

Лекция 2: перечислительная комбинаторика

Дискретная математика, ВШЭ, факультет компьютерных наук

(Осень 2014 – весна 2015)

Задачи перечислительной комбинаторики имеют типовой вид: «сколько способов сделать то-то и то-то». Сегодня мы рассмотрим основные примеры.

Но начнем с двух важных принципов подсчета.

Формула суммы. *Если есть несколько перечислительных задач со взаимно исключающими вариантами, то общее количество способов выбрать вариант для одной из задач равно сумме числа вариантов для каждой задачи.*

Здесь важно, что варианты для разных задач взаимно исключающие. Скажем, если в аудитории 50 юношей и 50 девушек, то общее количество людей в аудитории равно 100. Но если в аудитории 50 читателей Германа Гессе и 50 читателей Дарьи Донцовой, то общее количество людей в аудитории может быть и меньше 100 (кто-то может читать обоих авторов).

Второй принцип выглядит похожим образом.

Формула произведения. *Если есть несколько перечислительных задач, то общее количество способов выбрать вариант для каждой из задач равно произведению числа вариантов для каждой задачи.*

Для двух задач это представляется в виде наглядной картинки. Составим таблицу, строки в которой перечисляют варианты для одной задачи, а столбцы — для другой. Тогда каждая ячейка таблицы отвечает выбору одного из вариантов для первой задачи и одного из вариантов для второй. Общее количество ячеек в таблице как раз произведение числа строк на число столбцов.

Чтобы доказать формулу произведения в общем случае, удобно применить индукцию.

1 Количество слов в алфавите. Двоичные слова, подмножества

Слово в алфавите A — это последовательность символов из этого алфавита.

Задача 1. Сколько есть слов длины k в алфавите из n символов?

Это легкий пример использования формулы произведения. Для каждой из k позиций в слове есть n вариантов. Нужно выбрать один из них для каждой позиции (какой-то символ на этой позиции стоит). Получаем ответ

$$n^k.$$

Если $n = 2$, слова называются двоичными (а алфавит стандартно составляется из нулей и единиц). Итак, всего двоичных слов длины k ровно 2^k штук.

Оказывается, столько же есть подмножеств в k -элементном множестве. Это число является решением такой задачи.

Задача 2. Сколькими способами можно похвалить каких-то человек из группы в k человек?

Те, кого похвалили, как раз и образуют подмножество.

Почему подмножеств 2^k ? Вместо того, чтобы решать задачу заново, мы докажем, что ответ в ней тот же, что и в задаче про слова.

Для этого установим взаимно однозначное соответствие между словами и подмножествами. Составим по подмножеству двоичное слово, у которого в позициях из подмножества стоят единицы, а в остальных позициях нули.

Ясно, что это соответствие однозначно и, более того, по слову однозначно восстанавливается соответствующее ему подмножество.

2 Упорядоченные выборки

Пусть есть n объектов. Нас интересует, сколькими способами можно составить список k из них, в котором не будет повторений. Начнем с примеров.

Задача 3. Сколько способов выбрать командира в отряде из 30 человек?

Ответ очевиден: 30.

Задача 4. Сколько способов выбрать командира и политрука в отряде из 30 человек?

Тут рассуждаем так: способов выбора командира 30. Для каждого из них есть 29 способов выбрать из оставшихся людей политрука. Всего получаем $30 \cdot 29 = 870$ вариантов.

Фактически в последней задаче мы сделали индуктивный переход. Теперь уже можно доказать общий факт.

Теорема 1. *Способов составить список без повторений длины k , в котором содержатся какие-то объекты из числа n возможных, ровно*

$$(n)_k = n \cdot (n-1) \cdot \dots \cdot (n-k+1)$$

штук.

Выражение $(n)_k$ часто называют нижней степенью.

Доказательство. Индукция по k . Случай $k=1$ аналогичен задаче 3.

