Производящие функции и их использование в элементарной комбинаторике

1 Основные определения

1. Понятие производящей функции — это одно из основных понятий современной перечислительной комбинаторики. Задача этого пункта — дать вначале неформальное, а затем и вполне строгое определение этой функции.

1.1. Пусть вначале X — это некоторое конечное множество, элементы которого мы собираемся перечислять. Первое, что обычно хочется сделать — это указать мощность этого множества, т.е. количество всех элементов множества X. Однако, как правило, нам этой информации оказывается мало — на следующем шаге мы обычно хотим как-то классифицировать элементы этого множества.

1.2. Например, рассмотрим множество всех товаров на складе. Конечно, полезно знать общее их число. Однако гораздо полезнее разбить эти товары на категории или группы (например, бутылки вина, коробки мыла, пакеты с чаем), и подсчитать количество товаров в каждой группе по отдельности. При этом мы, конечно же, получим и общее число товаров, просто сложив количество товаров каждой категории.

В качестве еще одного примера можно привести гипотетическую задачу подсчета всех собак, живущих на Земле. Если пытаться практически решать эту задачу, то легче, видимо, начинать с подсчета собак данной породы, живущих в данном конкретном городе.

1.3. Рассмотрим теперь счетное множество X, например, множество всех графов. Здесь вообще ставить вопрос о количестве всех объектов данного множества бессмысленно. Однако и в этом случае задачу перечисления можно сделать вполне содержательной, и сделать это можно с помощью того же подхода. Именно, можно разбить множество X на конечные блоки по какому-то разумному признаку, и перечислять количество элементов в каждом блоке.

1.4. Например, бессмысленно перечислять все звезды во Вселенной. Однако перечислить все звезды в данной галактике — вполне разумная задача. Далее, столь же бессмысленно перечислять все графы. Однако перечисление всех различных графов на n вершинах есть также вполне содержательная и важная с практической точки зрения задача.

1.5. Как же чисто технически осуществить разбиение множества X на блоки? Наиболее популярный способ сделать это следующий: нужно приписать всем элементам x рассматриваемого множества некоторый вес так, чтобы элементы одного и того же блока имели одинаковый вес. Формально для этого следует ввести отображение

$$w : X \longrightarrow K$$

множества X в некоторое коммутативное кольцо K, при котором всем элементам x_i данного блока сопоставляется одно и то же значение $w(x_i) = k$. При этом все элементы X разбиваются на блоки мощности $c_k := |w^{-1}(k)|$:

$$X = \bigcup_{k \in K} \{x \mid w(x) = k \in K\}.$$

Заметим, что при таком подходе всему множеству X в кольце K отвечает вполне конкретный элемент
$$\sum_{x\in X} w(x) = \sum_{k\in K} c_k\cdot k, \qquad\qquad c_k = |w^{-1}(k)|,$$
называемый энумератором множества X.

1.6. `Пример.` Пусть на складе имеются три пакета с чаем, две бутылки вина и четыре коробки с мылом:
$$X = \{\text{чай}_1, \text{чай}_2, \text{чай}_3, \text{вино}_1, \text{вино}_2, \text{мыло}_1, \text{мыло}_2, \text{мыло}_3, \text{мыло}_4\}.$$
Введем теперь отображение $w : X \to \{\text{ч}, \text{в}, \text{м}\}$ следующим образом:
$$w(\text{чай}_i) = \text{ч}, \qquad w(\text{вино}_i) = \text{в}, \qquad w(\text{мыло}_i) = \text{м}.$$
Тогда все множество X разобьется на три блока, а энумератор этого множества примет вид
$$w(X) = 3\text{ч} + 2\text{в} + 4\text{м}.$$

1.7. Рассмотрим теперь частный случай отображения $w : X \to K$, при котором в качестве K выбирается множество $\mathbb{C}^{\mathbb{Z}_+}$ всех счетных последовательностей $(a_0, a_1, a_2, \ldots)$ комплексных чисел:
$$\mathbb{C}^{\mathbb{Z}_+} := \{(a_0, a_1, a_2, \ldots) \,|\, a_i \in \mathbb{C} \quad \forall\, i \in \mathbb{Z}_+\}.$$
На этом множестве можно ввести две базовые операции — сложения и умножения. Операция сложения числовых последовательностей определяется, как правило, однозначно:
$$(a_0, a_1, a_2, \ldots) + (b_0, b_1, b_2, \ldots) = (a_0 + b_0, a_1 + b_1, a_2 + b_2, \ldots). \tag{1}$$
Операцию же умножения
$$(a_0, a_1, a_2, \ldots) \cdot (b_0, b_1, b_2, \ldots) = (c_0, c_1, c_2, \ldots)$$
можно вводить по-разному. Чаще всего на практике используются следующие два способа умножения этих последовательностей: свертка
$$c_n = \sum_{i=0}^{n} a_i \cdot b_{n-i} \tag{2}$$
и биномиальная свертка
$$c_n = \sum_{i=0}^{n} \binom{n}{i} \cdot a_i \cdot b_{n-i} \tag{3}$$
числовых последовательностей. В обоих случаях множество $\mathbb{C}^{\mathbb{Z}_+}$ с введенными на нем операциями представляет собой коммутативное кольцо (точнее, коммутативную область целостности).

