

## 1 Выпуклые множества

### 1.1 Определение и основные примеры

**Определение 1** (Выпуклые множества). Пусть  $U$  — вещественное векторное пространство,  $Q$  — подмножество  $U$ . Множество  $Q$  называется *выпуклым*, если для любых двух точек  $x, y$  из множества  $Q$  и любого  $\lambda \in [0, 1]$  точка  $\lambda x + (1 - \lambda)y$  также принадлежит множеству  $Q$ .

Другими словами, множество  $Q$  называется выпуклым, если для каждой пары точек  $x, y \in Q$ , множество  $Q$  также содержит весь *отрезок*  $[x, y] := \{\lambda x + (1 - \lambda)y : 0 \leq \lambda \leq 1\}$ . Точка вида  $\lambda x + (1 - \lambda)y$  для  $\lambda \in [0, 1]$  называется *выпуклой комбинацией* точек  $x, y$ .

**Пример 1** (Тривиальные выпуклые множества). Все пространство  $U$ , множество из одного элемента  $\{a\}$  (где  $a \in U$ ) и пустое множество  $\emptyset$  являются выпуклыми.<sup>1</sup>

**Пример 2** (Множество решений системы линейных ограничений). Пусть в пространстве  $U$  задано (произвольное) скалярное произведение  $\langle \cdot, \cdot \rangle$ . Пусть  $\mathcal{A}$  — произвольное (не обязательно конечное и не обязательно счетное) индексное множество, и пусть для каждого индекса  $\alpha \in \mathcal{A}$  заданы вектор  $a_\alpha \in U$  и скаляр  $b_\alpha \in \mathbb{R}$ . Рассмотрим систему линейных неравенств

$$\langle a_\alpha, x \rangle \leq b_\alpha \quad \text{для всех } \alpha \in \mathcal{A}, \quad (1)$$

где  $x \in U$ . Тогда соответствующее множество всевозможных решений системы (1), т. е. множество

$$Q := \{x \in U : \langle a_\alpha, x \rangle \leq b_\alpha \text{ для всех } \alpha \in \mathcal{A}\}$$

является выпуклым.

Действительно, пусть  $x, y \in Q$  и  $\lambda \in [0, 1]$ . Тогда для любого  $\alpha \in \mathcal{A}$  выполняется

$$\langle a_\alpha, \lambda x + (1 - \lambda)y \rangle = \lambda \langle a_\alpha, x \rangle + (1 - \lambda) \langle a_\alpha, y \rangle \leq \lambda b_\alpha + (1 - \lambda)b_\alpha = b_\alpha.$$

<sup>1</sup>Последнее следует из того, что не существует такой пары точек  $x, y \in \emptyset$ , для которой отрезок  $[x, y]$  не принадлежит множеству — для пустого множества нельзя предъявить даже точку  $x \in \emptyset$ .

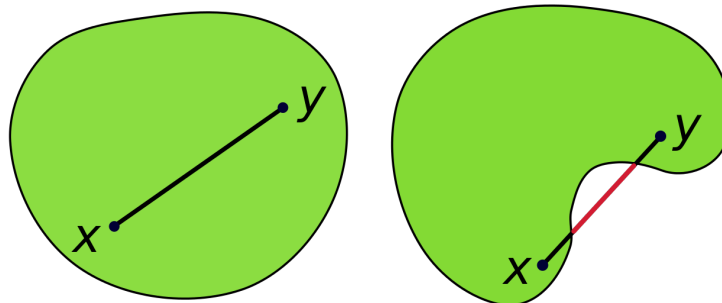


Рис. 1: Иллюстрация к определению выпуклого множества из Википедии. Слева: выпуклое множество; справа: невыпуклое множество.

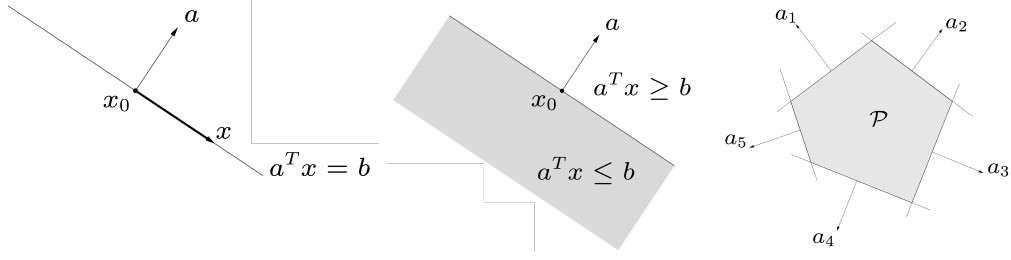


Рис. 2: Иллюстрация к примерам. Слева направо: гиперплоскость, полупространство, полиэдр.

Таким образом,  $\lambda x + (1 - \lambda)y \in Q$ .

Заметим, что в этом примере абсолютно не принципиально, что в системе линейных ограничений (1) используется именно неравенство  $\leq$ . Нетрудно видеть, что если некоторые (возможно, даже все) из неравенств  $\leq$  заменить на  $\geq$  или  $=$  (или  $<$ , или  $>$ ), то множество  $Q$  по-прежнему останется выпуклым. Таким образом, *множество решений произвольной системы линейных ограничений является выпуклым*.

