Выпуклые множества и функции

$\mathbb{R}^n$ — множество наборов из n вещественных чисел. Далее это множество будем называть *пространством*, его элементы — *точками*, точку с координатами $(x_1, \ldots, x_n)$ будем обозначать X. Покоординатно определим *сумму* точек $X + Y = Z$, так, что $z_i = x_i + y_i$ для всех $i = 1, \ldots, n$ и *произведение* вещественного числа и точки: $\lambda \cdot X = Z$, так что $\lambda \cdot x_i = z_i$ для всех i. Для пары точек X, Y *отрезком* XY будем называть множество всех точек вида $\lambda X + \mu Y$, где λ и μ пробегают все вещественные неотрицательные значения, удовлетворяющие $\lambda + \mu = 1$. $\mathbb{R}^n_{\geqslant 0}$ — множество точек пространства, все координаты которых неотрицательны.

Множество $\mathbb{A} \subseteq \mathbb{R}^n$ *выпукло*, если из того, что оно содержит точки X и Y, следует, что оно содержит весь отрезок XY. Очевидно, что $\mathbb{R}^n_{\geqslant 0}$ выпукло.

Функция $f : \mathbb{A} \to \mathbb{R}$, определённая на выпуклом множестве $\mathbb{A}$, называется *выпуклой*, если для любых точек X, Y из $\mathbb{A}$ и любых вещественных неотрицательных λ и μ, удовлетворяющих $\lambda + \mu = 1$, верно неравенство: $\lambda f(X) + \mu f(Y) \geqslant f(\lambda X + \mu Y)$. Геометрический смысл написанного: хорда, стягивающая любые две точки графика функции, лежит над графиком.

Теорема (*неравенство Йенсена*)**.** *Пусть $\mathbb{A}$ — выпуклое множество, f — выпуклая функция на нём, $X_1, \ldots, X_k$ — точки $\mathbb{A}$; $\lambda_1, \ldots, \lambda_k$ — вещественные неотрицательные числа, сумма которых* 1. *Тогда выполнено неравенство:*

$$\lambda_1 f(X_1) + \ldots + \lambda_k f(X_k) \geqslant f(\lambda_1 X_1 + \ldots + \lambda_k X_k).$$

► Докажем по индукции по k. База $k = 1$ очевидна. Докажем неравенство для параметра k в предположении, что для всех меньших значений параметра оно уже доказано. Выделим два последних слагаемых в левой части неравенства. Тогда, использовав определение выпуклости функции f для точек X_{k-1} и X_k и для значений параметров $\lambda = \frac{\lambda_{k-1}}{\lambda_{k-1}+\lambda_k}$ и $\mu = \frac{\lambda_k}{\lambda_{k-1}+\lambda_k}$ и переписав первые $k-2$ слагаемых без изменения, имеем:

$$\lambda_1 f(X_1) + \ldots + \lambda_{k-1} f(X_{k-1}) + \lambda_k f(X_k) \geqslant \lambda_1 f(X_1) + \ldots + (\lambda_{k-1} + \lambda_k) f(\frac{\lambda_{k-1}}{\lambda_{k-1} + \lambda_k} X_{k-1} + \frac{\lambda_k}{\lambda_{k-1} + \lambda_k} X_k).$$

Осталось только применить к правой части предположение индукции (неравенство для $k-1$ точки) для набора точек $X_1, \ldots, X_{k-2}, \frac{\lambda_{k-1}}{\lambda_{k-1}+\lambda_k} X_{k-1} + \frac{\lambda_k}{\lambda_{k-1}+\lambda_k} X_k$ и для коэффициентов $\lambda_1, \ldots, \lambda_{k-2}, \lambda_{k-1} + \lambda_k$.◄

Доказательство похоже на процедуру нахождения центра масс системы материальных точек.

Напомним, что для выпуклости дважды дифференцируемой функции, определённой на вещественной прямой, достаточно неотрицательности её второй производной во всех точках.

Для произвольного подмножества $\mathbb{M} \subseteq \mathbb{R}^n$ его *выпуклой оболочкой* назовём такое выпуклое множество $\mathbb{A}$, которое содержит $\mathbb{M}$ и содержится в любом выпуклом множестве, содержащем $\mathbb{M}$. Дальше будем обозначать выпуклую оболочку множества $\mathbb{M}$ как $conv(\mathbb{M})$.

Утверждение 1. *Выпуклая оболчка всегда существует и единственна.*

► Заметим, что пересечение любого семейства выпуклых множеств — выпуклое множество. Действительно, если это пересечение содержало точки X, Y, то каждое из множеств пересечения содержало эти точки $\Rightarrow$ каждое из множеств содержало весь отрезок $XY \Rightarrow$ этот отрезок лежит и в рассматриваемом пересечении.

