Линейное программирование

Н.И. Погодаев

3 ноября 2015

§1 Введение

Задача линейного программирования состоит в том, чтобы найти минимум (или максимум) линейной функции на множестве, которое задаётся системой линейных равенств и неравенств. Формально она может быть записана следующим образом:

$$c_1x_1 + c_2x_2 + \cdots + c_nx_n \to \min(\max) \tag{1.1}$$

$$\begin{cases} a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \geq b_i, & i \in I_1, \\ a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \leq b_i, & i \in I_2, \\ a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i, & i \in I_3. \end{cases} \tag{1.2}$$

Терминология:

- $(x_1, \ldots, x_n)$ — переменные задачи;
- $c_1x_1 + c_2x_2 + \cdots + c_nx_n$ — целевая функция или цена задачи;
- $a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \geq b_i$ — ограничение;
- $a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \leq b_i$ — ограничение;
- $a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i$ — ограничение;
- множество всех решений системы (1.2) — допустимое множество;
- любая точка из допустимого множества — допустимая точка.

Напомним, что точка $x^* \in X$ называется точкой минимума (максимума) функции f на множестве X, если

$$f(x^*) \leq f(x) \quad (f(x^*) \geq f(x)) \quad \forall x \in X.$$

Задачи линейного программирования часто возникают в экономике. Приведём ряд (игрушечных) примеров.

Пример 1 (задача о диете)**.** Пятилетняя девочка Дара ест без принуждения только макароны, сосиски и шоколад. Родители хотят, чтобы Дара получала каждый день все необходимые белки, жиры и углеводы. Содержание питательных веществ в 100 г. каждого продукта и суточная норма указаны в таблице.

	Белки	Жиры	Углеводы	Цена (р. за 100 г.)
Макароны	10	1	70	5
Сосиски	11	24	2	30
Шоколад	7	36	52	70
Суточная норма	68	70	272	

Какую минимальную сумму необходимо тратить на питание для Дары?
Решение:

Пусть

x_1 = число 100-граммовых порций макарон, съедаемых Дарой в день;
x_2 = число 100-граммовых порций сосисок, съедаемых Дарой в день;
x_3 = число 100-граммовых порций шоколада, съедаемых Дарой в день.

Стоимость такой продуктовой корзины равна

$$5x_1 + 30x_2 + 70x_3.$$

При этом Дара должна получить

$$10x_1 + 11x_2 + 7x_3 \geq 68$$

грамм белков,

$$x_1 + 24x_2 + 36x_3 \geq 70$$

грамм жиров и

$$70x_1 + 2x_2 + 52x_3 \geq 272$$

грамм углеводов.

Наконец, нужно учесть условия неотрицательности

$$x_1 \geq 0, \quad x_2 \geq 0, \quad x_3 \geq 0.$$

Таким образом, мы получаем задачу

$$5x_1 + 30x_2 + 70x_3 \to \min,$$
$$\begin{cases} 10x_1 + 11x_2 + 7x_3 \geq 68, \\ x_1 + 24x_2 + 36x_3 \geq 70, \\ 70x_1 + 2x_2 + 52x_3 \geq 272, \\ x_i \geq 0, \quad i = 1, 2, 3. \end{cases}$$

□

Её решением будет продуктовый набор, состоящий из 380 г. макарон и 276 г. сосисок, общей стоимостью 102 рубля.

Пример 2 (задача о расписании)**.** Госпиталю требуется составить расписание работы медсестёр. Сутки разбиты на 6 четырёхчасовых смен. Количество медсестёр, требуемое в каждую смену, задано в таблице.

Смена	Требуемое количество медсестёр
00:00 – 04:00	8
04:00 – 08:00	9
08:00 – 12:00	15
12:00 – 16:00	14
16:00 – 20:00	13
20:00 – 24:00	11

Каждая медсестра работает две четырёхчасовые смены подряд. После этого заступить на работу она может лишь через 16 часов. Какое минимальное число медсестёр требуется нанять госпиталю?

Решение:

Пусть

$$x_i = \text{количество медсестёр, заступивших на работу в } i\text{-ю смену.}$$

Тогда общее число медсестёр равно

$$x_1 + x_2 + x_3 + x_4 + x_5 + x_6.$$

В первую смену работает $x_1 + x_6$ медсестёр, во вторую — $x_1 + x_2$ и т.д.

Таким образом, мы приходим к задаче линейного программирования

$$x_1 + x_2 + x_3 + x_4 + x_5 + x_6 \to \min$$

$$\begin{cases}
\begin{array}{llllllr}
x_1 & & & & & +x_6 & \geq 8, \\
x_1 & +x_2 & & & & & \geq 9, \\
 & x_2 & +x_3 & & & & \geq 15, \\
 & & x_3 & +x_4 & & & \geq 14, \\
 & & & x_4 & +x_5 & & \geq 13, \\
 & & & & x_5 & +x_6 & \geq 11, \\
\end{array} \\
x_1 \geq 0, \quad x_2 \geq 0, \quad x_3 \geq 0, \quad x_4 \geq 0, \quad x_5 \geq 0, \quad x_6 \geq 0.
\end{cases}$$

Её решением является вектор $(8, 3, 12, 2, 11, 0)$, то есть в первую смену заступает 8 медсестёр, во вторую — 3 и т.д. Итого, требуется нанять 36 медсестёр. □

Пример 3 (транспортная задача)**.** Компания «Браавосэнерджи» владеет тремя электростанциями, которые обеспечивают электроэнергией четыре города (Браавос, Пентос, Норвос, Лорат). Электростанция 1 производит 35 миллионов кВт/ч электроэнергии, электростанция 2 — 50 миллионов, электростанция 3 — 40 миллионов. Пиковое потребление электроэнергии приходится на 14:00 и составляет в Браавосе — 45 миллионов кВт/ч, в Пентосе — 20 миллионов, в Норвосе — 30 миллионов, в Лорате — 30 миллионов. Цена передачи 1 кВт/ч электроэнергии от электростанции до города зависит от расстояния, которое электроэнергия должна пройти. Соответствующие цены приведены в таблице.

	Город 1	Город 2	Город 3	Город 4	Мощность
ЭС 1	\$8	\$6	\$10	\$9	35
ЭС 2	\$9	\$12	\$13	\$7	50
ЭС 3	\$14	\$9	\$16	\$5	40
Нужды	45	20	30	30	

Требуется минимизировать затраты компании на передачу электроэнергии во время пиковой нагрузки.

Решение:

Пусть

$$x_{ij} = \begin{cases} \text{количество электроэнергии (в млн. кВт/ч)} \\ \text{выработанное } i\text{-й ЭС и отправленное в } j\text{-й город.} \end{cases}$$

Полные затраты на передачу электроэнергии:

$$\begin{aligned}
&8x_{11}+6x_{12}+10x_{13}+9x_{14} && \text{(цена передачи электроэнергии для ЭС 1)}\\
&+9x_{21}+12x_{22}+13x_{23}+7x_{24} && \text{(цена передачи электроэнергии для ЭС 2)}\\
&+14x_{31}+9x_{32}+16x_{33}+5x_{34} && \text{(цена передачи электроэнергии для ЭС3)}
\end{aligned}$$

Так как переданная энергия не может превышать произведённую, получаем

$$\begin{aligned}
x_{11}+x_{12}+x_{13}+x_{14} &\le 35,\\
x_{21}+x_{22}+x_{23}+x_{24} &\le 50,\\
x_{31}+x_{32}+x_{33}+x_{34} &\le 40.
\end{aligned}$$

Необходимо, чтобы каждый город получил требуемое количество электроэнергии:

$$\begin{aligned}
x_{11}+x_{21}+x_{31} &\ge 45,\\
x_{12}+x_{22}+x_{32} &\ge 20,\\
x_{13}+x_{23}+x_{33} &\ge 30,\\
x_{14}+x_{24}+x_{34} &\ge 30.
\end{aligned}$$

Наконец, осталось добавить ограничения

$$x_{ij}\ge 0, \qquad i=1,2,3, \quad j=1,2,3,4.$$

□

Пример 4. (задача о рекламе) «Дориан авто» производит машины класса люкс и внедорожники. Компания предполагает, что наиболее вероятные покупатели её машин — это мужчины и женщины с высоким достатком. Чтобы привлечь внимание этих двух групп, «Дориан авто» хочет организовать масштабную рекламную кампанию на телевидение и купить 1-минутное рекламное время во время ток-шоу и футбольных матчей. По опросам каждый рекламный ролик во время ток-шоу смотрят 7 миллионов женщин и 2 миллиона мужчин. Каждый рекламный ролик во время футбольного матча смотрят 2 миллиона женщин и 12 миллиона мужчин. Одна минута во время ток-шоу стоит \$50000, а одна минута во время футбольной игры — \$100000. «Дориан авто» хочет, чтобы её рекламу увидело хотя бы 28 миллионов женщин и 24 миллиона мужчин. Какую минимальную сумму «Дориан авто» должна затратить для достижения этой цели?

Решение:
Пусть

$$x_1 = \text{количество рекламных роликов во время ток-шоу},$$
$$x_2 = \text{количество рекламных роликов во время матчей.}$$

Общая стоимость рекламных роликов равна

$$50x_1 + 100x_2$$

тысяч долларов. При этом ролики должны посмотреть

$$7x_1 + 2x_2 \geq 28$$

миллионов женщин и

$$2x_1 + 12x_2 \geq 24$$

миллионов мужчин. Наконец, добавим ограничения

$$x_1 \geq 0, \quad x_2 \geq 0.$$

Получаем задачу линейного программирования

$$50x_1 + 100x_2 \to \min,$$
$$\begin{cases} 7x_1 + 2x_2 \geq 28, \\ 2x_1 + 12x_2 \geq 24, \\ x_i \geq 0, \quad i = 1, 2. \end{cases}$$

Чтобы её решить, построим допустимое множество и линии уровня целевой функции (см. рис. 1).

Из рисунка видно, что допустимая точка с наименьшей ценой лежит на пересечении прямых

$$7x_1 + 2x_2 = 28, \quad 2x_1 + 12x_2 = 24.$$

Решая эту систему, получим $x_1 = 3.6$, $x_2 = 1.4$; цена этого решения — $50 \cdot 3.6 + 100 \cdot 1.6 = 320$ тыс. долларов. Однако по условиям задачи «Дориан авто» может купить лишь целое число минут. Опять же из рисунка видно, что допустимая точка с целыми координатами, имеющая минимальную цену, — это точка $(4, 2)$.

Ответ: «Дориан авто» должна купить 4 минуты рекламного времени во время ток-шоу и 2 минуты во время футбола, стоимость такого рекламного пакета — 400 тыс. долларов. $\square$

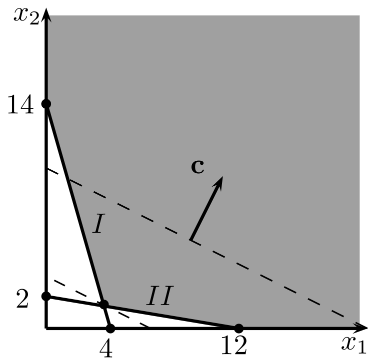


Рис. 1: Решение задачи о рекламе. Допустимое множество — это область, закрашенная серым цветом. Пунктирном изображены линии уровня целевой функции. Вектор $\mathbf{c} = (50, 100)$ указывает на направление возрастания целевой функции.