1.8. На практике с этими кольцами удобно работать как с кольцами $\mathbb{C}[[z]]$ и $\mathbb{C}_e[[z]]$ формальных степенных рядов. В таких кольцах для элемента $(0, 1, 0, 0, \ldots)$ вводится специальное обозначение
$$(0, 1, 0, 0, \ldots) =: z.$$
Рассмотрим вначале случай кольца $\mathbb{C}[[z]]$ с правилом умножения (2). Заметим, что
$$z^2 := z\cdot z = (0, 1, 0, 0, \ldots)\cdot(0, 1, 0, 0, \ldots) = (0\cdot 0, 0\cdot 1 + 1\cdot 0, 0\cdot 0 + 1\cdot 1 + 0\cdot 0, \ldots) = (0, 0, 1, 0, \ldots),$$

$$z^3 := z \cdot z \cdot z = (0,1,0,0,\ldots) \cdot (0,1,0,0,\ldots) \cdot (0,1,0,0,\ldots) = (0,1,0,0,\ldots) \cdot (0,0,1,0,\ldots) =$$
$$= (0 \cdot 0, 0 \cdot 0 + 1 \cdot 0, 0 \cdot 1 + 1 \cdot 0 + 0 \cdot 0, 0 \cdot 0 + 1 \cdot 1 + 0 \cdot 0 + 0 \cdot 0, \ldots) = (0,0,0,1,0,\ldots),$$

В общем случае, как несложно проверить по индукции,

$$z^n := \underbrace{z \cdot z \cdot \ldots \cdot z}_{n} = (0, \ldots, 0, 1, 0, \ldots),$$

где единица в последнем выражении стоит на n-м месте. Как следствие, любую последовательность чисел $(a_0, a_1, a_2, \ldots)$ можно записать в виде формального степенного ряда

$$(a_0, a_1, a_2, \ldots) = \sum_{n=0}^{+\infty} a_n z^n =: a(z) \in \mathbb{C}[[z]].$$

Аналогично, в случае кольца $\mathbb{C}_e[[z]]$ с правилом умножения (3) несложно убедиться в том, что

$$z^n := \underbrace{z \cdot z \cdot \ldots \cdot z}_{n} = (0, \ldots, 0, n!, 0, \ldots),$$

причем единственное ненулевое число $n!$ стоит в этой последовательности на n-м месте. Поэтому в кольце $\mathbb{C}_e[[z]]$ любая последовательность чисел $(a_0, a_1, a_2, \ldots)$ может быть записана так:

$$(a_0, a_1, a_2, \ldots) = \sum_{n=0}^{+\infty} a_n \frac{z^n}{n!} =: A(z) \in \mathbb{C}_e[[z]].$$

1.9. Вернемся теперь к перечисляемому нами множеству X и рассмотрим отображение

$$w : X \longrightarrow C[[z]],$$

которое любому элементу $x \in X$ сопоставляет некоторый формальный степенной ряд вида

$$w(x) = 0 \cdot 1 + 0 \cdot z + \ldots + 0 \cdot z^{k-1} + 1 \cdot z^k + 0 \cdot z^{k+1} + \ldots = z^k.$$

По сути дела, при таком подходе вес любого элемента $x \in X$ определяется степенью $k \in \mathbb{Z}_+$ особого элемента z кольца $C[[z]]$. Тогда суммирование образов $w(x)$ этого отображения по всем $x \in X$ даст нам энумератор вида

$$f(z) := \sum_{x \in X} w(x) = \sum_{n=0}^{+\infty} c_n z^n,$$

то есть формальный степенной ряд, коэффициенты c_n которого дают нам количество элементов $x \in X$, имеющих заданный вес $z^n = (0, 0, \ldots, 0, 1, 0, \ldots)$. Этот энумератор и называют *обыкновенной производящей функцией* множества X.

Аналогично, рассматривая отображение

$$w_e : X \longrightarrow C_e[[z]],$$

сопоставляющее любому элементу $x \in X$ формальный степенной ряд вида

$$w(x) = 0 \cdot 1 + 0 \cdot \frac{z}{1!} + \ldots + 0 \cdot \frac{z^{k-1}}{(k-1)!} + 1 \cdot \frac{z^k}{k!} + 0 \cdot \frac{z^{k+1}}{(k+1)!} + \ldots = \frac{z^k}{k!},$$

мы получаем в качестве энумератора X формальный степенной ряд

$$F(z) := \sum_{x\in X} w_e(x) = \sum_{n=0}^{+\infty} c_n \frac{z^n}{n!},$$

называемый *экспоненциальной производящей функцией* множества X.

1.10. Согласно определению, каждая производящая функция является элементом кольца формальных степенных рядов. Следовательно, любую пару производящих функций можно складывать и перемножать между собой.

Формальные правила сложения и умножения этих функций, естественно, совпадают с описанными выше правилами (1)–(3) сложения и умножения отвечающих этим рядам числовых последовательностей. Именно, пусть

$$f(z) = a_0 + a_1 z + a_2 z^2 + \ldots \qquad \text{и} \qquad g(z) = b_0 + b_1 z + b_2 z^2 + \ldots$$

есть пара обыкновенных производящих функций для множеств X и Y соответственно, а

$$F(z) = a_0 + a_1 \frac{z}{1!} + a_2 \frac{z^2}{2!} + \ldots \qquad \text{и} \qquad G(z) = b_0 + b_1 \frac{z}{1!} + b_2 \frac{z^2}{2!}$$

есть пара экспоненциальных производящих функций для этой пары множеств. Тогда суммой этих производящих функций называются формальные степенные ряды

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

и

$$F(z) + G(z) := (a_0 + b_0) + (a_1 + b_1)\frac{z}{1!} + (a_2 + b_2)\frac{z^2}{2!} + \ldots$$

Произведением производящих функций называются формальные степенные ряды

$$h(z) = f(z) \cdot g(z) := c_0 + c_1 z + c_2 z^2 + \ldots \qquad \text{и} \qquad H(z) = F(z) \cdot G(z) := c_0 + c_1 \frac{z}{1!} + c_2 \frac{z^2}{2!} + \ldots,$$

коэффициенты c_n которых вычисляются по формулам (2) и (3) соответственно.