Хотя бы одно выпуклое множество, содержащее $\mathbb{M}$, всегда существует — это само пространство. Теперь рассмотрим пересечение всех выпуклых множеств, содержащих $\mathbb{M}$. По построению это и есть выпуклая оболочка. Единственность очевидна: если существовало бы две выпуклые оболочки, то они, вследствие определения, должны были бы содержаться друг в друге $\Leftrightarrow$ совпадать.◄

Утверждение 2. $\mathbb{M} = \{X_1, \ldots, X_k\}$ — *конечный набор точек. Тогда $conv(M)$ — множество всех точек, представимых в виде $\lambda_1 X_1 + \ldots + \lambda_k X_k$ с некоторыми вещественными неотрицательными $\lambda_1, \ldots, \lambda_k$, дающими в сумме* 1.

► Обозначим множество всех точек, представимых в таком виде, буквой $\mathbb{A}$ и докажем, что $\mathbb{A} = conv(\mathbb{M})$. Покажем включения в обе стороны.

Убедимся, что $\mathbb{A}$ — выпуклое. Пусть оно содержит точки X и Y. Тогда для некоторых наборов коэффициентов λ_i и μ_i выполнены $X = \lambda_1 X_1 + \ldots + \lambda_k X_k$ и $Y = \mu_1 X_1 + \ldots + \mu_k X_k$. Сложим последние выражения, предварительно умножив на λ первое и на μ второе:

$$\lambda X + \mu Y = (\lambda\lambda_1 + \mu\mu_1)X_1 + \ldots + (\lambda\lambda_k + \mu\mu_k)X_k.$$

Получили, что для любых неотрицательных $\lambda + \mu = 1$ и для точки $\lambda X + \mu Y$ нашёлся набор неотрицательных коэффициентов в разложении по X_i с суммой $1 \Rightarrow \mathbb{A}$ содержит весь отрезок XY. Выпуклое. $\mathbb{A}$ содержит все X_i (достаточно положить λ_i равным 1, а остальные — 0) $\Rightarrow \mathbb{A}$ — некоторое выпуклое множество, содержащее $\mathbb{M} \Rightarrow \mathbb{A} \supseteq conv(\mathbb{M})$ по определению выпуклой оболочки.

Докажем, что $Z = \lambda_1 X_1 + \ldots + \lambda_k X_k \in conv(\mathbb{M})$ индукцией по числу ненулевых коэффициентов λ_i. Если такой коэффициент всего один, то тогда $Z = X_i \in conv(\mathbb{M})$. Шаг индукции: не умаляя общности, считаем $\lambda_{k-1} \neq 0$ и $\lambda_k \neq 0$. Рассмотрим точки X и Y, соответствующие наборам коэффициентов $\lambda_1, \ldots, \lambda_{k-2}, \lambda_{k-1} + \lambda_k, 0$ и $\lambda_1, \ldots, \lambda_{k-2}, 0, \lambda_{k-1} + \lambda_k$. По предположению индукции они лежат в $conv(\mathbb{M})$ (у них на один положительный коэффициент меньше). Но тогда и $Z = \frac{\lambda_{k-1}}{\lambda_{k-1}+\lambda_k} X + \frac{\lambda_k}{\lambda_{k-1}+\lambda_k} Y$ там лежит как точка отрезка XY. Таким образом, $\mathbb{A} \subseteq conv(\mathbb{M})$. ◄

Комментарий в сторону: в случае, когда $\mathbb{M}$ — аффинно независимы, т.е. являются вершинами симплекса, числа λ_i для каждой точки определяются однозначно и называются барицентрическими координатами этой точки. Геометрическая интерпретация второй части доказательства такая: сначала рисуем вершины симплекса, потом рёбра, затем грани и так далее; в конце рисуем внутренность.

Перестановки и симметрические многочлены

S_n — группа перестановок множества $\{1, \ldots, n\}$, т.е. множество всех обратимых отображений (функций) из $\{1, \ldots, n\}$ в себя. Она насчитывает $n!$ различных перестановок.

Однородным симметрическим многочленом, порождённым набором вещественных неотрицательных чисел $(x_1, \ldots, x_n)$ (который, напомню, мы обозначаем одной буквой X и называем точкой) назовём следующее выражение:

$$T_X(a_1, \ldots, a_n) := \sum_{\sigma \in S_n} a_1^{x_{\sigma(1)}} \cdot \ldots \cdot a_n^{x_{\sigma(n)}}.$$

Пример: $T_{(1,1,0)} = 2 \cdot (a_1 a_2 + a_2 a_3 + a_3 a_1)$.