Пример 5 (планирование производства)**.** Мастерская папы Карло производит два типа деревянных игрушек: Буратино и Пиноккио. Каждый Буратино продаётся за \$27 и требует материалов на \$10. Трудозатраты на производство одного Буратино составляют \$14. Пиноккио продаётся за \$21 и требует материалов на \$9. Трудозатраты на производство одного Пиноккио составляют \$10. Для производства игрушек требуется два типа работ: резка по дереву и шлифовка. На каждого Буратино требуется 2 часа шлифовки и 1 час резки по дереву. На каждого Пиноккио — 1 час шлифовки и 1 час резки по дереву. Каждую неделю мастерская может получить все необходимые материалы, но только 100 часов работы шлифовщика и 80 — резчика по дереву. Спрос на Пиноккио неограничен, однако за неделю покупают не больше 40 Буратин. Сколько игрушек каждого вида должна выпускать мастерская, чтобы получить наибольшую прибыль?

Решение:
Пусть

$$\begin{aligned} x_1 &= \text{количество произведённых Буратин}, \\ x_2 &= \text{количество произведённых Пиноккио}. \end{aligned}$$

Прибыль, которую нужно максимизировать, равна

$$(27 - 10 - 14)x_1 + (21 - 9 - 10)x_2 = 3x_1 + 2x_2.$$

На шлифовку тратится

$$2x_1 + x_2 \leq 100$$

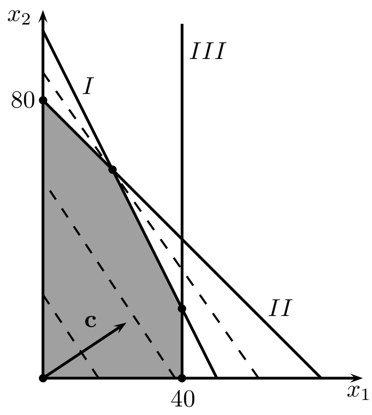


Рис. 2: Решение задачи планирования производства. Допустимое множество закрашено серым цветом. Пунктирном изображены линии уровня целевой функции, вектор **c** указывает направление возрастания.

часов, на резку по дереву —

$$x_1 + x_2 \leq 80$$

часов. Ограничение на количество Буратин:

$$x_1 \leq 40.$$

Неотрицательность:

$$x_1 \geq 0, \quad x_2 \geq 0.$$

Таким образом, мы получаем задачу

$$3x_1 + 2x_2 \to \max,$$
$$\begin{cases} 2x_1 + x_2 \leq 100, \\ x_1 + x_2 \leq 80, \\ x_1 \leq 40, \\ x_i \geq 0, \quad i = 1, 2. \end{cases}$$

Решая её графически (см. рис 2), заключаем, что оптимальное решение лежит на пересечении прямых

$$2x_1 + x_2 = 100, \quad x_1 + x_2 = 80.$$

Следовательно, $x_1 = 20$, $x_2 = 60$. Итак, каждую неделю мастерская должна производить 20 Буратин и 60 Пиноккио, при этом её прибыль этом составит $3 \cdot 20 + 2 \cdot 60 = 180\$$.

□

§2 Основные понятия многомерной геометрии

Далее мы обобщим классические геометрические понятия (отрезок, луч, прямая, плоскость и т.п.) на случай n-мерного пространства.

Определение 1. Пусть $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Рассмотрим следующие множества:

$$\begin{aligned} &\left\{\mathbf{x} + \lambda(\mathbf{y} - \mathbf{x}) \;:\; \lambda \in [0,1]\right\}, \\ &\left\{\mathbf{x} + \lambda(\mathbf{y} - \mathbf{x}) \;:\; \lambda \in [0,+\infty)\right\}, \\ &\left\{\mathbf{x} + \lambda(\mathbf{y} - \mathbf{x}) \;:\; \lambda \in (-\infty,+\infty)\right\}. \end{aligned}$$

Первое из них называется отрезком, соединяющим $\mathbf{x}$ и $\mathbf{y}$, второе — лучом, выходящим из $\mathbf{x}$ и проходящим через $\mathbf{y}$, третье — прямой, проходящей через $\mathbf{x}$ и $\mathbf{y}$.

Множества, которые мы определим ниже являются “строительными кирпичиками” многомерной геометрии.

Определение 2. Пусть $\mathbf{a} \in \mathbb{R}^n$ и $b \in \mathbb{R}$. Множество

$$H = \{\mathbf{x} \in \mathbb{R}^n \;:\; \mathbf{a}'\mathbf{x} = b\}$$

называется гиперплоскостью. Каждая гиперплоскость делит пространство $\mathbb{R}^n$ на две области

$$H^+ = \{\mathbf{x} \in \mathbb{R}^n \;:\; \mathbf{a}'\mathbf{x} \geq b\}, \quad H^- = \{\mathbf{x} \in \mathbb{R}^n \;:\; \mathbf{a}'\mathbf{x} \leq b\}.$$

Первая называется положительным, а вторая — отрицательным полупространством.

Определение 3. Мы скажем, что гиперплоскости

$$H_i = \{\mathbf{x} \in \mathbb{R}^n \;:\; \mathbf{a}_i'\mathbf{x} = b_i\}, \quad i = 1, \ldots, k,$$

линейно независимы, если набор векторов $\mathbf{a}_i$, $i = 1, \ldots, k$, линейно независим.

Определение 4. Множество, образованное пересечением $k \leq n$ линейно независимых гиперплоскостей, называется $(n-k)$-мерной плоскостью.

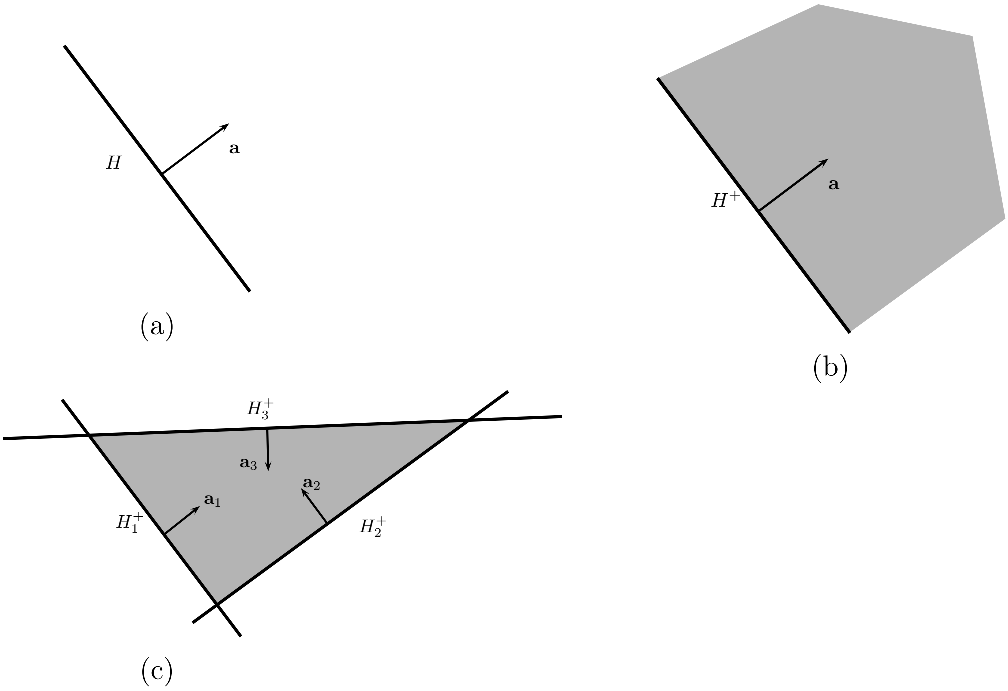


Рис. 3: Гиперплоскость (a), полупространство (b) и полиэдр (c)

Пусть $n = 3$. Тогда гиперплоскость — это обычная двумерная плоскость. Пересечение двух линейно независимых гиперплоскостей даёт одномерное множество — прямую (одномерная плоскость). Наконец, пересечение трёх линейно независимых гиперплоскостей даёт нульмерное множество — точку (нульмерная плоскость).

Вообще для любого n, $(n-1)$-мерная плоскость — это гиперплоскость, одномерная плоскость — это прямая, нульмерная плоскость — это точка.

Утверждение 1. *Пусть имеется произвольный набор гиперплоскостей H_i, $i = 1, \dots, m$. Их пересечение $\Pi = \bigcap_{i=1}^m H_i$ либо пусто, либо является $(n-k)$-мерной плоскостью, где k — максимальное число линейно независимых гиперплоскостей в наборе H_i, $i = 1, \dots, m$.*

Доказательство. Обозначим через $\mathbf{A}$ матрицу со строками $\mathbf{a}_i$, а через $\mathbf{b}$ — вектор с компонентами b_i. Тогда

$$\Pi = \{\mathbf{x} : \ \mathbf{A}\mathbf{x} = \mathbf{b}\},$$

то есть Π является множеством решений системы линейных уравнений. Стандартная теория говорит, что система либо несовместна, либо из неё можно выкинуть линейно зависимые строки. $\square$

Определение 5. Множество, образованное пересечением конечного числа (положительных) полупространств, называется *полиэдром*:

$$P = \bigcap_{i=1}^{m} H_i^+.$$

Примеры полиэдров: полупространство, гиперплоскость, k-мерная плоскость, полоса, угол, треугольник, прямоугольник, тетраэдр, куб и т.д.

Утверждение 2. *Допустимое множество в задаче линейного программирования является полиэдром.*

Доказательство. Достаточно заметить, что

$$a_{i1}x_1 + \dots + a_{in}x_n \le b_i \quad \Leftrightarrow \quad -a_{i1}x_1 - \dots - a_{in}x_n \ge -b_i$$

и

$$a_{i1}x_1 + \dots + a_{in}x_n = b_i \quad \Leftrightarrow \quad \begin{cases} a_{i1}x_1 + \dots + a_{in}x_n \ge b_i, \\ -a_{i1}x_1 - \dots - a_{in}x_n \ge -b_i. \end{cases}$$

□

Определение 6. Множество $S \subseteq \mathbb{R}^n$ называется *выпуклым*, если вместе с любыми двумя точками оно содержит отрезок, их соединяющий.

На первый взгляд кажется, что понятие выпуклого множества выбивается из контекста. Однако это не так. Позднее мы докажем, что замкнутое выпуклое множество можно определить как пересечение бесконечного числа полупространств.

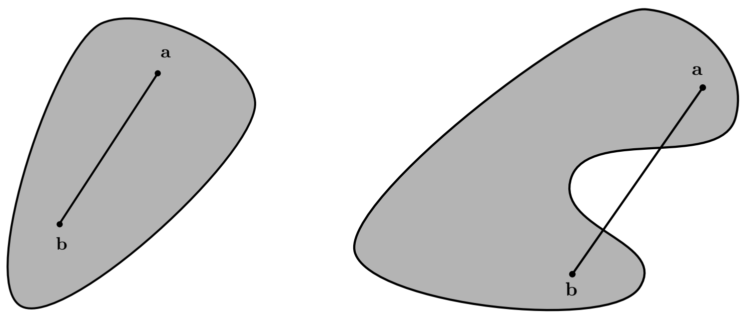
Упражнение 1. Докажите, что пересечение выпуклых множеств есть выпуклое множество.

Упражнение 2. Докажите, что полиэдр является замкнутым выпуклым множеством.

§3 Грани и вершины полиэдра

Для полиэдров можно определить понятие размерности.

Определение 7. Полиэдр имеет размерность k, если он содержится в некоторой k-мерной и не содержится ни в какой $(k-1)$-мерной плоскости.



(a) (b)

Рис. 4: Выпуклое множество (a) и невыпуклое множество (b).