Важно, что сложение и умножение любых двух производящих функций имеет и вполне определенный комбинаторный смысл. Выяснение этого смысла мы отложим до одного из следующих параграфов.

2. Итак, по определению, обыкновенная и экспоненциальная производящие функции являются элементами формальных степенных рядов $C[[z]]$ и $C_e[[z]]$. Поговорим немного подробнее об этих рядах.

2.1. Прежде всего, подчеркнем отличие этих формальных степенных рядов от рядов, встречающихся в математическом анализе. По сути дела, любой формальный степенной ряд — это некоторая картинка, удобная для изображения заданной числовой последовательности $(a_0, a_1, a_2, \ldots)$. Например, формальный степенной ряд

$$f(z) = \sum_{n=0}^{+\infty} n!\, z^n = 1 + z + 2z^2 + 6z^3 + 24z^4 + \ldots$$

можно рассматривать как обыкновенную производящую функцию, отвечающую числовой последовательности $(1, 1, 2, 4, 24, \ldots)$. Символы $1, z, z^2, z^3, \ldots$ в этой форме записи нужны лишь для того, чтобы указывать позиции, на которых эти элементы стоят. Например, если коэффициент при z^4 равен 24, то это означает лишь, что число 24 есть четвертый элемент рассматриваемой числовой последовательности.

2.2. Как следствие, в теории формальных степенных рядов символ z не надо рассматривать как комплексную переменную, и не надо вычислять значения производящих функций ни при каких конкретных значениях z. Наконец, в этой теории можно обойтись без понятия сходящихся или расходящихся рядов.

2.3. Несмотря на все эти различия, накопленный опыт работы с рядами из математического анализа часто бывает полезен и при анализе формальных степенных рядов. Связано это, прежде всего, с тем, что многие базовые операции над рядами, такие, как сложение или умножение, вводятся одинаково как для обычных, так и для формальных степенных рядов.

2.4. В качестве примера перемножим два формальных степенных ряда из $C[[z]]$:

$$f(z) = \sum_{n=0}^{+\infty} \alpha^n z^n \qquad \text{и} \qquad g(z) = 1 - \alpha z.$$

Коэффициент c_0 в произведении $h(z) = f(z) \cdot g(z)$, очевидно, равен $\alpha^0 \cdot 1 = 1$. Несложно убедиться, что остальные коэффициенты c_k равны нулю:

$$c_k = \alpha^k \cdot 1 - \alpha \cdot \alpha^{k-1} \equiv 0.$$

Следовательно,

$$f(z) \cdot g(z) = \Bigl(\sum_{n=0}^{+\infty} \alpha^n z^n\Bigr) \cdot (1 - \alpha z) = 1 \qquad \Longrightarrow \qquad f(z) = \sum_{n=0}^{+\infty} \alpha^n z^n = g^{-1}(z) = (1 - \alpha z)^{-1} =: \frac{1}{1 - \alpha z}.$$

Заметим, что формула

$$1 + \alpha z + \alpha^2 z^2 + \ldots = \frac{1}{1 - \alpha z} \qquad \forall\, \alpha \in \mathbb{Z} \tag{4}$$

есть хорошо известная формула для суммы геометрической прогрессии. В математическом анализе смысл этой формулы состоит в следующем: ряд, стоящий в левой ее части, сходится равномерно к функции, стоящей в ее правой части, при всех z, таких, что $|\alpha z| < 1$. В теории формальных степенных рядов ее смысл другой: формула означает, что функции $f(z)$ и $g(z)$ являются взаимно-обратными по отношению к операции умножения в кольце $C[[z]]$ формальных степенных рядов. Однако, несмотря на смысловые отличия, вид самого тождества и в том, и в другом случае одинаков.

2.5. Отмеченная в примере закономерность справедлива, как правило, и в общем случае. Именно, если имеется некоторое тождество со степенными рядами, которое выполняется, если эти ряды рассматривать как некоторые аналитические функции комплексного аргумента z, заданные в некоторой общей для них области определения (например, в некоторой малой окрестности нуля), то тогда высока вероятность того, что это же тождество остается верным как некоторое соотношение между формальными степенными рядами.

2.6. С этой точки зрения рассмотренный в примере результат можно было бы получить и так. Рассмотрим функции

$$f(z) = \sum_{n=0}^{+\infty} \alpha^n z^n \qquad \text{и} \qquad g(z) = 1 - \alpha z,$$

аналитические в области $|\alpha z| < 1$. В этой области для них справедливо тождество

$$f(z) \cdot g(z) = 1 \qquad \Longleftrightarrow \qquad \Big(\sum_{n=0}^{+\infty} \alpha^n z^n\Big) \cdot (1 - \alpha z) = 1.$$

Все, что теперь остается — это проверить, что данное тождество останется справедливым и как равенство между формальными степенными рядами в $C[[z]]$.

Далее мы рассмотрим еще несколько примеров из этой же серии.