Рассмотрим наиболее популярные частные случаи. Во всех примерах мы работаем в пространстве  $U := \mathbb{R}^n$  со скалярным произведением  $\langle x, y \rangle := x^T y$ .

1. *Гиперплоскость*, т. е. множество вида

$$\{x \in \mathbb{R}^n : a^T x = b\},$$

где  $a \in \mathbb{R}^n \setminus \{0\}$  и  $b \in \mathbb{R}$ , является выпуклым.

2. *Полупространство*, т. е. множество вида

$$\{x \in \mathbb{R}^n : a^T x \leq b\},$$

где  $a \in \mathbb{R}^n \setminus \{0\}$  и  $b \in \mathbb{R}$ , является выпуклым.

3. *Полиэдр*, т. е. множество вида

$$\{x \in \mathbb{R}^n : Ax \preceq b, Cx = d\},$$

где  $A \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^m$  и  $C \in \mathbb{R}^{p \times n}$ ,  $d \in \mathbb{R}^p$ , является выпуклым. (Здесь символ  $\preceq$  нужно понимать как поэлементное неравенство  $\leq$ .) Полиэдр представляет собой пересечение конечного числа полупространств и гиперплоскостей.

4. Рассмотрим особый случай полиэдра:

$$\Delta_n := \left\{ x \in \mathbb{R}^n : x \succeq 0, \sum_{i=1}^n x_i = 1 \right\}.$$

Это множество называется *стандартным симплексом* и является выпуклым.

**Пример 3** (Шар в произвольной норме). Пусть в пространстве  $U$  задана (произвольная) норма  $\|\cdot\|$ . Тогда (замкнутый) шар радиуса  $r > 0$  с центром в точке  $a \in U$ , т. е. множество

$$B_{\|\cdot\|}(a, r) := \{x \in U : \|x - a\| \leq r\},$$

является выпуклым.

Действительно, пусть  $x, y \in B_{\|\cdot\|}(a, r)$  и  $\lambda \in [0, 1]$ . Тогда

$$\|\lambda x + (1 - \lambda)y - a\| = \|\lambda(x - a) + (1 - \lambda)(y - a)\| \leq \lambda\|x - a\| + (1 - \lambda)\|y - a\| \leq \lambda r + (1 - \lambda)r = r.$$

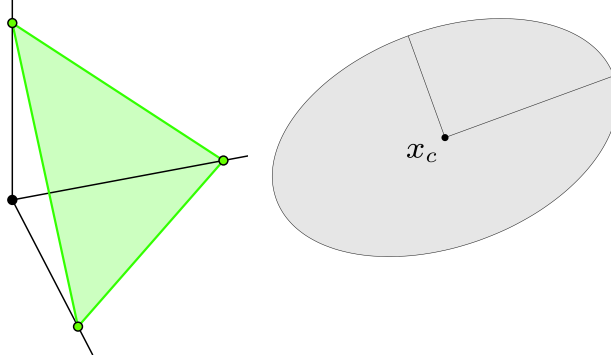


Рис. 3: Иллюстрация к примерам. Слева направо: стандартный симплекс, эллипсоид.

Здесь неравенство следует из неравенства треугольника для нормы и условия  $\lambda \in [0, 1]$ . Таким образом,  $\lambda x + (1 - \lambda)y \in B_{\|\cdot\|}(a, r)$ .

Рассмотрим наиболее популярные частные случаи. Во всех примерах мы работаем в пространстве  $U := \mathbb{R}^n$ .

1. (Евклидов шар) Пусть  $\|x\| := \|x\|_2 := \sqrt{x^T x}$  — евклидова норма. В этом случае рассмотренный выше абстрактный шар переходит в множество

$$B_2(a, r) := \{x \in \mathbb{R}^n : \|x - a\|_2 \leq r\},$$

которое называется (замкнутым) *евклидовым шаром* (с центром в точке  $a$  и радиусом  $r$ ) и является выпуклым.

2. (Эллипсоид) Пусть  $P \in \mathbb{S}_{++}^n$ , и пусть  $\|x\| := \|x\|_P := \sqrt{x^T P x}$ . В этом случае рассмотренный выше абстрактный шар переходит в множество

$$\mathcal{E} := \{x \in \mathbb{R}^n : (x - a)^T P (x - a) \leq r^2\},$$

которое называется *эллипсоидом* (с центром в точке  $a$  и радиусом  $r$ ) и является выпуклым.<sup>2</sup>

3. (Гипероктаэдр) Пусть  $\|x\| := \|x\|_1 := \sum_{i=1}^n |x_i|$  —  $\ell_1$ -норма. В этом случае рассмотренный выше абстрактный шар переходит в множество

$$B_1(a, r) := \left\{ x \in \mathbb{R}^n : \sum_{i=1}^n |x_i - a_i| \leq r \right\},$$

которое называется *гипероктаэдром* (с центром в точке  $a$  и радиусом  $r$ ) и является выпуклым.