Определение 8. Гранью полиэдра $P = \bigcap_{i=1}^{m} H_i^+$ называется любое множество вида

$$\Gamma = P \cap \left(\bigcap_{i \in I} H_i \right), \quad I \subset \{1, \ldots, m\}.$$

Любая грань полиэдра сама является полиэдром. Следовательно, для грани определено понятие размерности. Сам полиэдр P является гранью размерности $k = \dim P$. Грань размерности $(k-1)$ называется гипергранью. Грань размерности 1 называется ребром. Грань размерности 0 называется вершиной.

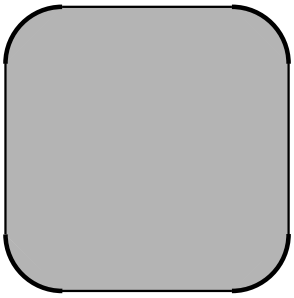
В дальнейшем нас будут интересовать именно вершины. Отметим, что полиэдр может их не иметь. Соответствующие примеры — гиперплоскость и полоса.

Определение 9. Точка $\mathbf{x} \in S$ называется *крайней точкой* выпуклого множества S, если она не является серединой отрезка, принадлежащего S.

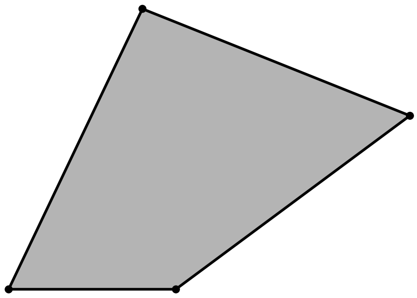
Теорема 1. *Крайние точки полиэдра — это в точности его вершины.*

Доказательство.

(крайняя точка) $\Rightarrow$ (вершина)
Положим $I = \{i \,:\, \mathbf{a}_i'\mathbf{x}^* = b_i\}$. Если $\mathbf{x}^*$ не вершина, то множество M, образованное пересечением гиперплоскостей $\{\mathbf{x} \,:\, \mathbf{a}_i\mathbf{x} = b_i\}$, $i \in I$, является линейным многообразием размерности не меньше единицы. Следовательно, существует ненулевой вектор $\mathbf{d}$, параллельный этому многообразию.



(a) Крайними точками являются дуги окружности



(b) Крайними точками являются вершины многоугольника

Рис. 5: Крайние точки выпуклых множеств.

Положим $\mathbf{y} = \mathbf{x}^* - \varepsilon \mathbf{d}$, $\mathbf{z} = \mathbf{x}^* + \varepsilon \mathbf{d}$. Очевидно, $\mathbf{y}, \mathbf{z} \in M$. Кроме того, при малых ε

$$\mathbf{a}_i'\mathbf{y} > b_i, \qquad \mathbf{a}_i'\mathbf{z} > b_i \qquad i \notin I.$$

Так как $\mathbf{x}^* = \frac{\mathbf{y}+\mathbf{z}}{2}$, то $\mathbf{x}^*$ не может быть крайней точкой.

(вершина) $\Rightarrow$ (крайняя точка)
Предположим, что $\mathbf{x}^* \in P$ не является крайней точкой, но при этом является вершиной. Это значит, что $\mathbf{x}^* \in P$ можно представить в виде

$$\mathbf{x}^* = \frac{\mathbf{y}+\mathbf{z}}{2}, \qquad \mathbf{y}, \mathbf{z} \in P, \qquad \mathbf{x}^* \neq \mathbf{y}, \qquad \mathbf{x}^* \neq \mathbf{z},$$

и существует набор $\mathbf{a}_i$, $i \in I$, из n линейно независимых векторов такой, что

$$\mathbf{a}_i'\mathbf{x}^* = b_i, \qquad i \in I.$$

Предположим, что $\mathbf{a}_i'\mathbf{y} > b_i$ для некоторого $i \in I$. Тогда должно выполняться неравенство $\mathbf{a}_i'\mathbf{z} < b_i$, но это невозможно поскольку $\mathbf{z} \in P$ Следовательно,

$$\mathbf{a}_i'\mathbf{y} = b_i, \qquad i \in I.$$

Поскольку $\mathbf{a}_i$ линейно независимы, то $\mathbf{x}^* = \mathbf{y}$, и мы пришли к противоречию. $\square$

Терминология: Если в точке $\mathbf{x}^*$ полиэдра

$$P = \{\mathbf{x} \in \mathbb{R}^n \ : \ \mathbf{a}_i'\mathbf{x} \geq b_i, \ i = 1, \ldots, m\}$$

j-е ограничение выполняется как равенство, то есть $\mathbf{a}_j'\mathbf{x}^* = b_j$, то говорят, что j-е ограничение *активно* в $\mathbf{x}^*$.

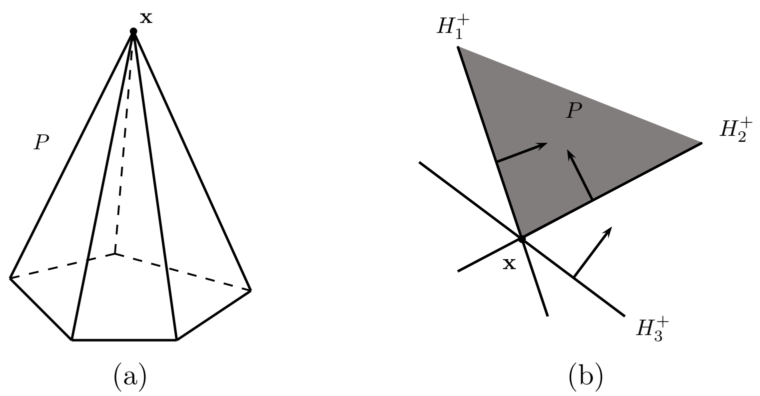


Рис. 6: Примеры вырожденных вершин. Вторая картинка показывает, что вырожденность не является чисто геометрическим свойством.

Таким образом, вершина — это точка полиэдра, в которой активно n линейно независимых ограничений.

Определение 10. Вершина полиэдра называется *невырожденной*, если в ней активны ровно n ограничений. В противном случае (если активных ограничений строго больше n) вершина называется *вырожденной*.

§4 Оптимальность вершины

Следующая теорема, по сути, является теоремой об альтернативе — либо полиэдр содержит прямую, либо у него есть вершины.

Теорема 2. *Пусть* $P = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{a}_i'\mathbf{x} \geq b_i, i = 1, \ldots, m\}$ *— непустой полиэдр. Тогда следующие утверждения эквивалентны:*

(a) полиэдр P не содержит прямой;

(b) у полиэдра P есть хотя бы одна вершина.

Доказательство.

(a) ⇒ (b)
Выберем в полиэдре P произвольную точку $\mathbf{x}^*$. Пусть I — множество, состоящее из номеров ограничений, активных в $\mathbf{x}^*$. Обозначим через k максимальное число линейно независимых ограничений в наборе

$$\mathbf{a}_i'\mathbf{x} = b_i, \quad i \in I. \tag{4.1}$$

Если $k = n$, то $\mathbf{x}^*$ — вершина по определению. Если $k < n$, то через $\mathbf{x}^*$ можно провести прямую, в каждой точке которой активны все ограничения (4.1). Эта прямая имеет вид

$$l = \{\mathbf{x}^* + \lambda \mathbf{d} \,:\, \lambda \in \mathbb{R}\}.$$

где $\mathbf{d}$ — произвольный ненулевой вектор, ортогональный линейной оболочке, натянутой на вектора $\mathbf{a}_i$, $i \in I$.

Поскольку P не содержит прямых, то при некотором значении λ построенная прямая покинет полиэдр. Это значит, что одно из ранее неактивных ограничений сначала станет активным, а затем нарушится. Иначе говоря, существуют λ^* и $j \notin I$ такие, что $\mathbf{a}'_j(\mathbf{x}^* + \lambda^* \mathbf{d}) = b_j$.

Таким образом, в точке $\mathbf{x}^* + \lambda^* \mathbf{d}$ активны ограничения

$$\mathbf{a}'_i \mathbf{x} = b_i, \quad i \in I \cup \{j\},$$

из которых можно выбрать $k+1$ линейно независимых. Действительно, если это не так, то $\mathbf{a}_j = \sum_{i \in I} \lambda_i \mathbf{a}_i$. Умножая последнее равенство на $\mathbf{d}$, получаем $\mathbf{a}'_j \mathbf{d} = 0$. Но тогда j-е ограничение должно быть активно в $\mathbf{x}^*$, поскольку $\mathbf{a}'_j \mathbf{x}^* = \mathbf{a}'_j(\mathbf{x}^* + \lambda^* \mathbf{d}) = b_j$, чего не может быть.

Продолжая данную процедуру, мы дойдём до точки, в которой активно n линейно независимых ограничений. Эта точка и будет вершиной.

(b) $\Rightarrow$ (a)

Пусть $\mathbf{x}^*$ — вершина. Тогда в ней должны быть активны n линейно независимых ограничений, скажем ограничения, задаваемые векторами $\mathbf{a}_1, \dots, \mathbf{a}_n$. Если полиэдр содержит прямую $\mathbf{x} + \lambda \mathbf{d}$, то при всех λ должны выполняться неравенства

$$\mathbf{a}'_i(\mathbf{x} + \lambda \mathbf{d}) \geq b_i, \quad i = 1, \dots, n.$$

Следовательно, $\mathbf{a}'_i \mathbf{d} = 0$ для всех i. Так как вектора $\mathbf{a}_1, \dots, \mathbf{a}_n$ линейно независимы, то $\mathbf{d} = 0$. $\square$

Теорема 3. *Пусть полиэдр P имеет вершины и функция $\mathbf{c}'\mathbf{x}$ достигает на нём минимума. Тогда минимум достигается в одной из вершин.*

Доказательство. Мы предположили, что множество решений Q задачи линейного программирования непусто. Если v — оптимальное значение целевой функции, то $Q = P \cap \{\mathbf{x} \in \mathbb{R}^n \,:\, \mathbf{c}'\mathbf{x} = v\}$. Следовательно, Q — полиэдр. Поскольку $Q \subset P$ и P не содержит прямых, то Q также не содержит прямых. Значит у Q есть вершина.

Обозначим через $\mathbf{x}^*$ какую-нибудь вершину полиэдра Q. Покажем, что $\mathbf{x}^*$ — вершина P. Если это не так, то должны существовать точки

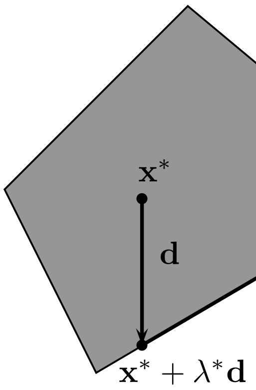


Рис. 7: Стартуя из произвольной точки $\mathbf{x}^*$, мы выбираем направление $\mathbf{d}$ и двигаемся вдоль него, пока не нарушится какое-нибудь ранее неактивное условие. Перед тем как нарушиться оно станет активным. Предположим это произойдёт в точке $\mathbf{x}^* + \lambda^* \mathbf{d}$. Значит в этой точке количество линейно независимых активных ограничений станет больше на единицу. Повторяя данную процедуру, мы дойдём до точки, в которой n линейно независимых условий активны. Это и есть вершина.