2.7. `Пример 1.` Запишем следующее достаточно тривиальное в математическом анализе тождество, справедливое для всех $z \in Z$:

$$\exp(z) \cdot \exp(-z) = 1 \qquad \Longleftrightarrow \qquad \Big(\sum_{n=0}^{+\infty} 1 \cdot \frac{z^n}{n!}\Big) \cdot \Big(\sum_{n=0}^{+\infty} (-1) \cdot \frac{z^n}{n!}\Big) = 1.$$

Несложно убедиться, что это же равенство справедливо и в кольце $\mathbb{C}_e[[z]]$ формальных степенных рядов. Действительно, используя правило умножения (3), получаем хорошо известное и очень полезное тождество для биномиальных коэффициентов

$$\sum_{k=0}^{n} (-1)^k \binom{n}{k} = \begin{cases} 0, & \text{если } n > 0; \\ 1, & \text{если } n = 0. \end{cases}$$

2.8. `Пример 2.` Пусть $F(z)$, $G(z)$ – пара аналитических функций, связанных в некоторой общей области их определения тождеством вида

$$F(z) = G(z) \cdot e^z.$$

Очевидно, что тогда

$$G(z) = F(z) \cdot e^{-z}.$$

На языке формальных степенных рядов эти равенства отвечают формулам обращения

$$f_n = \sum_{k=0}^{n} \binom{n}{k} g_k \qquad \Longleftrightarrow \qquad g_n = \sum_{k=0}^{n} (-1)^{n-k} \binom{n}{k} f_k.$$

3. Вернемся к формуле (4). Как уже отмечалось, функции $f(z)$ и $g(z)$ в этой формуле удовлетворяют равенству $f(z) \cdot g(z) = 1$, т.е. являются взаимно-обратными по отношению к операции умножения в кольце формальных степенных рядов $\mathbb{C}[[z]]$. Возникает вопрос: что нужно от функции $g(z) \in \mathbb{C}[[z]]$ для того, чтобы она имела обратный элемент в $\mathbb{C}[[z]]$? Оказывается, для этого необходимо и достаточно, чтобы коэффициент b_0 при $z^0 = 1$ был отличен от нуля.

3.1. Действительно, пусть $g(z) = b_0 + b_1 z + b_2 z^2 + \dots$, $b_0 \neq 0$. Покажем, что в этом случае обязательно существует такая функция $f(z) = a_0 + a_1 z + a_2 z^2 + \dots$, такая, что $f(z) \cdot g(z) = g(z) \cdot f(z) = 1$.

3.2. Для этого перемножим два этих формальных степенных ряда и приравняем коэффициенты при одинаковых степенях z:

$$\begin{aligned}
&z^0: \quad && a_0 \cdot b_0 = 1 && \Longrightarrow \quad && a_0 = \frac{1}{b_0} \neq 0;\\
&z^1: \quad && a_0 \cdot b_1 + a_1 \cdot b_0 = 0 && \Longrightarrow \quad && a_1 = -\frac{1}{b_0} \cdot (a_0\, b_1);\\
&z^2: \quad && a_0 \cdot b_2 + a_1 \cdot b_1 + a_2 \cdot b_0 = 0 && \Longrightarrow \quad && a_2 = -\frac{1}{b_0} \cdot (a_0\, b_2 + a_1\, b_1);\\
& && && && \ldots
\end{aligned}$$

Видно, что таким образом шаг за шагом можно восстановить все коэффициенты a_n формального степенного ряда $f(z)$.

3.3. Пусть теперь $h(z) = c_0 + c_1 z + c_2 z^2 + \ldots$ и $g(z) = b_0 + b_1 z + b_2 z^2 + \ldots$ — два произвольных элемента кольца $\mathbb{C}[[z]]$, причем $b_0 \neq 0$. В этом случае можно рассматривать формальные дроби вида $h(z)/g(z)$ как элементы $f(z) \in \mathbb{C}[[z]]$, такие, что $f(z) \cdot g(z) = h(z)$. Говорят, что все эти элементы $f(z)$ образуют *поле частных* кольца $\mathbb{C}[[z]]$. По сути же дела, их можно рассматривать как результат *деления* формального степенного ряда $h(z)$ на формальный степенной ряд $g(z)$. Так, в рассмотренном в пункте 2.4 примере ряд

$$f(z) = 1 + \alpha z + \alpha^2 z^2 + \ldots$$

можно рассматривать как результат деления формальных степенных рядов $h(z) = 1$ и $g(z) = 1 - \alpha z$.

3.4. В качестве еще одного чрезвычайно полезного для дальнейшего изложения примера приведем следующее обобщение формулы (4):

$$\frac{1}{(1-\alpha z)^k} = \sum_{n=0}^{+\infty} \left(\!\!\binom{k}{n}\!\!\right) (\alpha\, x)^n. \tag{5}$$

Для строгой проверки этой формулы следует, вообще говоря, доказать, что умножение стоящего в правой части равенства (5) ряда на $(1-\alpha z)^k$ дает единицу для любого k. Более простой способ убедиться в справедливости (5) — это вспомнить про описанную в конце пункта 2 аналогию между формальными степенными рядами и аналитическими функциями комплексного аргумента z, и воспользоваться хорошо известной из классического анализа формулой бинома Ньютона

$$(1+z)^q = \sum_{n=0}^{+\infty} \frac{(q)_n}{n!} z^n = \sum_{n=0}^{+\infty} \frac{q\,(q-1)\,\ldots\,(q-n+1)}{n\,(n-1)\,\ldots\,1} z^n, \tag{6}$$