Индуктивный переход (аналогично задаче 4). По предположению индукции вариантов начала списка длины $k-1$ есть ровно $(n)_{k-1}$ штук. Для каждого из них есть $n-(k-1) = n-k+1$ вариантов закончить список (добавляя любой из оставшихся объектов). Применим формулу произведения и получим $(n)_{k-1} \cdot (n-k+1) = (n)_k$. $\square$

Есть важный частный случай, когда список должен содержать все объекты ровно по разу.

Задача 5. Сколькими способами можно выстроить n человек в очередь?

Ответ уже получен. Это

$$(n)_n = n \cdot \dots \cdot 1 = n!$$

(читается «эн факториал», удобно считать, что $0! = 1$).

Упорядоченный список из n чисел, содержащих каждое из них ровно по разу, в математике называется перестановкой. Факториал как раз перечисляет количество перестановок.

Заметьте, что $(n)_k = 0$ при $k > n$. Это согласуется с нашими рассуждениями: в списке длины 10 из 9 объектов какой-то обязан повториться.

3 Неупорядоченные выборки. Сочетания

Сочетанием из n объектов по k в комбинаторике по традиции называется k -элементное подмножество n -элементного множества. Скажем, 1, 3, 5 — сочетание из 10 цифр по 3.

Теорема 2. *Количество k -элементных подмножеств n -элементного множества равно*

$$\frac{(n)_k}{k!} = \frac{n!}{k!(n-k)!}.$$

Доказательство. Установим соответствие между списками и подмножествами.

Каждому списку без повторений длины k соответствует подмножество — те объекты, которые вошли в список. А сколько списков отвечают одному и тому же подмножеству? Это в точности количество способов упорядочить k элементов. Таким образом, количество подмножеств в $k!$ раз меньше количества списков. $\square$

Взаимно однозначное соответствие между двоичными словами и подмножествами сразу дает решение такой задачи

Задача 6. Сколько есть двоичных слов длины n , содержащих k единиц?

4 Бином

Одно из открытий в перечислительной комбинаторике состоит в том, что считать нужно не числами, а многочленами (быть может, многочленами бесконечной степени — степенными рядами).

Применим эту идею к подсчету количества подмножеств с фиксированным числом элементов. Ответ дается многочленом $(1+x)^n$. Что это значит? Раскроем скобки и приведем подобные. Получим многочлен в стандартном виде

$$(1+x)^n = \sum_{k=0}^n \binom{n}{k} x^k. \quad (1)$$

Коэффициенты в правой части разложения называются биномиальными коэффициентами. Биномиальный коэффициент $\binom{n}{k}$ в точности равен количеству k -элементных подмножеств n -элементного множества. Действительно, найдем количество способов получить при раскрытии скобок (и до приведения подобных) одночлен степени k . Из каждой скобки можно выбрать либо 1, либо x . Выбрав один из вариантов для каждой скобки, получим слово в алфавите из символов 1, x . Степень полученного многочлена равна количеству символов x в слове. Итак, получаем формулу

$$\binom{n}{k} = \frac{(n)_k}{k!} = \frac{n!}{(n-k)!k!}.$$

Из последнего выражения можно заметить, что

$$\binom{n}{k} = \binom{n}{n-k}.$$

Однако есть и прямой комбинаторный способ проверить это равенство. Каждому двоичному слову взаимно однозначно соответствует инвертированное слово (ноль заменяем на единицу, единицу на ноль). При таком соответствии словам с k единицами соответствуют слова с $n-k$ единицами.

Хотя переменная x в формуле бинома имеет служебный смысл (степени x используются, чтобы представить всю последовательность биномиальных коэффициентов), формула (1) задает равенство многочленов. Подставляя вместо x число, получаем верные числовые равенства.

Например, при $x = 1$ получаем 2^n — количество подмножеств n -элементного множества. Если подставить $x = -1$, то окажется, что знакопеременная сумма биномиальных коэффициентов равна 0. Другими словами, количество подмножеств с четным числом элементов равно количеству множеств с четным числом элементов. Это довольно просто понять и без вычислений, если рассматриваются подмножества множества с нечетным числом элементов. А вот для подмножеств множества с четным числом элементов прямое комбинаторное объяснение этого равенства затруднительно.