4. (Гиперкуб) Пусть  $\|x\| := \|x\|_\infty := \max_{1 \leq i \leq n} |x_i|$  —  $\ell_\infty$ -норма. В этом случае рассмотренный выше абстрактный шар переходит в множество

$$B_\infty(a, r) := \{x \in \mathbb{R}^n : a_i - r \leq x_i \leq a_i + r \text{ для всех } 1 \leq i \leq n\},$$

которое называется *гиперкубом* (с центром в точке  $a$  и радиусом  $r$ ) и является выпуклым.

**Замечание 1.** Аналогично можно показать, что и *открытый шар*, т. е. множество  $\{x \in U : \|x - a\| < r\}$ , также является выпуклым. Однако *сфера*, т. е. множество  $\{x \in U : \|x - a\| = r\}$ , уже *не* является выпуклым (почему?).

<sup>2</sup>Другое популярное представление эллипсоида — это аффинный образ единичного евклидова шара:  $\mathcal{E} = \{Qx + a : \|x\|_2 \leq 1\}$ , где  $Q \in \mathbb{R}^{n \times n}$  — невырожденная матрица (такая, что  $Q^T P Q = I_n$ ). Это представление можно получить из приведенного ранее, сделав замену переменных.

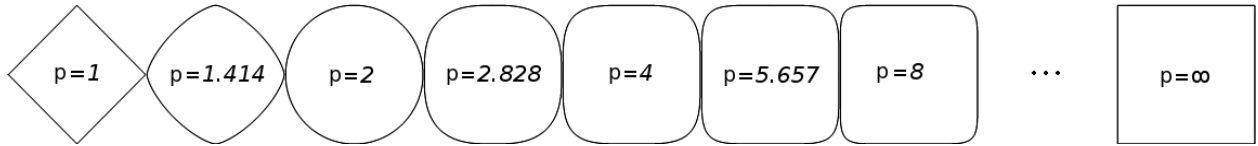


Рис. 4: Линии уровня  $f(x) = \|x\|_p$  для различных значений  $p$ .

**Пример 4** (Множество положительно полуопределенных матриц). В пространстве симметричных матриц  $\mathbb{S}^n$  множество положительно полуопределенных матриц  $\mathbb{S}_+^n$  является выпуклым.

Действительно, пусть  $X, Y \in \mathbb{S}_+^n$  — две положительно полуопределенные матрицы и  $\lambda \in [0, 1]$ . Тогда для любого  $u \in \mathbb{R}^n$  верно

$$u^T(\lambda X + (1 - \lambda)Y)u = \lambda u^T X u + (1 - \lambda)u^T Y u \geq 0.$$

Здесь последнее неравенство следует из положительной полуопределенности  $X$  и  $Y$  и условия  $\lambda \in [0, 1]$ . Таким образом, матрица  $\lambda X + (1 - \lambda)Y$  также будет положительно полуопределенной.

**Замечание 2.** Аналогично можно показать, что и внутренность множества  $\mathbb{S}_+^n$ , т. е. множество всех (строго) положительно определенных матриц  $\mathbb{S}_{++}^n$ , также является выпуклым.

## 1.2 Операции, сохраняющие выпуклость множеств

**Утверждение 1** (Операции, сохраняющие выпуклость множеств). *Следующие операции сохраняют выпуклость множеств.*

1. (Пересечение) Пусть  $U$  — вещественное векторное пространство. Пусть  $\mathcal{A}$  — произвольное (не обязательно конечное и не обязательно счетное) индексное множество, и пусть для каждого индекса  $\alpha \in \mathcal{A}$  задано множество  $Q_\alpha$  в пространстве  $U$ . Если каждое из множеств  $Q_\alpha$  является выпуклым, тогда и их пересечение

$$\bigcap_{\alpha \in \mathcal{A}} Q_\alpha := \{x : x \in Q_\alpha \text{ для всех } \alpha \in \mathcal{A}\}$$

также является выпуклым множеством в пространстве  $U$ .

2. (Прямое произведение) Пусть  $U_1, \dots, U_n$  — вещественные векторные пространства, и  $Q_1, \dots, Q_n$  — множества в пространствах  $U_1, \dots, U_n$  соответственно. Если каждое из множеств  $Q_1, \dots, Q_n$  является выпуклым, тогда и их прямое (декартово) произведение

$$Q_1 \times \dots \times Q_n := \{(x_1, \dots, x_n) : x_i \in Q_i \text{ для всех } 1 \leq i \leq n\}$$

является выпуклым множеством в пространстве  $U_1 \times \dots \times U_n$ .

3. (Проекция) Пусть  $U_1, \dots, U_n$  — вещественные векторные пространства, и  $Q_1, \dots, Q_n$  — множества в пространствах  $U_1, \dots, U_n$  соответственно. Пусть также  $1 \leq i \leq n$ . Если каждое из множеств  $Q_1, \dots, Q_n$  является выпуклым, тогда и проекция их прямого произведения  $Q_1 \times \dots \times Q_n$  на  $i$ -ую ось, т. е. множество

$$\{x_i : (x_1, \dots, x_n) \in Q_1 \times \dots \times Q_n\},$$

является выпуклым множеством в пространстве  $U_i$ .

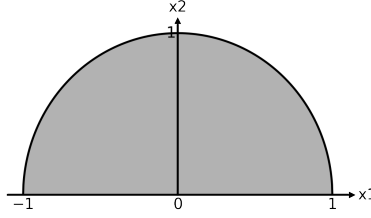


Рис. 5: Полуокруг в пространстве  $\mathbb{R}^2$ .