справедливой для любых $q, z \in \mathbb{C}$. Выбирая в (6) в качестве $q = -k$, $k \in \mathbb{N}$, и заменяя в ней z на αz, получаем следующую цепочку равенств, приводящую к формуле (5):

$$\frac{1}{(1-\alpha z)^k} = (1-\alpha z)^{-k} = \sum_{n=0}^{+\infty} \frac{(-k)_n}{n!} (-\alpha\, z)^n = \sum_{n=0}^{+\infty} \frac{(-k)\,(-k-1)\,\ldots\,(-k-(n-1))}{n\,(n-1)\,\ldots\,1} (-1)^n (\alpha\, z)^n =$$

$$= \sum_{n=0}^{+\infty} \binom{k+n-1}{n} (\alpha\, z)^n = \sum_{n=0}^{+\infty} \left(\!\!\binom{k}{n}\!\!\right) (\alpha\, x)^n.$$

Еще раз подчеркнем, что это не есть строгое доказательство формулы (5), а лишь наводящие соображения, позволяющие, зная формулу бинома Ньютона (6), быстро вывести формулу (5).

В частном случае $k=2$ и $k=3$ из (5) имеем следующие полезные формулы:

$$\frac{1}{(1-\alpha z)^2} = 1 + 2\,\alpha\, z + 3\,\alpha^2 z^2 + 4\,\alpha^3 z^3 + \ldots + (n+1)\alpha^n z^n + \ldots, \tag{7}$$

$$\frac{1}{(1-\alpha z)^3} = 1 + 3\,\alpha\, z + 6\,\alpha^2 z^2 + 10\,\alpha^3 z^3 + \ldots + \frac{(n+1)(n+2)}{2}\alpha^n z^n + \ldots. \tag{8}$$

3.5. *Замечание.* Все рассуждения, связанные с делением обыкновенных производящих функций как элементов кольца $C[[z]]$, легко переносятся и на экспоненциальные производящие функции, т.е. на элементы кольца $C_e[[z]]$.

4. Изучение производящих функций и операций над ними с точки зрения теории формальных степенных рядов удобно проводить параллельно с изучением рекуррентных соотношений. Этому и будут посвящены следующие несколько параграфов.

2 Рекуррентные соотношения

1. Начнем с простого примера. Популяция лягушек в озере увеличивается в четыре раза каждый год. В последний день каждого года 100 лягушек отлавливают и переправляют на другие озера. Предполагая, что в начале первого года в озере было 50 лягушек, найти количество лягушек в начале любого последующего года.

1.1. Обозначим через a_n количество лягушек в начале $(n{+}1)$-го года; $a_0 = 50$. Тогда, очевидно,

$$a_1 = 4\cdot 50 - 100 = 100, \qquad a_2 = 4\cdot 100 - 100 = 300,$$

а в общем случае

$$a_{n+1} = 4\cdot a_n - 100, \qquad n = 0, 1, 2, \ldots$$

Полученное равенство является простейшим примером *рекуррентного соотношения.*

1.2. *Определение.* Пусть $a_0, a_1, a_2, \ldots$ — произвольная числовая последовательность. Если для любого $n \geqslant m$ число a_{n+m} является некоторой функцией от m предыдущих членов последовательности, т.е.

$$a_{n+m} = f(a_n, a_{n+1}, \ldots, a_{n+m-1}), \tag{9}$$

то такая последовательность называется рекуррентной последовательностью, а соотношение (9) — рекуррентным соотношением m-го порядка.

1.3. В частном случае линейной функции f имеем так называемое линейное рекуррентное соотношение

$$a_{n+m} = b_1(n)\, a_{n+m-1} + b_2(n)\, a_{n+m-2} + \ldots + b_{m-1}(n)\, a_{n+1} + b_m(n)\, a_n + u_n. \tag{10}$$

В случае $u_n = 0$ оно называется однородным, в противном случае — неоднородным.

1.4. Самый простой случай рекуррентного соотношения – это линейное однородное рекуррентное соотношение с постоянными коэффициентами

$$a_{n+m} = b_1\, a_{n+m-1} + b_2\, a_{n+m-2} + \ldots + b_{m-1}\, a_{n+1} + b_m\, a_n. \tag{11}$$

1.5. Очевидно, что для однозначного определения всех a_n необходимо наряду с самим рекуррентным соотношением задать и первые m членов $a_0, a_1, \ldots, a_{m-1}$ данной последовательности, т.е., как говорят, задать начальные условия для рекуррентного соотношения.

1.6. Итак, наличие рекуррентного соотношения и начальных условий позволяет нам последовательно, шаг за шагом, определить любое наперед заданное количество n членов рекуррентной числовой последовательности. Однако иногда нам хочется получить явное аналитическое выражение для общего члена a_n этой последовательности, или, как говорят, *решить* данное рекуррентное соотношение. Это возможно далеко не всегда. Мы в данном параграфе покажем, что в случае линейного однородного рекуррентного соотношения с постоянными коэффициентами (11) такое решение можно построить всегда.