5 Треугольник Паскаля. Рекуррентные соотношения

Выпишем коэффициенты биномов в виде треугольника:

$$\begin{array}{ccccccc} & & & & 1 & & \\ & & & 1 & & 1 & \\ & & 1 & & 2 & & 1 \\ 1 & & 3 & & 3 & & 1 \\ \dots & & \dots & & \dots & & \dots \end{array}$$

Полученная таблица называется треугольником Паскаля. Договоримся считать, что слева и справа от написанных чисел подразумеваются нули. Тогда можно заметить, что каждое число в таблице, за исключением самой верхней единицы, равно сумме двух чисел в предыдущей строчке, стоящих слева и справа от данного числа.

Это свойство будет выполняться и дальше. Записать это наблюдение можно в виде равенства

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}. \quad (2)$$

Почему это равенство всегда верно? Есть простое комбинаторное рассуждение. Выделим среди n объектов один, назовем его нулем. Подмножества из k элементов разбиваются на две группы: не содержащие 0 и содержащие 0. Первых столько же, сколько k -элементных подмножеств в $(n-1)$ -элементном множестве, т.е. $\binom{n-1}{k}$. А вторых столько же, сколько $(k-1)$ -элементных подмножеств в $(n-1)$ -элементном множестве, т.е. $\binom{n-1}{k-1}$. Из формулы суммы получаем равенство (2).

Есть и другой, алгебраический, способ получить равенство (2). Нужно записать равенство

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

и раскрыть скобки в левой и правой частях. При каждой степени получаться в точности равенства (2).

Равенства (2) называются рекуррентными соотношениями. Они выражают некоторые числа в числовой последовательности (в данном случае с двумя индексами) через числа в той же последовательности, но с меньшими значениями. Это позволяет определить всю последовательность, если добавить начальные условия. Например, положить

$$\binom{n}{0} = 1; \quad \binom{n}{n} = 1.$$

Треугольник Паскаля подсказывает еще один способ понимания биномиальных коэффициентов. Чтобы было привычнее, развернем треугольник так, чтобы числа заполняли положительный ортант:

$$\begin{array}{ccccc} 1 & 5 & 15 & 35 & 70 \\ 1 & 4 & 10 & 20 & 35 \\ 1 & 3 & 6 & 10 & 15 \\ 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 1 & 1 & 1 \end{array}$$

Теперь рекуррентное соотношение говорит, что каждое число равно сумме соседей слева и снизу. Эти числа оказываются равными числу монотонных путей из начала координат в данную точку (каждый шаг пути либо влево, либо вправо на 1).

Теорема 3. *Количество монотонных путей из $(0, 0)$ в (a, b) равно*

$$\binom{a+b}{a}.$$

Доказать это утверждение можно двумя способами. Первый состоит в том, чтобы применить индукцию и рекуррентное соотношение для биномиальных коэффициентов. Индуктивное утверждение состоит в том, что теорема справедлива для заданного $a + b$ и любого a .

Индуктивный переход основан на формуле суммы: попасть в точку (a, b) за один шаг возможно лишь из точки $(a - 1, b)$, $(a, b - 1)$, у которых сумма координат меньше. Далее применяем формулу суммы и рекуррентное соотношение (2).

Второй способ состоит в том, чтобы записать каждый путь в виде слова, буквой x отмечая шаг вправо, а буквой y — шаг вверх. Путь заканчивается в точке (a, b) тогда и только тогда, когда в слове ровно a букв x и ровно b букв y .

6 Числа Фибоначчи

Последовательность чисел Фибоначчи определяется рекуррентно:

$$F_1 = F_2 = 1; \quad F_{n+2} = F_{n+1} + F_n.$$

Начало последовательности выглядит так:

$$1, 1, 2, 3, 5, 8, 13, \dots$$

Через числа Фибоначчи выражаются решения многих комбинаторных задач. Рассмотрим пример.

Задача 7. Сколько двоичных слов длины n , в которых нет двух нулей подряд?