4. (Линейная комбинация)<sup>3</sup> Пусть  $U$  — вещественное векторное пространство, и  $Q_1, \dots, Q_k$  — множества в пространстве  $U$ . Пусть также  $\alpha_1, \dots, \alpha_k$  — скаляры (произвольного знака). Если каждое из множеств  $Q_1, \dots, Q_k$  является выпуклым, тогда и их линейная комбинация

$$\alpha_1 Q_1 + \dots + \alpha_k Q_k := \left\{ \sum_{i=1}^k \alpha_i x_i : x_i \in Q_i \text{ для всех } 1 \leq i \leq k \right\}$$

также является выпуклым множеством в пространстве  $U$ .

5. (Образ при аффинном преобразовании) Пусть  $U$  и  $V$  — вещественные векторные пространства, и  $Q$  — множество в пространстве  $U$ . Пусть  $\mathcal{A} : U \rightarrow V$  — аффинное преобразование из пространства  $U$  в пространство  $V$ , т. е. преобразование вида  $\mathcal{A}(x) = Lx + a$ , где  $L : U \rightarrow V$  — линейное преобразование и  $a \in V$ . Если множество  $Q$  является выпуклым, тогда и его образ при аффинном преобразовании  $\mathcal{A}$ , т. е. множество

$$\mathcal{A}(Q) := \{\mathcal{A}(x) : x \in Q\},$$

является выпуклым множеством в пространстве  $V$ .

6. (Обратный образ при аффинном преобразовании) Пусть  $U$  и  $V$  — вещественные векторные пространства, и  $S$  — множество в пространстве  $V$ . Пусть  $\mathcal{A} : U \rightarrow V$  — аффинное преобразование из пространства  $U$  в пространство  $V$ , т. е. преобразование вида  $\mathcal{A}(x) = Lx + a$ , где  $L : U \rightarrow V$  — линейное преобразование и  $a \in V$ . Если множество  $S$  является выпуклым, тогда и его обратный образ при аффинном преобразовании  $\mathcal{A}$ , т. е. множество

$$\mathcal{A}^{-1}(S) := \{x \in U : \mathcal{A}(x) \in S\},$$

является выпуклым множеством в пространстве  $U$ .

**Пример 5.** Рассмотрим в пространстве  $\mathbb{R}^2$  полуокруг (рис. 5)

$$\{x \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 1, x_2 \geq 0\}.$$

Согласно утверждению 1, это множество является выпуклым как пересечение единичного круга  $\{x \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 1\}$  и полупространства  $\{x \in \mathbb{R}^2 : x_2 \geq 0\}$ .

**Пример 6.** Будем работать в пространстве  $\mathbb{R}^2$ . Пусть для каждого  $\alpha \in [-\pi, \pi]$  задано полупространство

$$Q_\alpha := \{x \in \mathbb{R}^2 : (\cos \alpha)x_1 + (\sin \alpha)x_2 \leq 1\},$$

порожденное касательной к единичной окружности в точке, соответствующей углу  $\alpha$ . Рассмотрим пересечение всех полупространств  $Q_\alpha$ , т. е. множество

$$Q := \bigcap_{\alpha \in [-\pi, \pi]} Q_\alpha.$$

<sup>3</sup>Данная операция в литературе часто встречается под названием *сумма Минковского*.

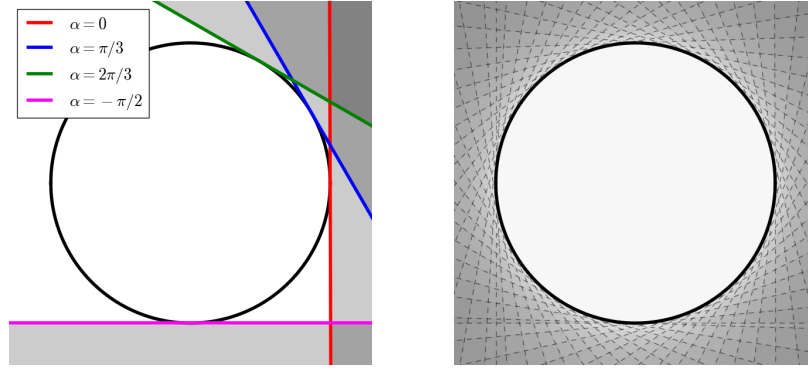


Рис. 6: Единичный круг как пересечение полупространств, образованных всевозможными касательными.

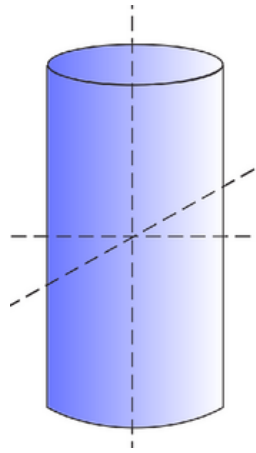


Рис. 7: Прямой круговой цилиндр.