$\mathbf{y}, \mathbf{z} \in P$ такие, что $\mathbf{y} \neq \mathbf{x}^*$, $\mathbf{z} \neq \mathbf{x}^*$ и $\mathbf{x}^* = \frac{\mathbf{y}+\mathbf{z}}{2}$. Поскольку v — оптимальное значение, то выполнены соотношения

$$v \leq \mathbf{c}'\mathbf{y}, \quad v \leq \mathbf{c}'\mathbf{z}, \quad v = \frac{\mathbf{c}'\mathbf{y} + \mathbf{c}'\mathbf{z}}{2}.$$

Откуда $\mathbf{c}'\mathbf{y} = \mathbf{c}'\mathbf{z} = v$, то есть $\mathbf{y}, \mathbf{z} \in Q$. Но последнее противоречит тому, что $\mathbf{x}^*$ — вершина Q. $\square$

Отметим, что решение всегда существует, если полиэдр ограничен. Решение не существует, только если значение целевой функции стремится к $-\infty$.

Итак, решение задачи линейного программирования нужно искать среди вершин полиэдра.

§5 Каноническая форма задачи линейного программирования

Задача линейного программирования

$$\mathbf{c}'\mathbf{x} \to \min\,(\max), \quad \mathbf{x} \in P$$

называется задачей в *канонической форме*, если её допустимое множество имеет вид

$$P = \{\mathbf{x} : \ \mathbf{A}\mathbf{x} = \mathbf{b}, \ \mathbf{x} \geq 0\}.$$

Другими словами, допустимое множество является пересечением линейного многообразия

$$M = \{\mathbf{x} : \ \mathbf{A}\mathbf{x} = \mathbf{b}\}$$

и конуса $\mathbf{x} \geq 0$.

Любую задачу линейного программирования можно свести к задаче в канонической форме. Для этого используются два приёма.

Исключение свободных переменных. Если x_j — свободная переменная (то есть на неё не наложено ограничение $x_j \geq 0$), то она заменяется на $x_j^+ - x_j^-$, где x_j^+ и x_j^- — новые переменные, удовлетворяющие условиям $x_j^+ \geq 0$, $x_j^- \geq 0$.

Введение фиктивных переменных. Если имеется ограничение

$$\sum_{j=1}^{n} a_{ij} x_j \leq b_i,$$

вводится фиктивная переменная $s_i = b_i - \sum_{j=1}^{n} a_{ij} x_j$. В итоге получаем

$$\sum_{j=1}^{n} a_{ij} x_j + s_i = b_i, \quad s_i \geq 0.$$

Для ограничения

$$\sum_{j=1}^{n} a_{ij} x_j \geq b_i,$$

вводим фиктивную переменную $s_i = \sum_{j=1}^{n} a_{ij} x_j - b_i$ и получаем

$$\sum_{j=1}^{n} a_{ij} x_j - s_i = b_i, \quad s_i \geq 0.$$

Пример 6. Привести к каноническому виду задачу

$$\begin{aligned} -x_1 - 2x_2 + x_3 &\to \max \\ x_1 + x_2 + \ \ x_3 &= 1 \\ x_1 - x_2 + 2x_3 &\leq 3 \\ x_1 &\geq 0. \end{aligned} \tag{5.1}$$

Решение:

1) Заменяем свободные переменные на разности:

$$\begin{gathered}
-x_1 - 2(x_2^+ - x_2^-) + (x_3^+ - x_3^-) \to \max \\
x_1 + (x_2^+ - x_2^-) + (x_3^+ - x_3^-) = 1 \\
x_1 - (x_2^+ - x_2^-) + 2(x_3^+ - x_3^-) \le 3 \\
x_1,\ x_2^+,\ x_2^-,\ x_3^+,\ x_3^- \ge 0.
\end{gathered}$$

2) Вводя фиктивную переменную

$$s_1 = 3 - x_1 + (x_2^+ - x_2^-) - 2(x_3^+ - x_3^-),$$

приходим к задаче в канонической форме

$$\begin{gathered}
-x_1 - 2x_2^+ + 2x_2^- + x_3^+ - x_3^- \to \max \\
x_1 + x_2^+ - x_2^- + x_3^+ - x_3^- = 1 \\
x_1 - x_2^+ + x_2^- + 2x_3^+ - 2x_3^- + s_1 = 3 \\
x_1,\ x_2^+,\ x_2^-,\ x_3^+,\ x_3^-,\ s_1 \ge 0.
\end{gathered} \tag{5.2}$$

$\square$

Мы скажем, что две задачи линейного программирования *эквивалентны*, если для любой допустимой точки одной задачи можно построить допустимую точку другой задачи, имеющую ту же цену.

Пример 7. Задачи (5.1) и (5.2) эквивалентны.

Решение:

Если (x_1, x_2, x_3) удовлетворяет всем ограничениям задачи (5.1), то точка

$$\big(x_1, \max\{0, x_2\}, \max\{0, -x_2\}, \max\{0, x_3\}, \max\{0, -x_3\}, 3 - x_1 + x_2 - 2x_3\big)$$

удовлетворяет всем ограничениям задачи (5.2), поскольку

$$x_2 = \max\{0, x_2\} - \max\{0, -x_2\} \ \text{ и } \ x_3 = \max\{0, x_3\} - \max\{0, -x_3\}.$$

Из этих же равенств следует, что обе точки имеют одинаковую цену —

$$-x_1 - 2x_2 + x_3.$$

Обратно, если $(x_1, x_2^+, x_2^-, x_3^+, x_3^-, s_1)$ удовлетворяет всем ограничениям задачи (5.2), то точка

$$\left(x_1, x_2^+ - x_2^-, x_3^+ - x_3^-\right)$$

удовлетворяет всем ограничениям задачи (5.1). Обе эти точки имеют одинаковую цену —

$$-x_1 - 2(x_2^+ - x_2^-) + (x_3^+ - x_3^-).$$

$\square$

Вообще нетрудно показать, что если одна задача линейного программирования получена из другой с помощью исключения свободных и введения фиктивных переменных, то обе эти задачи эквивалентны.

Переход от общей формы к канонической

Пусть

$$P = \{\mathbf{x} \in \mathbb{R}^n : \ \mathbf{A}\mathbf{x} \geq \mathbf{b}\}.$$

Рассмотрим отображение $f \colon \mathbb{R}^{2n} \to \mathbb{R}^n$, определённое равенством

$$f(\mathbf{y}, \mathbf{z}) = \mathbf{y} - \mathbf{z}.$$

Это линейное отображение. Его ядро образуется пересечением n линейно независимых гиперплоскостей пространства $\mathbb{R}^{2n}$. Следовательно, его размерность равна n.

Положим

$$\tilde{P} = \{(\mathbf{y}, \mathbf{z}) \in \mathbb{R}^{2n} : \ \mathbf{A}\mathbf{y} - \mathbf{A}\mathbf{z} \geq \mathbf{b}, \ \mathbf{y} \geq 0, \ \mathbf{z} \geq 0\}.$$

Легко проверить, что $f(\tilde{P}) = P$. Заметим, что f — расслоение (то есть действует фактически как проекция).

Пример 8. Пусть

$$P = \{x \in \mathbb{R} : \ x \leq 1\}, \qquad \tilde{P} = \{(x^-, x^+) \in \mathbb{R}^2 : \ x^+ - x^- \leq 1, \ x^- \geq 0, \ x^+ \geq 0\}.$$

Переход от канонической формы к стандартной

Пусть

$$P = \{\mathbf{x} \in \mathbb{R}^n : \ \mathbf{A}\mathbf{x} \geq \mathbf{b}, \ \mathbf{x} \geq 0\}.$$

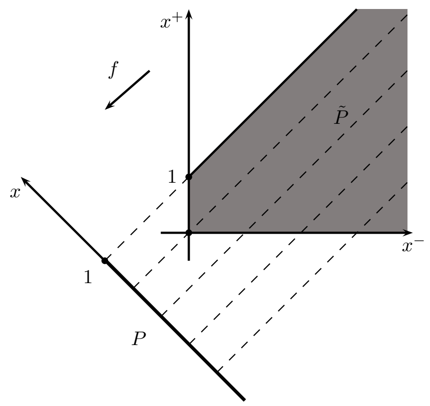


Рис. 8: Переход от общей формы к канонической.

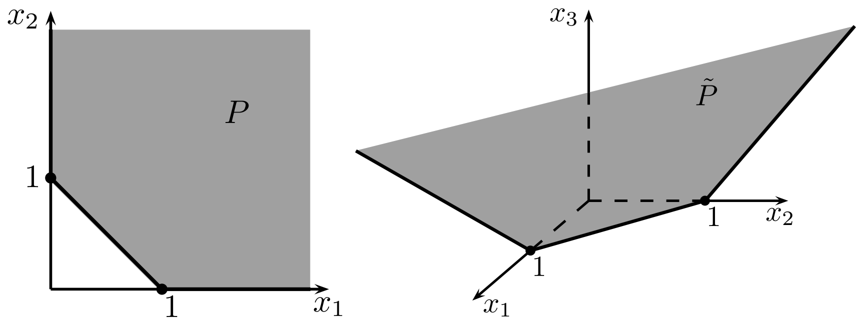


Рис. 9: Переход от канонической формы к стандартной.

Рассмотрим аффинное отображение

$$f\colon \mathbf{x} \mapsto \begin{pmatrix} \mathbf{x} \\ \mathbf{A}\mathbf{x} - \mathbf{b} \end{pmatrix}.$$

Легко видеть, что отображение f инъективно. Положим $\tilde{P} = f(P)$. Тогда

$$\tilde{P} = \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^{n+m} \,:\, \mathbf{A}\mathbf{x} - \mathbf{y} = \mathbf{b},\ \mathbf{x} \geq 0,\ \mathbf{y} \geq 0\,\}.$$

Отображение $f\colon P \to \tilde{P}$ — изоморфизм и, в частности, переводит вершины в вершины.

§6 Вершины полиэдра в канонической форме

В дальнейшем нам окажется удобно работать с задачей линейного программирования в канонической форме. Напомним, что её допустимым множеством является полиэдр, образованный пересечением линейного многообразия $\{\mathbf{x} : \ \mathbf{A}\mathbf{x} = \mathbf{b}\}$ с конусом $\mathbf{x} \geq 0$. Мы будем называть такой полиэдр *полиэдром в канонической форме.*

Пример 9. Изобразить полиэдр

$$P = \begin{cases} x_1 + x_2 + x_3 + \ \ x_4 = 5 \\ x_1 + x_2 - x_3 + 3x_4 = 7 \\ x_i \geq 0, \quad i = 1, 2, 3, 4. \end{cases}$$

Решение:
Прежде всего, найдём общее решение системы

$$\begin{cases} x_1 + x_2 + x_3 + \ \ x_4 = 5 \\ x_1 + x_2 - x_3 + 3x_4 = 7 \end{cases}$$

Оно имеет вид $(x_1, 6 - x_1 - 2x_4, -1 + x_4, x_4)$. Таким образом, линейное многообразие, задаваемое системой, является двумерным и состоит из точек вида

$$(x_1, 6 - x_1 - 2x_4, -1 + x_4, x_4) \in \mathbb{R}^4.$$

Построим проекцию полиэдра P на плоскость (x_1, x_4). Эта проекция — мы обозначим её $\tilde{P}$ — задаётся ограничениями

$$x_1 \geq 0, \quad 6 - x_1 - 2x_4 \geq 0, \quad -1 + x_4 \geq 0, \quad x_4 \geq 0.$$

Итак, $\tilde{P}$ является треугольником с вершинами $(0, 3)$, $(0, 1)$ и $(4, 1)$ (см. рис. 10).