Вообще говоря, многочленом это выражение может и не являться, так как показатели степеней могут быть нецелыми.

Для дальнейших доказательств нам потребуется отображение *перестановки координат*. Каждой перестановке $\sigma \in S_n$ сопоставим отображение $\widehat{\sigma} : \mathbb{R}^n \to \mathbb{R}^n$, действующее так:

$$\widehat{\sigma}(x_1, \ldots, x_n) = (x_{\sigma(1)}, \ldots, x_{\sigma(n)}).$$

Мажорирование последовательностей

Пусть даны два набора невозрастающих неотрицательных чисел $\overline{X} = (x_1, \ldots, x_n)$, $\overline{Y} = (y_1, \ldots, y_n)$ (чёрточку сверху будем ставить только над упорядоченными $x_1 \geqslant x_2 \geqslant \ldots$ наборами). Скажем, что набор $\overline{X}$ *мажорирует* набор $\overline{Y}$ (обозначение $\overline{X} \succ \overline{Y}$), если выполнен следующий набор соотношений:

$$
\begin{gathered}
x_1 \geqslant y_1; \\
x_1 + x_2 \geqslant y_1 + y_2; \\
x_1 + x_2 + x_3 \geqslant y_1 + y_2 + y_3; \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
x_1 + \ldots + x_{n-1} \geqslant y_1 + \ldots + y_{n-1}; \\
x_1 + \ldots + x_n = y_1 + \ldots + y_n.
\end{gathered}
$$

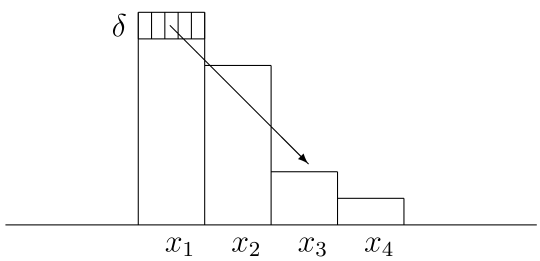
Легко показать, что если $\overline{X} \succ \overline{Y}$ и $\overline{Y} \succ \overline{Z}$, то $\overline{X} \succ \overline{Z}$.

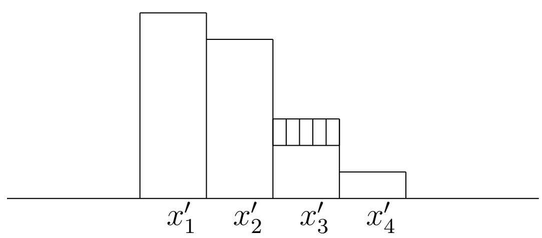
Операцией *сбросить δ-кирпич* с места i на место j $(1 \leqslant i < j \leqslant n)$ назовём преобразование набора $\overline{X}$ в набор $\overline{X'}$, заданное следующим образом:

$$(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_{j-1}, x_j, x_{j+1}, \ldots, x_n) \to (x_1, \ldots, x_{i-1}, x_i - \delta, x_{i+1}, \ldots, x_{j-1}, x_j + \delta, x_{j+1}, \ldots, x_n).$$

Операция осуществима, только если новый набор $\overline{X'}$ сохраняет свою монотонность (т.е. только при условиях $x_i - \delta \geqslant x_{i+1}$ и $x_{j-1} \geqslant x_j + \delta$, когда i и j не являются соседними, и только при условии $x_i - \delta \geqslant x_j + \delta$, когда $i + 1 = j$; это все места, где могла сломаться монотонность).

Смысл названия операции становится предельно понятен, если рисовать для последовательностей $\overline{X}$ *диаграммы Юнга.*





Лемма. *Пусть даны две невозрастающие последовательности $\overline{X}$ и $\overline{Y}$ неотрицательных чисел длины n и пусть $\overline{X} \succ \overline{Y}$. Тогда из $\overline{X}$ можно получить $\overline{Y}$ с помощью конечного числа операций сбрасывания кирпича.*

► Индукция по n. Если $n = 1$, то последовательности совпадают. Если $n > 1$ и $x_1 = y_1$, то по предположению индукции последовательности, полученные отбрасыванием первого элемента, можно преобразовать одну в другую. Условие мажорируемости для редуцированных последовательностей легко проверяется. Действуя теми же операциями в исходных последовательностях, получим пошаговое преобразование $\overline{X}$ в $\overline{Y}$ (ясно, что добавленный обратно первый самый большой элемент x_1 в начале последовательности никак не мог привести к потере монотонности на каком-либо шаге).