Ответ: F_{n+2} . Легко видеть, что для $n = 1$ получается $2 = F_3$ таких слова, а для $n = 2$ получается $3 = F_4$ таких слова (запрещено только слово 00).

Далее мы покажем справедливость того же рекуррентного соотношения для ответа в задаче. Отсюда по индукции будет следовать справедливость указанного ответа.

Чтобы вывести рекуррентное соотношение, обозначим через a_n количество слов длины n без двух нулей подряд, которые заканчиваются на 0, а через b_n — количество таких слов, которые заканчиваются на 1.

Слово без двух нулей подряд, которое заканчивается на нуль, можно получить дописыванием нуля только к слову без двух нулей подряд, которое заканчивается на 1, т.е.

$$a_{n+1} = b_n.$$

А слово без двух нулей подряд, которое заканчивается на единицу, можно получить приписыванием единицы к любому слову без двух нулей подряд, т.е.

$$b_{n+1} = a_n + b_n.$$

Теперь запишем рекуррентное соотношение для суммы $a_n + b_n$:

$$(a_{n+2} + b_{n+2}) = b_{n+1} + a_{n+1} + b_{n+1} = (a_{n+2} + b_{n+2}) = (a_{n+1} + b_{n+1}) + (a_n + b_n).$$

Это и есть соотношение для чисел Фибоначчи.

Для чисел Фибоначчи можно написать явную формулу:

$$F_n = \frac{1}{\sqrt{5}}(q^n - q^{-n}), \quad \text{где } q = \frac{1 + \sqrt{5}}{2},$$

т.е. корень уравнения $q^2 - q - 1 = 0$ (так что q^{-1} — тоже корень). Доказать эту формулу нетрудно по индукции.

7 Числа Каталана

Эти числа 1, 2, 5, 14, 42, ... встречаются в ответах комбинаторных задач еще чаще. Для чисел Каталана C_n тоже есть рекуррентное соотношение, но оно нам не понадобится. Поэтому зададим C_n как количество последовательностей, состоящих из n штук $+1$ и n штук -1 , для которых сумма любого начального отрезка последовательности неотрицательна (для краткости будем называть такие последовательности неотрицательными). Например:

$$\begin{aligned} n = 1 : & \quad (+1, -1), \\ n = 2 : & \quad (+1, +1, -1, -1), \quad (+1, -1, +1, -1), \\ n = 3 : & \quad (+1, +1, +1, -1, -1, -1), \quad (+1, +1, -1, -1, +1, -1), \\ & \quad (+1, +1, -1, +1, -1, -1), \quad (+1, -1, +1, +1, -1, -1), \\ & \quad (+1, -1, +1, -1, +1, -1). \end{aligned}$$

Главный пример, в котором возникают числа Каталана, это количество правильных скобочных выражений. Когда нужно указать точный порядок выполнения арифметических действий, расставляются скобки. Если убрать остальную часть формулы, то и получится правильное скобочное выражение.

Формально правильные скобочные выражения — это слова из открывающей и закрывающей скобки, которые определяются рекурсивно: $()$ — правильное скобочное выражение из 1 пары скобок, и любое правильное скобочное выражение из $n + 1$ пары скобок имеет вид

$$(S_1)S_2,$$

где S_1 — правильное скобочное выражение из k пар скобок, а S_2 — правильное скобочное выражение из $n - k$ пар скобок, $k \geq 0$. (Отсюда и получается рекуррентное соотношение для числа правильных скобочных выражений.)

Теорема 4. *Количество правильных скобочных выражений из n пар скобок равно C_n .*

Доказательство. Для позиции в скобочном выражении определим скобочный итог как разность между количеством открывающих и закрывающих скобок, предшествующих этой позиции.

По индукции доказывается, что в правильном скобочном выражении скобочный итог неотрицателен в любой позиции, а в конце слова равен 0. Поэтому замена открывающей скобки на $+1$, а закрывающей на -1 превращает правильное скобочное выражение в неотрицательную последовательность из ± 1 .

Верно и обратное: замена в неотрицательной последовательности ± 1 на открывающую и закрывающую скобки дает правильное скобочное выражение. Доказательство индукцией по n .