Заметим, что функция

$$\varphi : \ (x_1, x_4) \mapsto (x_1, 6 - x_1 - 2x_4, -1 + x_4, x_4).$$

взаимно однозначно отображает $\tilde{P}$ на P (легко проверить, что она инъективна и сюръективна). Следовательно, полиэдры $\tilde{P}$ и P изоморфны, а значит φ переводит вершины в вершины. Отсюда заключаем, что полиэдр P — это треугольник с вершинами

$$\varphi(0, 3) = (0, 0, 2, 3), \quad \varphi(0, 1) = (0, 4, 0, 1), \quad \varphi(4, 1) = (4, 0, 0, 1);$$

его стороны образованы пересечением линейного многообразия с гиперплоскостями $x_1 = 0$, $x_2 = 0$ и $x_3 = 0$. $\square$

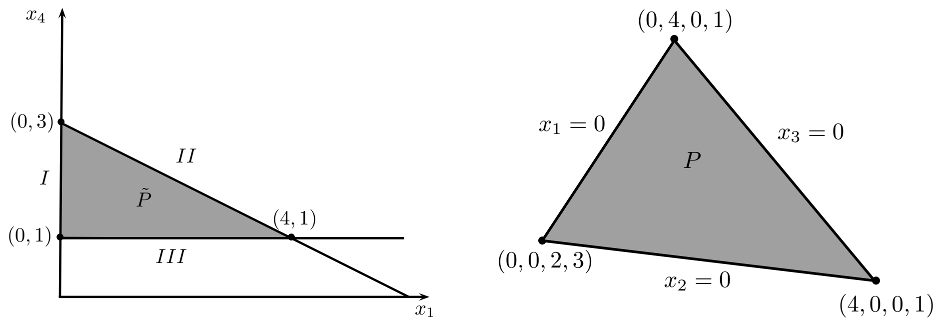


Рис. 10: Полиэдр в канонической форме P и его проекция $\tilde{P}$.

Предположения: Далее мы будем считать, что $\mathbf{x}$ — $(m+n)$-мерный вектор, $\mathbf{A}$ — матрица, состоящая из m строк и $m+n$ столбцов. При этом ранг матрицы $\mathbf{A}$ максимален, то есть равен m.

Отсюда получаем два следствия:

- размерность линейного многообразия $\{\mathbf{x} : \mathbf{Ax} = \mathbf{b}\}$ равна n;
- вершина полиэдра — это точка, в которой пересекаются m гиперплоскостей, составляющих линейное многообразие $\{\mathbf{x} : \mathbf{Ax} = \mathbf{b}\}$, и n гиперплоскостей вида $x_i = 0$ при условии, что все эти $m+n$ гиперплоскостей линейно независимы.

Теорема 4. *(критерий вершины) Пусть P — полиэдр в стандартной форме и $\mathbf{x} \in P$. Тогда следующие утверждения эквивалентны*

(a) точка $\mathbf{x} = (x_1, \dots, x_{m+n})$ является вершиной полиэдра P;

(b) существует набор $B(1), \dots, B(m)$ из m индексов такой, что столбцы $\mathbf{A}_{B(1)}, \dots, \mathbf{A}_{B(m)}$ линейно независимы и $x_i = 0$ при $i \neq B(1), \dots B(m)$.

Доказательство.

(a) ⇒ (b)

Пусть $\mathbf{x}$ — вершина полиэдра. Тогда в точке $\mathbf{x}$ активны m ограничений, определяемых системой $\mathbf{Ax} = \mathbf{b}$, и не менее чем n ограничений вида $x_i = 0$. Пусть $B(1), \dots, B(k)$ — номера неактивных ограничений, то есть $x_{B(1)}, \dots x_{B(k)}$ — компоненты вектора $\mathbf{x}$, отличные от нуля. Заметим, что $k \leq m$.

Поскольку $\mathbf{x}$ — вершина, система

$$\mathbf{Ax} = \mathbf{b}, \quad x_i = 0, \quad i \neq B(1), \dots, B(k),$$

образованная активными ограничениями, должна иметь единственное решение — $\mathbf{x}$. Последнее эквивалентно тому, что система

$$\sum_{i=1}^{k} \mathbf{A}_{B(i)} x_{B(i)} = \mathbf{b}$$

имеет единственное решение. Но для этого матрица $[\mathbf{A}_{B(1)} \cdots \mathbf{A}_{B(k)}]$ должна иметь полный ранг — k, а значит вектора $\mathbf{A}_{B(1)}, \ldots, \mathbf{A}_{B(k)}$ линейно независимы.

Так как $\operatorname{rank} \mathbf{A} = m$, то к векторам $\mathbf{A}_{B(1)}, \ldots, \mathbf{A}_{B(k)}$ можно добавить ещё $m - k$ векторов $\mathbf{A}_{B(k+1)}, \ldots, \mathbf{A}_{B(m)}$ так, чтобы получившийся набор $\mathbf{A}_{B(1)}, \ldots, \mathbf{A}_{B(m)}$ был линейно независимым.

(b) $\Rightarrow$ (a)
В точке $\mathbf{x}$ активны $m + n$ ограничений

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \qquad x_i = 0, \qquad i \neq B(1), \ldots, B(m). \tag{6.1}$$

Если данная система имеет единственное решение, то ограничения, входящие в неё, линейно независимы. Но система (6.1) эквивалентна системе

$$\mathbf{A}\mathbf{x} = \sum_{i=1}^{m} \mathbf{A}_{B(i)} x_{B(i)} = \mathbf{b},$$

а последняя имеет единственное решение в силу линейной независимости векторов $\mathbf{A}_{B(1)}, \ldots, \mathbf{A}_{B(m)}$.

□

Терминология: Пусть $\mathbf{x}$ — вершина, $B(1), \ldots, B(m)$ — набор индексов, удовлетворяющих условию **(b)** теоремы. Тогда мы будем говорить, что

- $B(1), \ldots, B(m)$ — базисные индексы;
- $\mathbf{A}_{B(1)}, \ldots, \mathbf{A}_{B(m)}$ — базисные столбцы;
- $x_{B(1)}, \ldots, x_{B(m)}$ — базисные переменные.

Кроме того, мы введём следующие обозначения:

$$\mathbf{B} = \left[\mathbf{A}_{B(1)} \cdots \mathbf{A}_{B(m)}\right], \qquad \mathbf{x}_B = \begin{bmatrix} x_{B(1)} \\ \vdots \\ x_{B(m)} \end{bmatrix}.$$

Принимая во внимание предыдущую теорему, мы предложим следующую процедуру поиска вершин полиэдра, записанного в стандартной форме.

Процедура поиска вершин

1. Выбрать m линейно независимых столбцов $\mathbf{A}_{B(1)}, \dots, \mathbf{A}_{B(m)}$.
2. Положить $x_i = 0$ при $i \neq B(1), \dots, B(m)$.
3. Найти $x_{B(1)}, \dots, x_{B(m)}$ из системы $\mathbf{B}\mathbf{x}_B = \mathbf{b}$.
4. Проверить выполнение условия $\mathbf{x} \geq 0$.

Пример 10. Найти все вершины полиэдра из примера 9 с помощью приведённой выше процедуры поиска вершин.

Решение:

Запишем матрицу $\mathbf{A}$:

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & 3 \end{bmatrix}.$$

Из столбцов этой матрицы можно составить 6 различных пар, каждая из которых может быть базисом какой-нибудь вершины:

$$\mathbf{A}_1\mathbf{A}_2, \quad \mathbf{A}_1\mathbf{A}_3, \quad \mathbf{A}_1\mathbf{A}_4, \quad \mathbf{A}_2\mathbf{A}_3, \quad \mathbf{A}_2\mathbf{A}_4, \quad \mathbf{A}_3\mathbf{A}_4.$$

Проверим все варианты.

$\mathbf{A}_1\mathbf{A}_2$: Это не базис, так как линейно зависимы.

$\mathbf{A}_1\mathbf{A}_3$: $x_2 = 0$, $x_4 = 0$, $\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 7 \end{bmatrix}$. Откуда $\mathbf{x} = (6, 0, -1, 0) \ngeq 0$.

$\mathbf{A}_1\mathbf{A}_4$: $x_2 = 0$, $x_3 = 0$, $\begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_4 \end{bmatrix} = \begin{bmatrix} 5 \\ 7 \end{bmatrix}$. Откуда $\mathbf{x} = (4, 0, 0, 1) \geq 0$.

$\mathbf{A}_2\mathbf{A}_3$: $x_1 = 0$, $x_4 = 0$, $\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 7 \end{bmatrix}$. Откуда $\mathbf{x} = (0, 6, -1, 0) \ngeq 0$.

$\mathbf{A}_2\mathbf{A}_4$: $x_1 = 0$, $x_3 = 0$, $\begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix} \cdot \begin{bmatrix} x_2 \\ x_4 \end{bmatrix} = \begin{bmatrix} 5 \\ 7 \end{bmatrix}$. Откуда $\mathbf{x} = (0, 4, 0, 1) \geq 0$.

$\mathbf{A}_3\mathbf{A}_4$: $x_1 = 0$, $x_2 = 0$, $\begin{bmatrix} 1 & 1 \\ -1 & 3 \end{bmatrix} \cdot \begin{bmatrix} x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 5 \\ 7 \end{bmatrix}$. Откуда $\mathbf{x} = (0, 0, 2, 3) \geq 0$.

Итак, вершинами полиэдра являются точки $(4, 0, 0, 1)$, $(0, 4, 0, 1)$, $(0, 0, 2, 3)$.
$\square$

Нетрудно понять, что вершина вырождена тогда и только тогда, когда соответствующий ей вектор базисных переменных $\mathbf{x}_B$ содержит нулевые компоненты. Отметим также, что каждой невырожденной вершине соответствует единственный базис, а у вырожденной — базисов может быть несколько.

Пример 11. Рассмотрим два полиэдра:

$$P_1 = \begin{cases} x_1 + \ \ x_2 + x_3 \qquad\quad = 1 \\ x_1 + 2x_2 \qquad\quad - x_4 = 0 \\ x_i \geq 0, \quad i = 1, \ldots, 4, \end{cases}, \quad P_2 = \begin{cases} x_1 + x_2 + x_3 + x_4 = 2 \\ x_1 - x_2 + x_3 + x_4 = 2 \\ x_i \geq 0, \quad i = 1, \ldots, 4. \end{cases}$$

Точка $(0, 0, 1, 0)$ является вырожденной вершиной полиэдра P_1; ей соответствует три базиса: $\mathbf{A}_1\mathbf{A}_3$, $\mathbf{A}_2\mathbf{A}_3$ и $\mathbf{A}_3\mathbf{A}_4$.

Точка $(2, 0, 0, 0)$ является вырожденной вершиной полиэдра P_2; ей соответствует только один базис: $\mathbf{A}_1\mathbf{A}_2$.

§7 Симплекс метод для невырожденного случая

Пусть дана задача линейного программирования в канонической форме:

$$\mathbf{c}'\mathbf{x} \to \min, \quad \mathbf{x} \in P,$$

где

$$P = \{\mathbf{x} : \ \mathbf{A}\mathbf{x} = \mathbf{b}, \ \mathbf{x} \geq 0\}.$$

По-прежнему будем считать, что $\mathbf{A}$ — матрица полного ранга, состоящая из m строк и $m+n$ столбцов. Кроме того предположим, что все вершины полиэдра P невырожденные.

Согласно теореме 2 допустимое множество этой задачи всегда содержит хотя бы одну вершину. В то же время теорема 3 утверждает, что решение задачи (если оно вообще существует) нужно искать среди вершин полиэдра.

Наша цель — описать «умный» алгоритм перебора вершин, который приведёт нас к оптимальному решению. В основе этого алгоритма лежит понятие допустимого направления.