2. Прежде чем рассматривать общий случай уравнения (11), рассмотрим наиболее простой его вариант — линейное однородное рекуррентное соотношение первого порядка

$$a_{n+1} = b_1 \cdot a_n, \qquad n = 0, 1, 2, \ldots; \qquad a_0 \text{ — заданное число.} \tag{12}$$

2.1. Решение этого уравнения построить легко. Действительно,

$$a_1 = b_1 \cdot a_0; \qquad a_2 = b_1 \cdot a_1 = b_1^2 \cdot a_0; \qquad \ldots \qquad a_n = b_1^n \cdot a_0.$$

2.2. Предположим теперь, что нам кто-то сразу подсказал вид решения, а именно, что решение нашего уравнения (12) степенным образом зависит от n:

$$a_n = r^n \qquad \text{для некоторого } r\,.$$

Воспользовавшись этой подсказкой, подставим это выражение в исходное рекуррентное соотношение (12). В результате получается равенство вида

$$r^{n+1} = b_1 \cdot r^n,$$

из которого сразу же следует, что $r = b_1$; при этом a_n оказывается равным b_1^n. Это означает, что при $n = 0$ число $a_0 = b_1^0 = 1$. Иными словами, $a_n = b_1^n$ есть решение исходной задачи (12) в частном случае $a_0 = 1$. Такое решение называется *частным решением* уравнения (12).

2.3. Теперь попытаемся, зная частное решение, построить общее решение задачи (12). Для этого заметим, что любое частное решение уравнения (12), умноженное на произвольную постоянную c, по-прежнему этому уравнению удовлетворяет:

$$c \cdot b_1^{n+1} \equiv c \cdot b_1 \cdot b_1^n.$$

Решение же вида $c \cdot b_1^n$ позволяет удовлетворить любому начальному условию. Действительно, подставляя его в начальное условие для уравнения (12), получим:

$$c \cdot b_1^0 = a_0 \qquad \Longrightarrow \qquad c = a_0 \qquad \Longrightarrow \qquad a_n = a_0 \cdot b_1^n.$$

По этой причине решение вида $c \cdot b_1^n$ называется *общим решением* уравнения (12).

2.4. Подведем предварительные итоги. Предположив, что уравнение (12) степенным образом зависит от n, мы получили частное его решение вида $a_n = b_1^n$. Домножение этого решения на произвольную постоянную дало нам общее решение уравнения (12). Начальное условие позволило нам определить точное значение этой постоянной.

3. Рассмотрим теперь линейное однородное рекуррентное соотношение второго порядка

$$a_{n+2} = b_1 \cdot a_{n+1} + b_2 \cdot a_n, \qquad n = 0, 1, 2, \ldots, \qquad a_0\,,\ a_1\ \text{— заданные числа,} \tag{13}$$

и попытаемся применить к этому уравнению алгоритм, описанный в конце предыдущего пункта.

3.1. Для этого вновь предположим, что частное решение уравнения (13) имеет вид $a_n = r^n$. Подставляя это выражение в рекуррентное соотношение, получим

$$r^{n+2} = b_1 \cdot r^{n+1} + b_2 \cdot r^n \qquad \Longrightarrow \qquad r^2 - b_1\,r - b_2 = 0,$$

т.е. квадратное уравнение на r, любое решение r_0 которого есть некоторое частное решение уравнения (13). По этой причине данное уравнение называется характеристическим уравнением для рекуррентного соотношения (13).

3.2. Рассмотрим вначале случай, когда характеристическое уравнение имеет два различных вещественных корня r_1 и r_2. Покажем, что в таком случае выражение

$$a_n = c_1\,r_1^n + c_2\,r_2^n, \tag{14}$$

где c_1 и c_2 — произвольные постоянные, является *общим решением* соотношения (13) в том смысле, что любое решение (13) с заданными начальными условиями в нем содержится.

Действительно, покажем вначале, что выражение (14) действительно удовлетворяет рекуррентному соотношению (13):

$$c_1\,r_1^{n+2} + c_2\,r_2^{n+2} = b_1\,c_1\,r_1^{n+1} + b_1\,c_2\,r_2^{n+1} + b_2\,c_1\,r_1^n + b_2\,c_2\,r_2^n =$$

$$= c_1\,(r_1^2 - b_1\,r_1 - b_2) + c_2\,(r_1^2 - b_1\,r_1 - b_2) \equiv 0.$$

Покажем теперь, что это — действительно общее решение, т.е. что мы всегда можем подобрать константы c_1 и c_2 так, чтобы решение вида (14) удовлетворяло любым заданным начальным условиям. Для этого рассмотрим это выражение при $n = 0$ и $n = 1$:

$$\begin{aligned} a_0 &= c_1\,r_1^0 + c_2\,r_2^0 = c_1 + c_2, \\ a_1 &= c_1\,r_1^1 + c_2\,r_2^1 = c_1\,r_1 + c_2\,r_2. \end{aligned}$$

Эти выражения следует рассматривать как систему линейных уравнений для определения неизвестных постоянных c_1 и c_2. Определитель этой системы

$$\begin{vmatrix} 1 & 1 \\ r_1 & r_2 \end{vmatrix} = r_2 - r_1 \neq 0,$$

поэтому система всегда имеет единственное решение.

3.3. `Упражнение.` Показать, что в случае равных корней характеристического уравнения общее решение уравнения (13) имеет вид

$$a_n = c_1\,r^n + c_2\,n\,r^n,$$

а в случае комплексных корней это решение записывается в виде

$$a_n = c_1\,\rho^n\cos(n\,\vartheta) + c_2\,\rho^n\sin(n\,\vartheta),$$

где ρ и ϑ — модуль и аргумент одного из двух комплексных корней характеристического уравнения.

