, чем он, варианты в это ядро уже не попадут. В «простых» случаях ядро дает естественное решение задачи выбора. Так, если граф G не содержит циклов, а отношение предпочтения транзитивно, то единственным ядром графа G служит множество вершин уровня 0 (множество недоминируемых альтернатив). Ядро дает решение задачи выбора и в более сложных случаях.

Пример. Пусть имеются шесть вариантов $\{1, 2, 3, 4, 5, 6\}$, которые оцениваются по четырем критериям. Возможные оценки по каждому критерию: 0, 1, 2, 3, 4.

В следующей таблице приведены оценки вариантов.

Показатели / Варианты	П1	П2	П3	П4
1	3	3	0	2
2	3	1	1	3
3	2	1	2	1
4	4	2	1	1
5	2	3	1	3
6	2	2	0	4

Будем считать, что вариант *x* лучше варианта *y*, если *x* не содержит ни одной нулевой оценки и число показателей, по которым он лучше варианта *y*, больше, чем число показателей, по которым *y* лучше, чем *x*. На рис. 5 изображен соответствующий граф предпочтений.

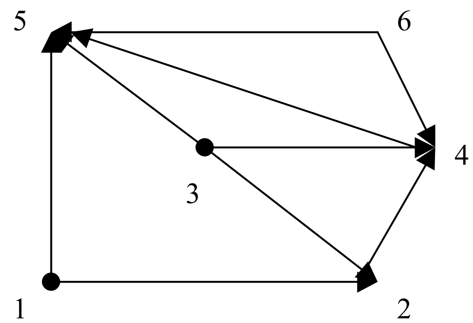


Рис. 5

Единственным ядром графа служит множество {2, 5}. Это и есть множество «хороших» вариантов. Чтобы сделать выбор между вторым и пятым вариантом, требуется привлечь дополнительную информацию.

Пример. Граф на рис. 6 обладает двумя ядрами: {1, 5, 7} и {2, 4, 6, 8}.

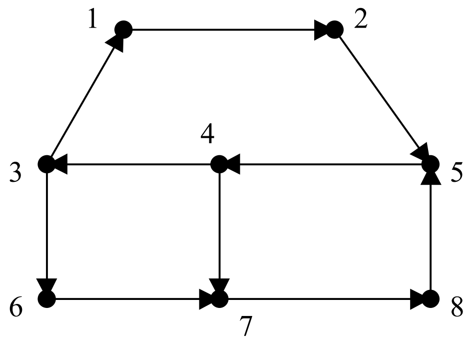


Рис. 6