Определение 11. Пусть $\bar{\mathbf{x}} \in P$. Вектор $\mathbf{d}$ называется допустимым направлением в точке $\bar{\mathbf{x}}$, если $\bar{\mathbf{x}} + \varepsilon\mathbf{d} \in P$ при всех достаточно малых $\varepsilon > 0$.

Пусть $\bar{\mathbf{x}}$ — вершина, $\mathbf{A}_{B(1)}, \ldots, \mathbf{A}_{B(m)}$ — её базис. Множества базисных и небазисных индексов обозначим соответственно $\mathcal{B}$ и $\mathcal{N}$. Поскольку вершина $\bar{\mathbf{x}}$ невырожденая, в ней активны m ограничений $\mathbf{Ax} = \mathbf{b}$ и ровно n ограничений $x_i = 0$, $i \in \mathcal{N}$.

Выберем какой-нибудь небазисный индекс $j \in \mathcal{N}$ и определим вектор $\mathbf{d}$ так, чтобы выполнялись равенства

$$d_j = 1, \quad d_i = 0, \quad i \in \mathcal{N} \setminus \{j\}, \quad \mathbf{Ad} = 0. \tag{7.1}$$

Иногда вместо $\mathbf{d}$ мы будем писать $\mathbf{d}^j$, чтобы подчеркнуть зависимость вектора от небазисного индекса j.

Утверждение 3. *Вектор* $\mathbf{d}$ *определён однозначно и является допустимым направлением в невырожденной вершине* $\bar{\mathbf{x}}$.

Доказательство. Из условий (7.1) следует, что

$$\sum_{i \in \mathcal{B}} \mathbf{A}_i d_i + \mathbf{A}_j = 0.$$

Поскольку все базисные столбцы линейно независимы, компоненты d_i, $i \in \mathcal{B}$, определяются однозначно.

Теперь нам нужно показать, что $\bar{\mathbf{x}} + \varepsilon\mathbf{d} \in P$ при малых ε. Из равенства $\mathbf{Ad} = 0$ вытекает, что $\mathbf{A}(\bar{\mathbf{x}} + \varepsilon\mathbf{d}) = \mathbf{b}$. Так как вершина $\bar{\mathbf{x}}$ невырождена, то $\bar{x}_i > 0$ для всех $i \in \mathcal{B}$. Следовательно, $\bar{\mathbf{x}} + \varepsilon\mathbf{d} \geq 0$ для малых ε. $\square$

Пусть $\mathbf{d}_B$ — вектор, состоящий из компонент d_i, которые соответствуют базисным индексам i. Тогда $\mathbf{d}$ вычисляется по формулам

$$d_j = 1, \quad d_i = 0, \quad i \in \mathcal{N} \setminus \{j\}, \quad \mathbf{d}_B = -\mathbf{B}^{-1}\mathbf{A}_j. \tag{7.2}$$

Пример 12. Пусть P — полиэдр из примера 9. Нарисовать допустимые направления, соответствующие небазисным индексам вершины $(4, 0, 0, 1)$.

Решение:

Поскольку вершина $\mathbf{x}$ невырождена, то $\mathcal{B} = \{1, 4\}$, $\mathcal{N} = \{2, 3\}$. Значит

$$\mathbf{B} = \begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix}, \quad \mathbf{B}^{-1} = \begin{bmatrix} \frac{3}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

Для каждого небазисного индекса вычислим допустимое направление.

2: $d_2 = 1$, $d_3 = 0$, $\begin{bmatrix} d_1 \\ d_4 \end{bmatrix} = -\begin{bmatrix} \frac{3}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Значит $\mathbf{d}^2 = (-1, 1, 0, 0)$.

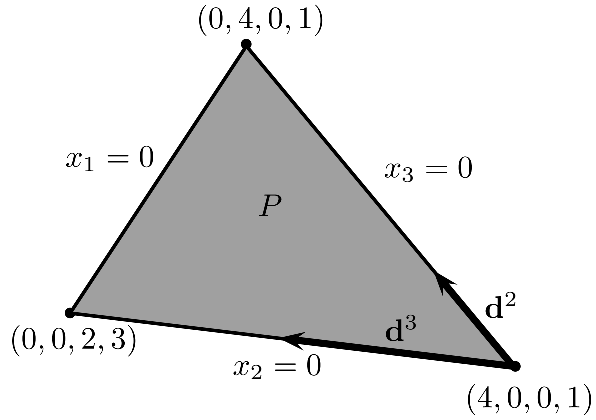


Рис. 11: Направления $\mathbf{d}^2$ и $\mathbf{d}^3$.

3: $d_3 = 1$, $d_2 = 0$, $\begin{bmatrix} d_1 \\ d_4 \end{bmatrix} = -\begin{bmatrix} \frac{3}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \end{bmatrix}$. Значит $\mathbf{d}^3 = (-2, 0, 1, 1)$.

Заметим, что

$(4, 0, 0, 1) + 4(-1, 1, 0, 0) = (0, 4, 0, 1), \quad (4, 0, 0, 1) + 2(-2, 0, 1, 1) = (0, 0, 2, 3).$

Следовательно вектора $\mathbf{d}^2$ и $\mathbf{d}^3$ направлены вдоль рёбер полиэдра (рис. 11). □

Тот факт, что вектор $\mathbf{d}$, определённый формулой (7.2), оказался направлен вдоль ребра полиэдра, не случаен.

Утверждение 4. *Вектор $\mathbf{d}^j$ направлен вдоль ребра полиэдра, которое получается отбрасыванием ограничения $x_j = 0$.*

Доказательство. Достаточно заметить, что при движении из $\bar{\mathbf{x}}$ в направлении вектора $\mathbf{d}^j$ активен набор из $m + n - 1$ линейно независимых ограничений: $\mathbf{Ax} = \mathbf{b}$, $x_i = 0$, $i \in \mathcal{N} \setminus \{j\}$. □

Поставим следующий вопрос: что происходит с целевой функцией при движении вдоль направления $\mathbf{d}$? Введём обозначение

$$\bar{c}_j = \mathbf{c}'\mathbf{d}.$$

Ясно, что если $\bar{c}_j > 0$, то целевая функция возрастает вдоль направления $\mathbf{d}$, а если $\bar{c}_j < 0$, то убывает.

Следовательно при $\bar{c}_j < 0$ значение целевой функции можно уменьшить, двигаясь из вершины $\bar{\mathbf{x}}$ вдоль допустимого направления $\mathbf{d}$. Как долго можно продолжать движение? Здесь возможны два случая:

- Если $\mathbf{d} \geq 0$, то $\mathbf{x}+\varepsilon\mathbf{d} \geq 0$ для всех $\varepsilon > 0$. Это значит, что $\mathbf{x}+\varepsilon\mathbf{d} \in P$ для всех $\varepsilon > 0$. Поскольку $\mathbf{c}'(\mathbf{x}+\varepsilon\mathbf{d}) \to -\infty$ при $\varepsilon \to +\infty$, то задача не имеет решения.

- Если для $d_l < 0$ некоторого $l \in \mathcal{B}$, то движение можно продолжать до точки $\mathbf{x}^* = \bar{\mathbf{x}} + \varepsilon^*\mathbf{d}$, где

$$\varepsilon^* = \max\{\varepsilon : \bar{\mathbf{x}} + \varepsilon\mathbf{d} \geq 0\}.$$

Утверждение 5. *Точка $\mathbf{x}^* = \bar{\mathbf{x}} + \varepsilon^*\mathbf{d}$ является вершиной полиэдра P.*

Доказательство. Во всех точках луча $\bar{\mathbf{x}} + \varepsilon\mathbf{d}$ активны ограничения

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \quad x_i = 0, \quad i \in \mathcal{N} \setminus \{j\}.$$

Кроме того, имеется ограничение $x_l = 0$, неактивное в $\bar{\mathbf{x}}$, но активное в $\mathbf{x}^* = \bar{\mathbf{x}} + \varepsilon^*\mathbf{d}$. Тогда, используя те же рассуждения, что и при доказательстве теоремы 2, мы заключаем, что $m + n$ ограничений

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \quad x_i = 0, \quad i \in \mathcal{N} \cup \{l\} \setminus \{j\}$$

линейно независимы и активны в точке $\mathbf{x}^*$, а потому $\mathbf{x}^*$ — вершина. □

Может однако случиться, что $\bar{c}_j \geq 0$ для всех $j \in \mathcal{N}$. Оказывается, что в этом случае вершина $\bar{\mathbf{x}}$ оптимальна.

Утверждение 6. *Невырожденная вершина $\bar{\mathbf{x}}$ оптимальна тогда и только тогда, когда $\bar{c}_j \geq 0$ для всех $j \in \mathcal{N}$.*

Доказательство.

$\Rightarrow$:
Так как $\bar{\mathbf{x}}$ — невырожденная вершина, то все направления $\mathbf{d}^j$, $j \in \mathcal{N}$, допустимы. Из оптимальности вершины $\bar{\mathbf{x}}$ следует, что $\bar{c}_j = \mathbf{c}'\mathbf{d}^j \geq 0$ для всех $j \in \mathcal{N}$.
$\Leftarrow$:
Пусть $\mathbf{y}$ — произвольная точка полиэдра, обозначим $\mathbf{d} = \mathbf{y}-\bar{\mathbf{x}}$. Поскольку обе точки принадлежат линейному многообразию, то $\mathbf{A}\mathbf{d} = 0$ или, что то же самое,

$$\mathbf{B}\mathbf{d}_B + \sum_{j\in\mathcal{N}} \mathbf{A}_j d_j = 0.$$

Отсюда получаем, что

$$\mathbf{d}_B = -\sum_{j\in\mathcal{N}} \mathbf{B}^{-1}\mathbf{A}_j d_j$$

и

$$\mathbf{c}'\mathbf{d} = \mathbf{c}_B\mathbf{d}_B + \sum_{j\in\mathcal{N}} c_j d_j = \sum_{j\in\mathcal{N}} \left(c_j - \mathbf{c}_B\mathbf{B}^{-1}\mathbf{A}_j\right) d_j = \sum_{j\in\mathcal{N}} \bar{c}_j d_j.$$

Поскольку $\bar{x}_i = 0$ для всех $j \in \mathcal{N}$, то $d_j \geq 0$ для всех $j \in \mathcal{N}$ (иначе точка $\mathbf{y}$ была бы недопустима). Учитывая неравенства $\bar{c}_j \geq 0$, $j \in \mathcal{N}$, мы получаем, что $\mathbf{c}'\mathbf{d} \geq 0$. Так как последнее неравенство справедливо для произвольного $\mathbf{y} \in P$, то $\bar{\mathbf{x}}$ оптимально. $\square$

Теперь мы можем записать алгоритм целиком.

Симплекс метод

1. В начале нам дана вершина $\bar{\mathbf{x}}$, множества базисных и небазисных индексов $\mathcal{B}$, $\mathcal{N}$.
2. Для каждого $j \in \mathcal{N}$ вычисляем вектор $\mathbf{d}^j$ по формулам $\mathbf{d}^j_B = -\mathbf{B}^{-1}\mathbf{A}_j$, $d^j_j = 1$, $d^j_i = 0$, $i \in \mathcal{N} \setminus \{j\}$.
3. Для каждого $j \in \mathcal{N}$ вычисляем $\bar{c}_j = \mathbf{c}'\mathbf{d}^j$.
4. Если $\bar{c}_j \geq 0$ для всех $j \in \mathcal{N}$, то алгоритм прерывается — мы в оптимальной вершине.
5. Выбираем какой-нибудь индекс $j \in \mathcal{N}$, для которого $\bar{c}_j < 0$.
6. Если $\mathbf{d}^j \geq 0$, то алгоритм прерывается — решения нет.
7. Вычисляем
$$\varepsilon^* = \max\{\varepsilon : \bar{\mathbf{x}} + \varepsilon\mathbf{d} \geq 0\}.$$
8. Точка $\mathbf{x}^* = \bar{\mathbf{x}} + \varepsilon^*\mathbf{d}^j$ — это новая вершина.