Согласно утверждению 1, множество  $Q$  является выпуклым как пересечение (произвольного числа) выпуклых множеств (полупространств). Нетрудно понять, что множество  $Q$ , на самом деле, является единичным кругом, который, как мы уже знаем, действительно, является выпуклым (см. рис. 6).

**Пример 7.** Рассмотрим в пространстве  $\mathbb{R}^3$  *прямой круговой цилиндр* (рис. 7)

$$\{x \in \mathbb{R}^3 : x_1^2 + x_2^2 \leq 1, -1 \leq x_3 \leq 1\}.$$

Согласно утверждению 1, это множество является выпуклым как прямое произведение единичного круга  $\{x \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 1\}$  в пространстве  $\mathbb{R}^2$  и отрезка  $[-1, 1]$  в пространстве  $\mathbb{R}$ .

**Пример 8.** Рассмотрим в пространстве  $\mathbb{R}^3$  следующий эллипсоид:

$$\{x \in \mathbb{R}^3 : x_1^2 + x_2^2 + 5x_3^2 \leq 1\}.$$

Проекцией этого множества на плоскость  $(x_1, x_2)$  будет множество

$$\{x \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 1\},$$

представляющее собой единичный круг.

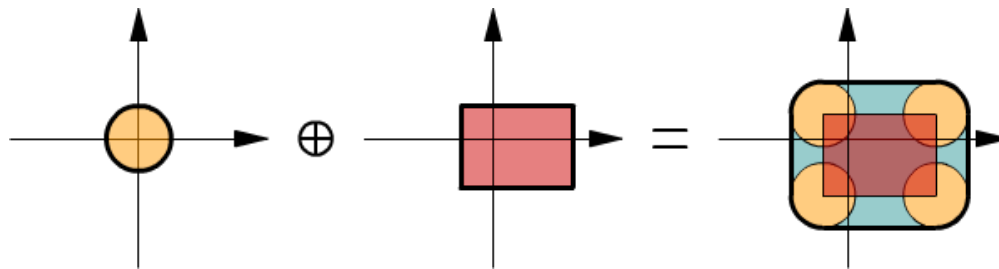


Рис. 8: Сумма круга и прямоугольника представляет собой прямоугольник большего размера с закругленными углами.

**Пример 9.** Будем работать в пространстве  $\mathbb{R}^2$ . Пусть  $Q_1 := \{x \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 1\}$  — единичный круг с центром в нуле и  $Q_2 := \{x \in \mathbb{R}^2 : -1 \leq x_1 \leq 4, -3 \leq x_2 \leq 1\}$  — прямоугольник. Сумма множеств  $Q_1$  и  $Q_2$  будет представлять собой увеличенный прямоугольник  $Q_2$  с закругленными углами (рис. 8). Согласно утверждению 1, множество  $Q_1 + Q_2$  будет выпуклым.

**Пример 10.** Рассмотрим в пространстве  $\mathbb{R}^n$  эллипсоид

$$\{Lx + a : x \in \mathbb{R}^n, \|x\|_2 \leq 1\},$$

где  $L \in \mathbb{R}^{n \times n}$  — невырожденная матрица и  $a \in \mathbb{R}^n$ . Согласно утверждению 1, это множество является выпуклым как образ единичного евклидова шара  $B_2(0, 1)$  при аффинном преобразовании  $\mathcal{A} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ , заданного формулой  $\mathcal{A}(x) = Lx + a$ .

**Пример 11** (Множество решений LMI). Пусть  $A_1, \dots, A_k$  и  $B$  — матрицы в пространстве  $\mathbb{S}^n$ . Рассмотрим линейное матричное неравенство (LMI)

$$x_1 A_1 + \dots + x_k A_k \preceq B,$$

где  $x_1, \dots, x_k \in \mathbb{R}$ . Рассмотрим множество всевозможных решений этого линейного матричного неравенства, т. е. множество

$$\{x \in \mathbb{R}^k : x_1 A_1 + \dots + x_k A_k \preceq B\}.$$

Согласно утверждению 1, это множество является выпуклым в пространстве  $\mathbb{R}^k$  как обратный образ множества положительно полуопределенных матриц  $\mathbb{S}_+^n$  при аффинном преобразовании  $\mathcal{A} : \mathbb{R}^k \rightarrow \mathbb{S}^n$ , заданного как

$$\mathcal{A}(x) := B - x_1 A_1 - \dots - x_k A_k.$$

(Почему это преобразование аффинное?)

## 2 Выпуклые функции

### 2.1 Определение и примеры

**Определение 2** (Выпуклые функции). Пусть  $U$  — вещественное векторное пространство,  $Q$  — непустое выпуклое множество в  $U$ . Функция  $f : Q \rightarrow \mathbb{R}$  называется *выпуклой*, если для любых  $x, y \in Q$  и любого  $\lambda \in [0, 1]$  выполняется

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y). \quad (2)$$

Если это неравенство является строгим для всех  $x \neq y$  и  $0 < \lambda < 1$ , то функция  $f$  называется *строго выпуклой*.

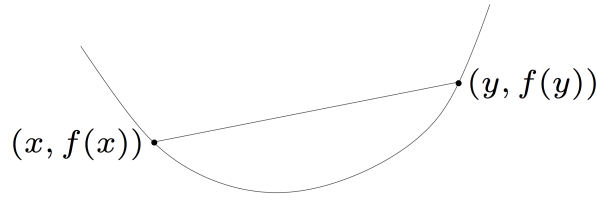


Рис. 9: Иллюстрация к определению выпуклой функции. Хорда лежит целиком над графиком функции.