3.4. Пожалуй, самым известным и важным примером рекуррентного соотношения (13) является соотношение вида

$$F_{n+2} = F_{n+1} + F_n, \qquad n = 0, 1, 2, \dots, \qquad F_0 = 0,\ \ F_1 = 1,$$

определяющее так называемые числа Фибоначчи F_n. Характеристическое уравнение для этого рекуррентного уравнения имеет вид

$$r^2 = r + 1 \qquad \Longrightarrow \qquad r_{1,2} = \frac{1 \pm \sqrt{5}}{2} \qquad \Longrightarrow \qquad F_n = c_1 \left(\frac{1+\sqrt{5}}{2}\right)^n + c_2 \left(\frac{1-\sqrt{5}}{2}\right)^n.$$

Константы c_1 и c_2 определим из начальных условий:

$$\begin{aligned} &F_0 = 0 = c_1 + c_2 \\ &F_1 = 1 = c_1\,\frac{1+\sqrt{5}}{2} + c_2\,\frac{1-\sqrt{5}}{2} = c_1\left[\frac{1+\sqrt{5}-1+\sqrt{5}}{2}\right] = c_1\sqrt{5} \qquad \Longrightarrow \end{aligned}$$

$$\Longrightarrow \qquad c_1 = +\frac{1}{\sqrt{5}}, \qquad c_2 = -\frac{1}{\sqrt{5}}.$$

Таким образом, окончательно получаем следующее явное выражение для чисел Фибоначчи:

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

4. В принципе, описанный выше способ явного решения рекуррентных соотношений можно пытаться распространять и на соотношения более высокого порядка, и на неоднородные линейные рекуррентные соотношения. Вместо этого, однако, мы в следующем параграфе продемонстрируем более универсальный и удобный способ решения таких соотношений, использующий аппарат производящих функций.

3 Линейные рекуррентные соотношения с постоянными коэффициентами и обыкновенные производящие функции

1. Вернемся к простому примеру в начале предыдущего параграфа — задаче о лягушках. Для этого примера нами было получено следующее рекуррентное соотношение, описывающее изменение популяции лягушек в озере:

$$a_{n+1} = 4\,a_n - 100, \qquad n = 0, 1, 2, \dots, \qquad a_0 = 50. \tag{15}$$

Покажем, как решать это уравнение с помощью обыкновенных производящих функций.

1.1. Введем обыкновенную производящую функцию

$$f(z) = a_0 + a_1\, z + a_2\, z^2 + \ldots$$

для искомой числовой последовательности a_n, удовлетворяющей уравнению (15). Наша задача — отыскать явный вид этой производящей функции.

1.2. С этой целью домножим (15) на z^{n+1} и просуммируем полученное уравнение по n от 0 до $+\infty$:

$$\sum_{n=0}^{+\infty} a_{n+1}\, z^{n+1} = 4 \sum_{n=0}^{+\infty} a_n\, z^{n+1} - 100 \sum_{n=0}^{+\infty} z^{n+1}. \tag{16}$$

Левая часть этого равенства — это "почти" $f(z)$; переходя в этой сумме к новому индексу суммирования $k = n+1$, $k = 1, \ldots, +\infty$, получаем

$$\sum_{k=1}^{+\infty} a_k\, z^k = f(z) - a_0.$$

Первая сумма в правой части равенства (16), очевидно, равна

$$4 \sum_{n=0}^{+\infty} a_n\, z^{n+1} = 4\, z \sum_{n=0}^{+\infty} a_n\, z^n = 4\, z\, f(z).$$

Наконец, последнюю сумму в (16) можно свернуть и записать в виде дроби

$$\sum_{n=0}^{+\infty} z^{n+1} = z \sum_{n=0}^{+\infty} z^n = \frac{z}{1-z}.$$

Поэтому окончательно равенство (16) переписывается в виде

$$f(z) - a_0 = 4\, z\, f(z) - 100 \frac{z}{1-z} \qquad \Longrightarrow$$

$$\Longrightarrow \qquad f(z) = \frac{a_0}{1-4z} - \frac{100\, z}{(1-z)(1-4z)}. \tag{17}$$

1.3. Итак, мы построили производящую функцию для числовой последовательности a_n. Наша же исходная задача заключалась в отыскании явного выражения для этих чисел. Оказывается, что теперь это сделать очень просто — достаточно разложить левую и правую часть (17) в ряд по z, а затем сравнить члены при одинаковых степенях z.

С первым слагаемым в правой части (17) справиться легко — мы знаем, что

$$g(z) = \frac{1}{1-4z} = 1 + 4z + (4z)^2 + \ldots .$$

Для того, чтобы проделать ту же операцию со вторым слагаемым, нам предварительно необходимо разложить эту дробь на простейшие:

$$\frac{z}{(1-z)(1-4z)} = \frac{A}{1-z} + \frac{B}{1-4z} = \frac{A - 4Az + B - Bz}{(1-z)(1-4z)} \Longrightarrow \begin{cases} A + B = 0 \\ 4A + B = -1 \end{cases} \Longrightarrow \begin{cases} A = -1/3 \\ B = 1/3 \end{cases}$$

Как следствие,
$$-\frac{100\,z}{(1-z)(1-4z)} = -\frac{100}{3}\left[\sum_{n=0}^{+\infty}(4^n-1)\,z^n\right],$$
и мы окончательно для a_n получаем следующее явное аналитическое выражение:
$$a_n = a_0\,4^n - \frac{100}{3}\,[4^n-1] = \frac{50}{3}\,[4^n+2]\,, \qquad n=0,1,2,\ldots .$$

2. Итак, опишем общий алгоритм использования аппарата обыкновенных производящих функций для решения линейных рекуррентных соотношений m-го порядка с постоянными коэффициентами.