Пример 13. Минимизировать функцию

$$-x_2 + x_3 + x_4$$

на полиэдре из примера 9, стартуя из вершины $(4, 0, 0, 1)$.
Решение:
Прежде всего выпишем вектор $\mathbf{c}$ и матрицу $\mathbf{A}$:

$$\mathbf{c} = (0, -1, 1, 1), \quad \mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & 3 \end{bmatrix}.$$

Итерация 1.

1. Опорная вершина: $\bar{\mathbf{x}} = (4, 0, 0, 1)$. Множества базисных и небазисных индексов: $\mathcal{B} = \{1, 4\}$, $\mathcal{N} = \{2, 3\}$.

2. Допустимые направления уже найдены: $\mathbf{d}^2 = (-1, 1, 0, 0)$, $\mathbf{d}^3 = (-2, 0, 1, 1)$.

3. Вычисляем оценки: $\bar{c}_2 = -1 < 0$, $\bar{c}_3 = 2 \geq 0$. Следовательно, $\mathbf{d}^2$ — направление убывания.

4. Теперь $\bar{\mathbf{x}} + \varepsilon \mathbf{d}^2 = (4 - \varepsilon, \varepsilon, 0, 1)$. Значит $\varepsilon^* = 4$ и $\mathbf{x}^* = (0, 4, 0, 1)$.

Итерация 2.

1. Теперь у нас новая опорная вершина: $\bar{\mathbf{x}} = (0, 4, 0, 1)$. Множества базисных и небазисных индексов: $\mathcal{B} = \{2, 4\}$, $\mathcal{N} = \{1, 3\}$.

2. Вычисляем допустимые направления $\mathbf{d}^1$ и $\mathbf{d}^3$. Здесь
$$B = \begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix}, \quad B^{-1} = \begin{bmatrix} \frac{3}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$
$\mathbf{d}^1$: $d_1 = 1, d_3 = 0$, $\begin{bmatrix} d_2 \\ d_4 \end{bmatrix} = - \begin{bmatrix} \frac{3}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Откуда $\mathbf{d}^1 = (1, -1, 0, 0)$.

   $\mathbf{d}^3$: $d_3 = 1$, $d_1 = 0$, $\begin{bmatrix} d_2 \\ d_4 \end{bmatrix} = - \begin{bmatrix} \frac{3}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \end{bmatrix}$. Откуда $\mathbf{d}^3 = (0, -2, 1, 1)$.

3. Вычисляем оценки: $\bar{c}_1 = 1 \geq 0$, $\bar{c}_3 = 4 \geq 0$. Следовательно, $\bar{\mathbf{x}} = (0, 4, 0, 1)$ — оптимальное решение. Алгоритм завершён.

$\square$

§8 Симплекс метод для вырожденного случая

Предположим теперь, что P содержит вырожденные вершины. Модифицируем предложенный выше алгоритм так, чтобы он работал и в этом случае.

Основная трудность состоит в том, что у вырожденных вершин может быть неединственный базис. Отсюда первая модификация:

- На первом шаге алгоритма необходимо зафиксировать один из возможных базисов (если их несколько).

У множественности базисов есть другое, неочевидное, следствие. А именно: направление $\mathbf{d}$, соответствующее небазисному индексу, может оказаться недопустимым. Действительно, у вырожденных вершин хотя бы одна базисная переменная равна нулю. Пусть, например, $\bar{x}_i = 0$ для некоторого $i \in \mathcal{B}$. Если при этом $d_i < 0$, то направление $\mathbf{d}$ недопустимо.

Как следствие, на шаге 7 мы получим $\varepsilon^* = 0$, а значит алгоритм не даёт нам новой вершины. В этом случае, чтобы двигаться дальше, потребуется ещё одна модификация:

- Если $\varepsilon^* = 0$, то необходимо перейти на шаг 1 и сменить базис.

Пример 14. Дана задача линейного программирования

$$\begin{aligned} -x_1 + x_2 + x_3 + x_4 &\to \min \\ x_1 + \;\; x_2 + x_3 \qquad &= 1 \\ x_1 + 2x_2 \qquad - x_4 &= 0 \\ x_i \geq 0, \quad i = 1, \ldots, 4, \end{aligned}$$

Улучшить вершину $(0, 0, 1, 0)$ с помощью симплекс метода.
Решение:

Выпишем вектор $\mathbf{c}$ и матрицу $\mathbf{A}$:

$$\mathbf{c} = (-1, 1, 1, 1), \quad \mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 2 & 0 & -1 \end{bmatrix}.$$

Итерация 1.

1. Опорная вершина: $\bar{\mathbf{x}} = (0, 0, 1, 0)$. В качестве базиса возьмём $\mathbf{A}_2\mathbf{A}_3$, то есть $\mathcal{B} = \{2, 3\}$, $\mathcal{N} = \{1, 4\}$.

2. Вычисляем допустимые направления $\mathbf{d}^1$ и $\mathbf{d}^4$. Здесь

$$B = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}, \quad B^{-1} = \begin{bmatrix} 0 & \frac{1}{2} \\ 1 & -\frac{1}{2} \end{bmatrix}.$$

$$\mathbf{d}^1\colon\; d_1 = 1,\, d_4 = 0,\; \begin{bmatrix} d_2 \\ d_3 \end{bmatrix} = -\begin{bmatrix} 0 & \frac{1}{2} \\ 1 & -\frac{1}{2} \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} \Rightarrow \mathbf{d}^1 = \left(1, -\tfrac{1}{2}, -\tfrac{1}{2}, 0\right).$$

$$\mathbf{d}^4\colon\ d_4 = 1,\ d_1 = 0,\ \begin{bmatrix} d_2 \\ d_3 \end{bmatrix} = -\begin{bmatrix} 0 & \frac{1}{2} \\ 1 & -\frac{1}{2} \end{bmatrix} \cdot \begin{bmatrix} 0 \\ -1 \end{bmatrix} \Rightarrow \mathbf{d}^4 = \left(0, \tfrac{1}{2}, -\tfrac{1}{2}, 1\right).$$

3. Вычисляем оценки: $\bar{c}_1 = -\frac{3}{2} < 0$, $\bar{c}_4 = 1 \geq 0$. Следовательно, $\mathbf{d}^1$ — направление убывания.

4. Теперь $\bar{\mathbf{x}} + \varepsilon\mathbf{d}^1 = (\varepsilon, -\frac{\varepsilon}{2}, 1 - \frac{\varepsilon}{2}, 0)$. Значит $\varepsilon^* = 0$.

Итерация 2.

1. Опорная вершина по-прежнему: $\bar{\mathbf{x}} = (0, 0, 1, 0)$. Сменим базис на $\mathbf{A}_3\mathbf{A}_4$, то есть $\mathcal{B} = \{3, 4\}$, $\mathcal{N} = \{1, 2\}$.

2. Вычисляем допустимые направления $\mathbf{d}^1$ и $\mathbf{d}^2$. Здесь

$$B = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad B^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

$$\mathbf{d}^1\colon\ d_1 = 1,\ d_2 = 0,\ \begin{bmatrix} d_3 \\ d_4 \end{bmatrix} = -\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} \Rightarrow \mathbf{d}^1 = (1, 0, -1, 1).$$

$$\mathbf{d}^2\colon\ d_2 = 1,\ d_1 = 0,\ \begin{bmatrix} d_3 \\ d_4 \end{bmatrix} = -\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix} \Rightarrow \mathbf{d}^2 = (0, 1, -1, 2).$$

3. Вычисляем оценки: $\bar{c}_1 = -1 < 0$, $\bar{c}_2 = 2 \geq 0$. Следовательно, $\mathbf{d}^1$ — направление убывания.

4. Теперь $\bar{\mathbf{x}} + \varepsilon\mathbf{d}^1 = (\varepsilon, 0, 1 - \varepsilon, \varepsilon)$. Значит $\varepsilon^* = 1$ и $\mathbf{x}^* = (1, 0, 0, 1)$. Новая вершина найдена.

$\square$

§9 Теорема об отделимости

Два множества A и B называются строго отделимыми, если существует такое замкнутое полупространство H^+, что $A \subset H^+$ и $B \cap H^+ = \varnothing$.

Теорема 5. *Замкнутое выпуклое множество C из $\mathbb{R}^n$ можно строго отделить от точки a, ему не принадлежащей.*

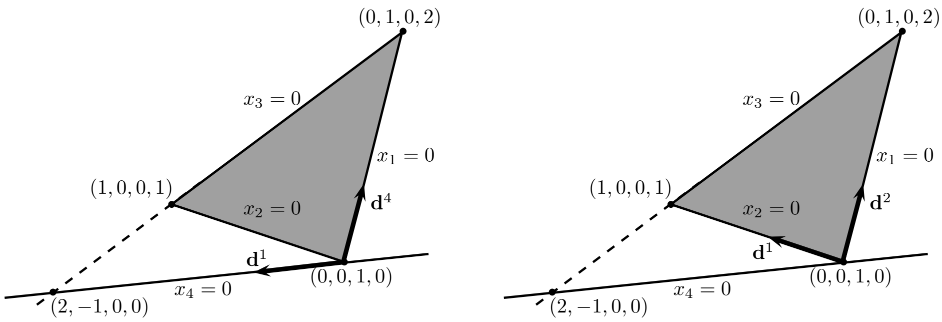


Рис. 12: Первая и вторая итерация симплекс метода. На первой итерации из-за неудачного выбора базиса направление $\mathbf{d}^1$ является недопустимым и указывает на фиктивную вершину $(2, -1, 0, 0)$.

Доказательство. 1) Положим $f(x) = \|x - a\|$. Покажем, что задача

$$f(x) \to \min, \quad x \in C \tag{9.1}$$

имеет решение. Ясно, что f непрерывна. Значит, если C ограничено, то существование решения следует из теоремы Вейерштрасса. Если же C неограничено, то заменим C на пересечение $C \cap B(a, r)$, где $B(a, r)$ — шар радиуса r, описанный вокруг точки a. Задача

$$f(x) \to \min, \quad x \in C \cap B(a, r)$$

имеет решение в силу теоремы Вейерштрасса. Обозначим это решение через x_*. Тогда

$$f(x_*) \begin{cases} \le f(x) & x \in C \cap B(a, r), \\ = \|x_* - a\| \le r \le \|x - a\| = f(x), & x \in C \setminus B(a, r). \end{cases}$$

Следовательно, x_* — решение задачи (9.1).

2) Покажем, что полупространство $H = \{x \colon c'x \ge b\}$, где

$$c = x_* - a, \quad b = \frac{1}{2}\Big(\|x_*\|^2 - \|a\|^2\Big),$$

удовлетворяет условиям теоремы.

Если $a \in H$, то

$$\langle x_* - a, a \rangle \ge \frac{1}{2}\Big(\|x_*\|^2 - \|a\|^2\Big).$$

Следовательно,

$$\|x_* - a\|^2 \le 0.$$

Откуда $a = x_* \in C$ и мы приходим к противоречию.