**Замечание 3.** Заметим, что в этом определении подразумевается, что для любых двух допустимых точек  $x, y \in Q$  функцию  $f$  возможно «вычислить» в любой промежуточной точке отрезка  $[x, y]$ . Именно поэтому в определении требуется, чтобы область определения  $Q$  функции  $f$  являлась выпуклым множеством.

Абсолютно аналогично вводится понятие вогнутой функции. Единственное отличие по сравнению с определением выпуклой функции состоит в том, что неравенство  $\leq$  заменяется на  $\geq$ .

**Определение 3 (Вогнутые функции).** Пусть  $U$  — вещественное векторное пространство,  $Q$  — непустое выпуклое множество в  $U$ . Функция  $f : Q \rightarrow \mathbb{R}$  называется *вогнутой*, если для любых  $x, y \in Q$  и любого  $\lambda \in [0, 1]$  выполняется

$$f(\lambda x + (1 - \lambda)y) \geq \lambda f(x) + (1 - \lambda)f(y).$$

Если неравенство (2) является строгим для всех  $x \neq y$  и  $0 < \lambda < 1$ , то функция  $f$  называется *строго вогнутой*.

**Замечание 4.** Из определения легко видеть, что функция  $f$  является (строго) выпуклой тогда и только тогда, когда функция  $-f$  является (строго) вогнутой.

**Пример 12 (Аффинная функция).** Пусть в пространстве  $U$  задано (произвольное) скалярное произведение  $\langle \cdot, \cdot \rangle$ . Пусть  $f : U \rightarrow \mathbb{R}$  — *аффинная функция*

$$f(x) := \langle a, x \rangle + \beta,$$

где  $a \in U$  и  $\beta \in \mathbb{R}$ . Заметим, что для этой функции неравенство (2) переходит в равенство. Таким образом, аффинная функция является одновременно и выпуклой, и вогнутой (но не строго).

**Пример 13 (Произвольная норма).** Пусть в пространстве  $U$  задана (произвольная) норма  $\|\cdot\|$ . Тогда функция  $f : U \rightarrow \mathbb{R}$ , заданная формулой

$$f(x) := \|x\|,$$

является выпуклой.

Действительно, пусть  $x, y \in U$  и  $\lambda \in [0, 1]$ . Тогда

$$f(\lambda x + (1 - \lambda)y) = \|\lambda x + (1 - \lambda)y\| \leq \lambda \|x\| + (1 - \lambda)\|y\| = \lambda f(x) + (1 - \lambda)f(y).$$

Здесь неравенство следует из неравенства треугольника для нормы и условия  $\lambda \in [0, 1]$ .

**Замечание 5.** Рассматриваемая функция  $x \mapsto \|x\|$  не является строго выпуклой (почему?).

## 2.2 Простейшие свойства выпуклых функций

**Утверждение 2 (Неравенство Йенсена).** Пусть  $U$  — вещественное векторное пространство,  $Q$  — непустое выпуклое множество в  $U$ , и  $f : Q \rightarrow \mathbb{R}$  — выпуклая функция. Пусть также  $x_1, \dots, x_k$  —



точки во множестве  $Q$  и  $\lambda_1, \dots, \lambda_k$  — неотрицательные коэффициенты, суммирующиеся в единицу:  $\lambda_i \geq 0$  и  $\sum_{i=1}^k \lambda_i = 1$ . Тогда справедливо следующее неравенство:

$$f\left(\sum_{i=1}^k \lambda_i x_i\right) \leq \sum_{i=1}^k \lambda_i f(x_i),$$

причем равенство достигается тогда и только тогда, когда функция  $f$  является аффинной или когда все точки совпадают:  $x_1 = \dots = x_k$ .

Другими словами, неравенство Йенсена говорит о том, для выпуклой функции значение функции от выпуклой комбинации точек не превосходит соответствующей выпуклой комбинации значений функции.

**Замечание 6.** Неравенство Йенсена также обобщается и на случай выпуклой комбинации бесконечно-го (счетного или несчетного) числа точек. В случае счетного числа точек  $x_1, x_2, \dots \in Q$  соответствующие суммы переходят в бесконечные суммы  $\sum_{i=1}^{\infty} \lambda_i x_i$  и  $\sum_{i=1}^{\infty} \lambda_i f(x_i)$ , где  $\lambda_1, \lambda_2, \dots \geq 0$  и  $\sum_{i=1}^{\infty} \lambda_i = 1$ . В наиболее общей форме (для несчетного числа точек) неравенство Йенсена формулируется в терминах вероятностных интегралов или математических ожиданий. Действительно, требование о том, что веса в выпуклой комбинации должны быть неотрицательными и суммироваться в единицу, в общем случае означает, что на множестве  $Q$  должно быть задано вероятностное распределение, по которому выполняется усреднение точек множества. Сформулируем неравенство Йенсена для случайных величин. Пусть функция  $f$  выпуклая, и  $X$  — случайная величина, принимающая значения во множестве  $Q$ . Тогда справедливо неравенство

$$f(\mathbb{E}X) \leq \mathbb{E}f(x),$$

при условии, что соответствующие математические ожидания существуют.