Далее считаем, что $x_1 > y_1$. Теперь будем строить операции так, чтобы количество индексов i, при которых $x_i \neq y_i$, уменьшалось хотя бы на 1. Если мы сможем построить такие операции, то когда-нибудь окажется $x'_1 = y_1$ и мы воспользуемся предположением индукции.

Построение операции следующее: ищем наименьший индекс q, такой что $x_q < y_q$ (существует в силу равенства сумм всех x_i и y_i и неравенства $x_1 > y_1$); находим наибольший индекс p, меньший q, при котором $x_p > y_q$ (существует в силу неравенства на первые элементы). Обозначаем $\delta_p := x_p - y_p > 0$, $\delta_q := y_q - x_q > 0$, $\delta = \min(\delta_p, \delta_q)$. Сбрасываем δ-кирпич с места p на место q.

Убедимся, что операция корректна и обладает заявленными свойствами.

Во-первых, операция не ломает монотонности:

Рассмотрим два случая: а) $p + 1 < q$ и б) $p + 1 = q$. В случае а) оцениваем:

$$x'_p = x_p - \delta \geqslant x_p - \delta_p = y_p \geqslant y_{p+1} = x_{p+1} = x'_{p+1};$$
$$x'_q = x_q + \delta \leqslant x_q + \delta_q \leqslant y_q \leqslant y_{q-1} = x_{q-1} = x'_{q-1}.$$

Последние два равенства в этих двух строках используют, что индексы p и q не являются соседними.

В случае б):
$$x'_p = x_p - \delta \geqslant x_p - \delta_p = y_p \geqslant y_q = x_q + \delta_q \geqslant x_q + \delta = x'_q.$$

Во-вторых, операция действительно сравнивает одну пару ранее не равных чисел (с индексами p или q в зависимости от того, чему именно оказалось равно δ), не теряя при этом других равенств.

В третьих, для нового набора останется верным $\overline{X'} \succ \overline{Y}$. Действительно в неравенствах из определения мажорирования при $i \geq q$ сумма слева не изменится, а при меньших i все слагаемые $x'_i \geqslant y_i$ (так как q — первый индекс, где это не верно, а после операции для $i = p$ это осталось верным) $\Rightarrow$ первые $q-1$ неравенство из определения мажорирования тоже справедливы. ◄

По сути, построена такая цепочка $\overline{X} \succ \overline{Z_1} \succ \overline{Z_2} \succ \ldots \succ \overline{Z_M} \succ \overline{Y}$ промежуточных последовательностей, что любые две соседние в цепочке последовательности отличает друг от друга сбрасывание кирпича.

Неравенство Мюрхеда

Теорема (*Неравенство Мюрхеда*)**.** *Пусть даны две невозрастающие последовательности $\overline{X}$ и $\overline{Y}$ неотрицательных чисел длины n и пусть $\overline{X} \succ \overline{Y}$. Тогда для любых положительных вещественных чисел $a_1, \ldots, a_n$ выполнено неравенство:*

$$T_{\overline{X}}(a_1, \ldots, a_n) \geqslant T_{\overline{Y}}(a_1, \ldots, a_n).$$

▷ Каждому упорядоченному набору $\overline{X}$ естественным образом сопоставим точку X в пространстве (и то, и то — наборы из n чисел). На каждую из полученных точек подействуем всевозможными перестановками координат; множества полученных точек будем обозначать соответствующими толстыми буквами, т.е. $\mathbb{X} = \{\widehat{\sigma}(X) \mid \sigma \in S_n\}$ (в этих множествах могут быть совпадающие элементы, без учёта совпадений их должно быть $n!$).

Утверждение 3. *Если $\overline{X} \succ \overline{Y}$, то в введённых обозначениях $conv(\mathbb{X}) \supseteq conv(\mathbb{Y})$.*

► Построим цепочку $\overline{X} \succ \overline{Z_1} \succ \overline{Z_2} \succ \ldots \succ \overline{Z_M} \succ \overline{Y}$ упорядоченных наборов из доказательства предыдущей леммы. Наша ближайшая цель — показать, что $conv(\mathbb{X}) \supseteq conv(\mathbb{Z}_1) \supseteq \ldots \supseteq conv(\mathbb{Y})$. Отсюда тривиальным образом будет следовать утверждение.