1) Для числовой последовательности $\{a_n\}$ вводится обыкновенная производящая функция $f(z)$.
2) Рекуррентное соотношение трансформируется в уравнение для $f(z)$ следующим образом: это соотношение домножается на z^{n+m}, суммируется по n от 0 до $+\infty$, а затем каждая из сумм выражается через $f(z)$.
3) Полученное уравнение формально разрешается относительно $f(z)$.
4) Числа a_n находятся как коэффициенты при z^n в разложении $f(z)$ по степеням z.

Применим этот алгоритм к линейному неоднородному рекуррентному соотношению m-го порядка с постоянными коэффициентами, записанному в следующем виде:
$$b_0a_{n+m}+b_1a_{n+m-1}+\ldots+b_ma_n=u_n,\quad n=0,1,2,\ldots,\quad b_0\neq 0,\ \ a_0,a_1,\ldots,a_{m-1}\ \text{– заданы.} \tag{18}$$

2.1. Введем для последовательности $a_0,a_1,a_2,\ldots$ обыкновенную производящую функцию
$$f(z)=a_0+a_1\,z+a_2\,z^2+\ldots=\sum_{n=0}^{+\infty}a_nz^n.$$

2.2. Домножим (18) на z^{n+m} и просуммируем полученное уравнение по n от 0 до $+\infty$. В полученном соотношении разберем каждое слагаемое отдельно:

$$b_0 \sum_{n=0}^{+\infty} a_{n+m} z^{n+m} = b_0 \left[f(z) - a_0 - a_1 z - \ldots - a_{m-1} z^{m-1}\right],$$

$$b_1 z \sum_{n=0}^{+\infty} a_{n+m-1} z^{n+m-1} = b_1 z \left[f(z) - a_0 - a_1 z - \ldots - a_{m-2} z^{m-2}\right],$$

$$\ldots$$

$$b_{m-2} z^{m-2} \sum_{n=0}^{+\infty} a_{n+2} z^{n+2} = b_{m-2} z^{m-2} \left[f(z) - a_0 - a_1 z\right],$$

$$b_{m-1} z^{m-1} \sum_{n=0}^{+\infty} a_{n+1} z^{n+1} = b_{m-1} z^{m-1} \left[f(z) - a_0\right],$$

$$b_m z^m \sum_{n=0}^{+\infty} a_n z^n = b_m z^m f(z),$$

$$\sum_{n=0}^{+\infty} u_n z^n =: z^m u(z).$$

Введем также следующие обозначения:

$$g(z) := b_0 + b_1 z + b_2 z^2 + \ldots + b_{m-1} z^{m-1} + b_m z^m,$$

$$c_n = \sum_{i=0}^{n} a_i b_{n-i}, \qquad h(z) := \sum_{n=0}^{m-1} c_n z^n.$$

С учетом этих обозначений и сделанных выше преобразований получается следующее уравнение на производящую функцию $f(z)$:

$$f(z) \cdot g(z) = b_0 a_0 + (b_0 a_1 + b_1 a_0) z + \ldots + (b_0 a_{n+m-1} + b_1 a_{n+m-2} + \ldots + b_{m-2} a_1 + b_{m-1} a_0) z^{m-1} + z^m u(z) =$$

$$= \sum_{n=0}^{m-1} c_n z^n + z^m u(z) = h(z) + z^m u(z) \qquad \Longrightarrow$$

$$\Longrightarrow \qquad f(z) = \frac{h(z) + z^m u(z)}{g(z)}.$$

Заметим, что деление на $g(z) = b_0 + b_1 z + \ldots + b_m z^m$ законно — коэффициент b_0 отличен от нуля.

2.3. В случае *однородного* линейного рекуррентного соотношения с постоянными коэффициентами m-го порядка производящая функция для рекуррентной последовательности $a_0, a_1, a_2, \ldots$ представляет собой рациональную функцию

$$f(z) = \frac{h(z)}{g(z)} = \frac{\sum\limits_{n=0}^{m-1} c_n z^n}{\sum\limits_{n=0}^{m} b_n z^n}.$$

По сути, этим устанавливается взаимно-однозначное соответствие между линейными однородными рекуррентными соотношениями на коэффициенты a_n и рациональными производящими функциями.

2.4. Как правило, при вычислении явного вида коэффициентов a_n вместо формального деления степенных рядов целесообразно разложить дробь на простейшие — элементарные дроби вида

$$\frac{A}{1-\alpha\, z}, \qquad \frac{B}{(1-\alpha\, z)^2}, \qquad \dots \qquad \frac{D}{(1-\alpha\, z)^m},$$

а затем воспользоваться готовыми формулами вида (5).

4 Комбинаторный смысл сложения и умножения производящих функций

1. Пусть теперь X , Y — пара конечных или счетных множеств, и пусть каждому из этих множеств поставлены в соответствие обыкновенные ($f(z)$ и $g(z)$) или экспоненциальные ($F(z)$ и $G(z)$) производящие функции. Так как производящие функции являются элементами кольца формальных степенных рядов, то их можно складывать и перемножать между собой. Эти формальные операции имеют вполне определенный комбинаторный смысл, к изучению которого мы и перейдем.