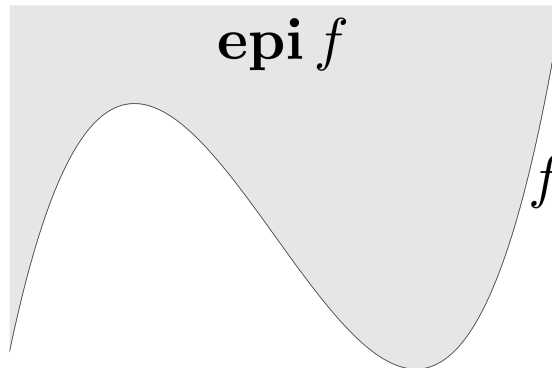


Рис. 10: Иллюстрация к определению надграфика функции.

Оказывается, что выпуклые функции и выпуклые множества тесно связаны. В частности, исследование выпуклости заданной функции всегда может быть сведено к исследованию выпуклости специального множества, ассоциированного с функцией, которое называется *надграфиком*.

**Определение 4 (Надграфик).** Пусть  $U$  — вещественное векторное пространство,  $Q$  — непустое множество в  $U$ . *Надграфиком* функции  $f : Q \rightarrow \mathbb{R}$  называется множество

$$\text{Epi}(f) := \{(x, t) \in Q \times \mathbb{R} : f(x) \leq t\}.$$

(Это множество лежит в пространстве  $U \times \mathbb{R}$ .)

Следующее утверждение можно считать альтернативным определением выпуклости функции.

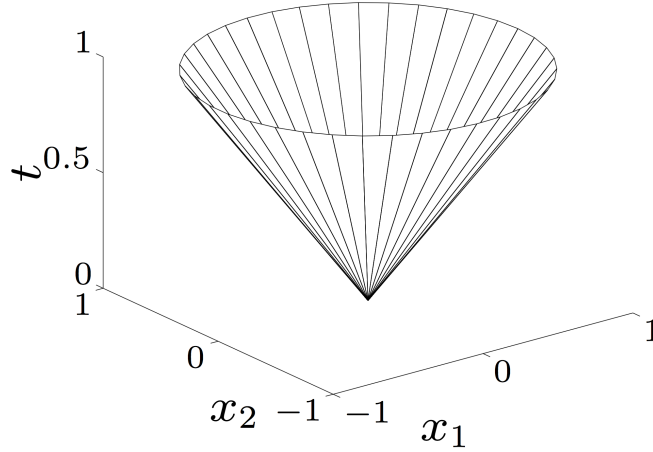


Рис. 11: Конус Лоренца для  $x \in \mathbb{R}^2$ .

**Утверждение 3** (Определение выпуклости через надграфик). Пусть  $U$  — вещественное векторное пространство,  $Q$  — непустое выпуклое множество в  $U$ . Функция  $f : Q \rightarrow \mathbb{R}$  является выпуклой тогда и только тогда, когда ее надграфик  $\text{Epi}(f)$  является выпуклым множеством в пространстве  $U \times \mathbb{R}$ .

**Пример 14** (Конус нормы). Пусть в пространстве  $U$  задана (произвольная) норма  $\|\cdot\|$ . Рассмотрим множество

$$K := \{(x, t) \in U \times \mathbb{R}_+ : \|x\| \leq t\},$$

представляющее собой надграфик функции  $x \mapsto \|x\|$ . Это множество называется *конусом нормы*. Согласно утверждению 3, множество  $K$  является выпуклым.

В случае, когда  $U = \mathbb{R}^n$  и  $\|x\| = \|x\|_2$  (евклидова норма), абстрактное множество  $K$  переходит в множество

$$\{(x, t) \in \mathbb{R}^n \times \mathbb{R}_+ : \|x\|_2 \leq t\} = \{(x, t) \in \mathbb{R}^n \times \mathbb{R}_+ : \sqrt{x^T x} \leq t\}.$$

Это множество называется *конусом Лоренца* (рис. 11). Альтернативные названия: *конус второго порядка* или *конус мороженого*.

Следующее утверждение показывает, что у выпуклой функции все линии уровня являются выпуклыми множествами.

**Утверждение 4** (Выпуклость множества линий уровня). Пусть  $U$  — вещественное векторное пространство,  $Q$  — непустое множество в  $U$ , и пусть  $f : Q \rightarrow \mathbb{R}$  — выпуклая функция. Тогда для любого  $\alpha \in \mathbb{R}$  соответствующее множество линий уровня

$$\text{Lev}_f(\alpha) := \{x \in Q : f(x) \leq \alpha\}$$

является выпуклым.

Из этого утверждения сразу же следует, что в выпуклой задаче оптимизации множество оптимальных решений является выпуклым.

**Следствие 1.** Пусть  $U$  — вещественное векторное пространство,  $Q$  — непустое множество в  $U$ , и пусть  $f : Q \rightarrow \mathbb{R}$  — выпуклая функция. Обозначим  $f^* := \inf_{x \in Q} f(x)$ . Тогда множество

$$X^* := \{x \in Q : f(x) = f^*\}$$

является выпуклым.