Покажем, что $C \subset H$. Возьмём произвольную точку $x \in C$, тогда в силу выпуклости множества C, имеем

$$x_* + \lambda(x - x_*) \in C, \quad \lambda \in (0, 1].$$

Поскольку x_* доставляет минимум f на C, получаем

$$\begin{aligned}\|x_* - a\|^2 &\leq \|x_* + \lambda(x - x_*) - a\|^2 \\ &= \|x_* - a\|^2 + 2\lambda\langle x - x_*, x_* - a\rangle + \lambda^2\|x - x_*\|^2.\end{aligned}$$

Разделив всё на λ и устремив λ к нулю, получаем

$$\langle x - x_*, x_* - a\rangle \geq 0.$$

Значит

$$\langle c, x\rangle = \langle x_* - a, x\rangle \geq \langle x_*, x_* - a\rangle \geq \frac{1}{2}\Big(\|x_*\|^2 - \|a\|^2\Big) = b.$$

Доказательство завершено. $\square$

В качестве приложения мы докажем лемму Фаркаша

Лемма 1. *(Фаркаш) Пусть $A \in M^{m,n}$ и $b \in \mathbb{R}^m$. Только одно из следующих множеств непусто:*

1) $P = \{x \in \mathbb{R}^n : \ Ax = b, \ x \geq 0\}$,

2) $Q = \{y \in \mathbb{R}^m : \ A'y \geq 0, \ b'y < 0\}$.

Чтобы доказать лемму Фаркаша, нужно всего лишь понять её геометрический смысл. Сделать это нам поможет следующее определение.

Определение 12. Конической оболочкой множества S называется множество

$$\mathrm{cone}(S) = \left\{\sum_{i=1}^{n} \alpha_i x_i \ : \ \alpha_i \geq 0, \ x_i \in S\right\}.$$

Доказательство. Утверждение $P \neq \varnothing$ означает, что b принадлежит конической оболочке столбцов матрицы A:

$$b \in \mathrm{cone}(A_1, \ldots, A_n).$$

Утверждение $Q \neq \varnothing$ означает, что b и $\mathrm{cone}(A_1, \ldots, A_n)$ строго отделимы. Действительно, возьмём $y \in Q$ и положим $H^+ = \{q \ : \ q'y \geq 0\}$. Из определения множества Q следует, что $b \notin H^+$. С другой стороны, пусть

$$q = Ax, \quad x \geq 0.$$

Тогда

$$q'y = (Ax)'y = x'A'y = \sum_{i=1}^{n}(A'y)_i x_i \geq 0,$$

то есть $q \in H^+$.

Таким образом, лемма утверждает, что либо $b \in \mathrm{cone}(A_1, \ldots A_n)$, либо b и $\mathrm{cone}(A_1, \ldots, A_n)$ строго отделимы. Доказательство этого утверждения будет следовать из теоремы об отделимости, как только мы установим, что коническая оболочка

$$C = \mathrm{cone}(A_1, \ldots, A_n) = \{Ax \ : \ x \geq 0\}$$

выпукла и замкнута. Заметим, что множество C есть образ полиэдра $x \geq 0$ под действием линейного преобразования A. Следовательно, множество C само является полиэдром (это нетривиальный факт!). Полиэдр же, как мы знаем, есть выпуклое и замкнутое множество. $\square$

Следствие 1. *Пусть $A_1, \ldots, A_n$ и b — заданные вектора и имеет место импликация*

$$\forall i = 1, \ldots, n \quad A_i'y \geq 0 \quad \Rightarrow \quad y'b \geq 0.$$

Тогда b принадлежит конической оболочке $\mathrm{cone}(A_1, \ldots, A_n)$.

Доказательство. Рассмотрим матрицу A, в которой вектора A_i являются столбцами и применим лемму Фаркаша. В ней $Q = \varnothing$. $\square$

§10 Двойственность

Теорема 6. *Сильная двойственность.*

Доказательство. Пусть x^* — оптимальная вершина, $I = \{i \ : \ a_i'x^* = b_i\}$ — индексы активных ограничений. Покажем, что имеет место импликация

$$\forall i \in I \quad a_i'd \geq 0 \quad \Rightarrow \quad c'd \geq 0.$$

Пусть d удовлетворяет неравенствам в левой части импликации. Тогда

$$a_i'(x^* + \varepsilon d) \geq a_i x^* = b_i \quad \forall i \in I,$$

где ε — произвольное положительное число. С другой стороны, если ε мало, то

$$a_i' x^* > b_i \quad \Rightarrow \quad a_i'(x^* + \varepsilon d) > b_i$$

для всех $i \notin I$. Следовательно, $x^* + cd$ — допустимое решение. Теперь из оптимальности x^* следует, что $c'd \geq 0$.

Согласно лемме Фаркаша, существуют $y_i \geq 0$ такие, что

$$c = \sum_{i \in I} y_i a_i$$

Для всех $i \notin I$ положим $y_i = 0$. Тогда $c = A'y$ и $y \geq 0$, то есть y — допустимая точка в сопряжённой задаче. При этом

$$y'b = \sum_{i \in I} y_i b_i = \sum_{i \in I} y_i a_i' x^* = c' x^*.$$

□

Снова рассмотрим задачу линейного программирования в канонической форме:

$$\begin{gathered} \mathbf{c}'\mathbf{x} \to \min, \\ \mathbf{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{x} \geq 0. \end{gathered} \tag{$\mathcal{P}$}$$

Составим лагранжиан

$$L(\mathbf{x}, \mathbf{y}) = \mathbf{c}'\mathbf{x} + \mathbf{y}'(\mathbf{b} - \mathbf{A}\mathbf{x}), \qquad \mathbf{x} \in \mathbb{R}^{m+n}, \ \mathbf{y} \in \mathbb{R}^m.$$

Тем самым мы штрафуем ограничение $\mathbf{A}\mathbf{x} = \mathbf{b}$ в задаче $(\mathcal{P})$, причём $\mathbf{y}$ определяет цену, которую мы платим за нарушение этого ограничения.

Рассмотрим теперь семейство задач

$$\begin{gathered} L(\mathbf{x}, \mathbf{y}) \to \min, \\ \mathbf{x} \geq 0. \end{gathered} \tag{$\mathcal{P}_\mathbf{y}$}$$

При каждом фиксированном $\mathbf{y}$ задача $(\mathcal{P}_\mathbf{y})$ может быть легко решена. Действительно, обозначив через $\varphi(\mathbf{y})$ минимальное значение целевой функции в задаче $(\mathcal{P}_\mathbf{y})$, мы получим

$$\begin{aligned} \varphi(\mathbf{y}) &= \min_{\mathbf{x} \geq 0} L(\mathbf{x}, \mathbf{y}) \\ &= \mathbf{y}'\mathbf{b} + \min_{\mathbf{x} \geq 0} \left\{ (\mathbf{c}' - \mathbf{y}'\mathbf{A})\mathbf{x} \right\} \\ &= \mathbf{y}'\mathbf{b} + \begin{cases} 0, & \text{если } \mathbf{c}' - \mathbf{y}'\mathbf{A} \geq 0, \\ -\infty, & \text{иначе.} \end{cases} . \end{aligned}$$

Легко видеть, что для любого $\mathbf{y}$ справедлива оценка

$$\varphi(\mathbf{y}) \le \min(\mathcal{P}),$$

а значит

$$\max_{\mathbf{y}} \varphi(\mathbf{y}) \le \min(\mathcal{P}). \tag{10.1}$$

Но, ввиду найденного ранее выражения для $\varphi(\mathbf{y})$, задача $\max_{\mathbf{y}} \varphi(\mathbf{y})$ есть не что иное, как задача линейного программирования вида

$$\begin{gathered} \mathbf{y}'\mathbf{b} \to \max, \\ \mathbf{y}'\mathbf{A} \le \mathbf{c}'. \end{gathered} \tag{$\mathcal{P}^*$}$$

Определение 13. Задача $(\mathcal{P}^*)$ называется *двойственной* к исходной (или прямой) задаче $(\mathcal{P})$.

Формула (10.1) устанавливает взаимосвязь между прямой и двойственной задачей:

$$\max(\mathcal{P}^*) \le \min(\mathcal{P}). \tag{10.2}$$

Отсюда мы немедленно получаем следующее утверждение.

Утверждение 7. *Пусть $\mathbf{x}^*$ — допустимая точка в задаче $(\mathcal{P})$, $\mathbf{y}^*$ — допустимая точка в задаче $(\mathcal{P}^*)$ и $\mathbf{c}'\mathbf{x}^* = \mathbf{b}'\mathbf{y}^*$. Тогда $\mathbf{x}^*$ — оптимальное решение задачи $(\mathcal{P})$, а $\mathbf{y}^*$ — оптимальное решение задачи $(\mathcal{P}^*)$.*

Оказывается, что справедливо более сильное утверждение

Теорема 7. *Если $(\mathcal{P})$ имеет оптимальное решение, то $(\mathcal{P}^*)$ также имеет оптимальное решение и*

$$\max(\mathcal{P}^*) = \min(\mathcal{P}).$$

Доказательство. Пусть $\mathbf{x}$ — оптимальное решение задачи $(\mathcal{P})$. Мы докажем теорему в предположении, что $\mathbf{x}$ — невырожденная вершина. В этом случае

$$\bar{c}_j = \mathbf{c}'\mathbf{d}^j = c_j - \mathbf{c}_B\mathbf{B}^{-1}\mathbf{A}_j \ge 0, \quad j \in \mathcal{N}.$$

Теми же формулами можно определить $\bar{c}_j$ для небазисных индексов. В этом случае мы получим

$$\bar{c}_j = c_j - \mathbf{c}'_B\mathbf{B}^{-1}\mathbf{A}_j = 0, \quad j \in \mathcal{B}.$$

Составив вектор $\bar{\mathbf{c}}$ из всех $\bar{c}_j$, мы получим

$$\bar{\mathbf{c}}' = \mathbf{c}' - \mathbf{c}'_B\mathbf{B}^{-1}A \ge 0.$$

Положим $\mathbf{y}' = \mathbf{c}'_B\mathbf{B}^{-1}$. Тогда $\mathbf{y}'\mathbf{A} \le \mathbf{c}'$ и, следовательно, $\mathbf{y}$ — допустимая точка в задаче $(\mathcal{P}^*)$. Кроме того,

$$\mathbf{y}'\mathbf{b} = \mathbf{c}'_B\mathbf{B}^{-1}\mathbf{b} = \mathbf{c}'_B\mathbf{x}_B = \mathbf{c}'\mathbf{x}.$$

□

Теневые цены

Пусть $\mathbf{x}^*$ — оптимальная невырожденная вершина в $(\mathcal{P})$, $\mathbf{B}$ — соответствующая базисная матрица, $\mathbf{x}_B = \mathbf{B}^{-1}\mathbf{b}$ — вектор базисных переменных. Заменим $\mathbf{b}$ на вектор $\mathbf{b}+\mathbf{d}$, где $\mathbf{d}$ — малая добавка. Так как $\mathbf{B}^{-1}\mathbf{b} > 0$, то $\mathbf{B}^{-1}(\mathbf{b}+\mathbf{d}) > 0$ при малых $\mathbf{d}$. Следовательно, тот же базис приводит к вершине в возмущённой задаче. Значит на вектор $\bar{\mathbf{c}}$ такое возмущение не влияет. Следовательно, оптимальный базис в возмущённой задаче тот же. Тогда оптимальная цена в возмущённой задаче

$$\mathbf{c}_B'\mathbf{B}^{-1}(\mathbf{b}+\mathbf{d}) = \mathbf{y}'(\mathbf{b}+\mathbf{d}).$$