Возьмем первую $+1$ и найдем ближайшую к ней -1 такую, что сумма становится равной 0 (должна быть, так как общая сумма равна 0). Между этой парой стоит неотрицательная последовательность (так как мы выбрали ближайшую -1) и после выбранной -1 также стоит неотрицательная последовательность. Применяя предположение индукции, видим, что это соответствует определению правильного скобочного выражения. $\square$

Задача 8. Найдите количество монотонных путей из $(0, 0)$ в (n, n) , которые не опускаются ниже биссектрисы координатного квадранта. (Т.е. проходят только через точки (x, y) с $x \leq y$.)

Ответ: C_n . Шагу вверх сопоставим $+1$, шагу вправо сопоставим -1 . Шаг вверх увеличивает на 1 разность $y - x$, а шаг вправо уменьшает эту разность на 1. Поэтому монотонному пути в треугольнике отвечает неотрицательная последовательность и это соответствие взаимно однозначно.

Числа Каталана выражаются через биномиальные коэффициенты.

Теорема 5.
$$C_n = \frac{1}{n+1} \binom{2n}{n}.$$

Доказательство. Число Каталана C_n есть разность общего количества монотонных путей в (n, n) (их $\binom{2n}{n}$ штук) и количества тех путей, которые пересекают хотя бы в одной точке прямую $y = x - 1$.

Это число оказывается равным биномиальному коэффициенту $\binom{2n}{n+1}$, т.е. количеству путей в точку $(n+1, n-1)$. Установим между путями двух типов взаимно однозначное соответствие.

Для этого заметим, что пути обоих типов обязательно пересекают прямую $y = x - 1$. Выберем последнюю точку пересечения и отразим остаток пути относительно прямой $y = x - 1$. Путь первого типа перейдет в путь второго типа и наоборот.

Итак,

$$C_n = \binom{2n}{n} - \binom{2n}{n+1} = \binom{2n}{n} \left(1 - \frac{n}{n+1}\right) = \frac{1}{n+1} \binom{2n}{n},$$

что и требовалось доказать. $\square$

8 Сочетания с повторениями

В заключение рассмотрим еще одну перечислительную задачу, которая часто применяется при решении других задач.

Задача 9 (задача Муавра). Сколько есть решений уравнения $x_1 + x_2 + \dots + x_n = k$ в целых неотрицательных числах?

В комбинаторике это количество называется обычно числом сочетаний с повторениями из n элементов по k . В обычных сочетаниях элементы различны и количество сочетаний равно числу решений уравнения $x_1 + x_2 + \dots + x_n = k$, когда переменные принимают значения 0 или 1. Теперь элемент i разрешается использовать несколько раз, x_i — это его кратность в сочетании с повторениями.

Ответ на задачу Муавра:

$$\binom{n+k-1}{k}.$$

Для решения задачи годится «метод точек и перегородок». Расставим последовательно k точек и $n-1$ перегородку. По такой расстановке прочитаем решение задачи Муавра:

• || ••• | •

соответствует решению уравнения

$$x_1 = 1; x_2 = 0; x_3 = 3; x_4 = 1$$

при $k = 5$ и $n = 4$ (здесь нет ограничения $k \leq n$ как для сочетаний с повторениями).

Правило соответствия ясно из примера: количество точек до первой перегородки равно значению первой переменной, количество точек между первой и второй перегородками равно значению второй переменной и т.д. Количество точек после последней перегородки равно значению последней переменной.

Легко видеть, что это правило задает взаимно однозначное соответствие между решениями уравнения и словами из точек и перегородок, в которых k точек и $n-1$ перегородка.

Сочетания с повторениями возникают в элементарной комбинаторике, когда нужно подсчитывать количество способов разложить одинаковые предметы на несколько кучек.

Задача 10. Сколькими способами 5 разбойников могут разделить добычу в 6 рублей, если каждый должен получить целое число рублей?

Ответ: $\binom{5+6-1}{6} = 210$. Обозначим x_i количество рублей у i -го разбойника. Получаем задачу Муавра.