Пусть $\overline{Z_{i+1}}$ получена из $\overline{Z_i}$ сбрасыванием кирпича с индекса p на индекс q. Утверждается, что точка Z_{i+1} лежит на отрезке, соединяющем Z_i и Z'_i, где Z'_i — точка, полученная из Z_i перестановкой p-й и q-й координаты. Действительно, все остальные координаты точек Z_{i+1}, Z_i и Z'_i совпадают, а если выписать отдельно две оставшиеся координаты, то они устроены так: $(x_p - \delta, x_q + \delta)$, (x_p, x_q), (x_q, x_p). Отсюда ясно, что первая точка лежит отрезке между второй и третьей ($x_p > x_p - \delta \geqslant x_q + \delta > x_q$, нарисуйте на плоскости), а значит и в $conv(\mathbb{Z}_i)$. Аналогично для любой $\sigma \in S_n$ точка $\widehat{\sigma}(Z_{i+1})$ лежит на отрезке $\widehat{\sigma}(Z_i)\widehat{\sigma}(Z'_i)$. Получаем, что все точки множества $\mathbb{Z}_{i+1}$ лежат в $conv(\mathbb{Z}_i)$ и, следовательно, $conv(\mathbb{Z}_{i+1}) \subseteq conv(\mathbb{Z}_i)$.

Итак, вложенность соседних выпуклых оболочек доказана. В частности, $conv(\mathbb{X}) \supseteq conv(\mathbb{Y})$.◄

Зафиксируем набор положительных чисел $a_1, \ldots, a_n$. Рассмотрим функцию $f : \mathbb{R}^n_{\geqslant 0} \to \mathbb{R}$, заданную формулой: $f(x_1, \ldots, x_n) = a_1^{x_1} \cdot \ldots \cdot a_n^{x_n}$.

Утверждение 4. *Определённая выше f выпукла на своей области определения.*

► Выберем точки X и Y. Положим $\alpha := x_1 \ln(a_1) + \ldots + x_n \ln(a_n)$, $\beta := y_1 \ln(a_1) + \ldots + y_n \ln(a_n)$. Тогда $\lambda f(X) + \mu f(Y) = \lambda e^\alpha + \mu e^\beta \geqslant e^{\lambda\alpha+\mu\beta} = f(\lambda X + \mu Y)$. Неравенство в середине — следствие выпуклости e^x как функции одной переменной. ◄

Рассмотрим точку Y. Согласно утверждению 3 она лежит внутри $conv(\mathbb{X})$. Согласно утверждению 2 это означает, что для неё найдётся набор вещественных неотрицательных чисел m_σ ("масс"), дающих в сумме 1, таких что $Y = \sum\limits_{\sigma \in S_n} m_\sigma \widehat{\sigma}(X)$ (так как точек $n!$, удобно сами перестановки использовать в качестве индексов у буквы m). А теперь в предыдущем равенстве перемешаем координаты всеми возможными способами: для произвольной перестановки $\tau \in S_n$ напишем:

$$\widehat{\tau}(Y) = \sum_{\sigma \in S_n} m_\sigma \widehat{\tau}(\widehat{\sigma}(X)).$$

Согласно утверждению 4, f — выпуклая функция, для каждого из таких равенств можно написать неравенство Йенсена:

$$f(\widehat{\tau}(Y)) \leqslant \sum_{\sigma \in S_n} m_\sigma f(\widehat{\tau}(\widehat{\sigma}(X))).$$

Наконец, сложим все эти неравенства по всем перестановкам $\tau \in S_n$:

$$\sum_{\tau \in S_n} f(\widehat{\tau}(Y)) \leqslant \sum_{\tau \in S_n} \sum_{\sigma \in S_n} m_\sigma f(\widehat{\tau}(\widehat{\sigma}(X))) = \sum_{\sigma \in S_n} m_\sigma \sum_{\tau \in S_n} f(\widehat{\tau}(\widehat{\sigma}(X))).$$

Остаётся лишь заметить, что левая и правая части как раз равны тому, что нужно:

$$\sum_{\tau \in S_n} f(\widehat{\tau}(Y)) = \sum_{\tau \in S_n} a_1^{y_{\tau(1)}} \cdot \ldots \cdot a_n^{y_{\tau(n)}} = T_{\overline{Y}}(a_1, \ldots, a_n);$$

$$\sum_{\sigma \in S_n} m_\sigma \sum_{\tau \in S_n} f(\widehat{\tau}(\widehat{\sigma}(X))) = \sum_{\sigma \in S_n} m_\sigma \cdot T_{\overline{X}}(a_1, \ldots, a_n) = T_{\overline{X}}(a_1, \ldots, a_n).$$

Комментарий к последнему равенству: если τ пробегает все перестановки, то при фиксированной σ выражение $\tau(\sigma)$ также пробегает все перестановки.◁