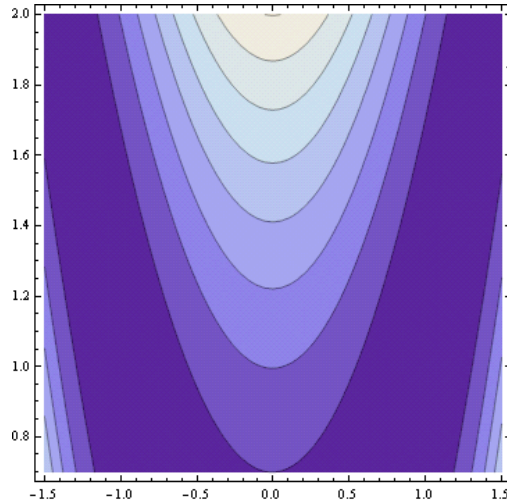


Рис. 12: Линии уровня функции Розенброка.

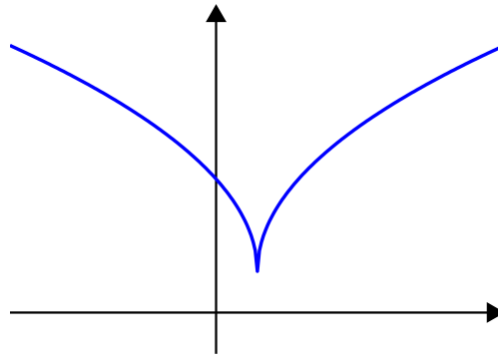


Рис. 13: Пример квазивыпуклой функции, которая не является выпуклой (из Википедии).

**Пример 15.** С помощью утверждения 4 иногда можно устанавливать *невыпуклость* функции. Например, функция Розенброка, линии уровня которой приведены на рис. 12, не может быть выпуклой, потому что имеет невыпуклые линии уровня.

**Замечание 7.** Функции, которые обладают указанным выше свойством, т. е. что множество линий уровня  $\text{Lev}_f(\alpha)$  является выпуклым для любого  $\alpha \in \mathbb{R}$ , называются *квазивыпуклыми*. Как показывает это утверждение, любая выпуклая функция является квазивыпуклой. Однако обратное утверждение не верно: существуют квазивыпуклые функции, которые не являются выпуклыми (см. рис. 13). Нам не понадобится понятие квазивыпуклой функции в этом курсе.

### 2.3 Расширение выпуклой функции на все пространство

При работе с выпуклыми функциями  $f : Q \rightarrow \mathbb{R}$  оказывается удобным считать, что функция задана не только на своей «истинной» области определения  $Q$ , но также и за ее пределами. В этом случае говорят о расширении выпуклой функции на все пространство, и считают, что за пределами своей «истинной» области определения функция принимает значение  $+\infty$ .

**Определение 5** (Эффективная область определения). Пусть  $U$  — векторное пространство, и  $f : U \rightarrow \mathbb{R}^*$  — функция, принимающая значения во множестве расширенных вещественных чисел  $\mathbb{R}^* :=$

$\mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$ . Будем называть *эффективной областью определения* функции  $f$  множество всех точек, в которых функция принимает конечные значения:

$$\text{Dom } f := \{x \in U : |f(x)| < +\infty\}.$$

**Определение 6** (Выпуклые и вогнутые расширеннозначные функции). Пусть  $U$  — векторное пространство. Пусть функция  $f : U \rightarrow \mathbb{R} \cup \{+\infty\}$  задана на всем пространстве и принимает расширенные вещественные значения. Функция  $f$  называется *выпуклой*, если для любых  $x, y \in U$  и любых  $\lambda \in (0, 1)$  выполнено

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y). \quad (3)$$

Аналогично, если  $f : U \rightarrow \mathbb{R} \cup \{-\infty\}$ , и для всех  $x, y \in U$  и  $\lambda \in (0, 1)$  указанное выше неравенство выполнено с противоположным знаком, то функция  $f$  называется *вогнутой*.

**Замечание 8.** Согласно определению, выпуклая расширеннозначная функция может принимать только одно расширенное значение — значение  $+\infty$ . Аналогично, вогнутая вещественнозначная функция может принимать только  $-\infty$ . Функции, которые в некоторых точках принимают  $+\infty$ , а в некоторых  $-\infty$ , не рассматриваются.

**Замечание 9** (Операции в  $\mathbb{R}^*$ ). Операции во множестве расширенных вещественных чисел подчиняются следующим естественным правилам.

1. Операции с вещественными числами понимаются в обычном смысле.
2. (Порядок) Любое вещественное число строго меньше  $+\infty$ , а также  $-\infty < +\infty$ . Аналогично для  $-\infty$ .
3. (Сумма) Сумма  $+\infty$  и любого вещественного числа, а также двух  $+\infty$  равна  $+\infty$ . Аналогично для  $-\infty$ . Сумма  $+\infty$  и  $-\infty$  не определена.
4. (Произведение) Произведение  $+\infty$  и положительного вещественного числа, а также произведение  $+\infty$  и  $+\infty$  равно  $+\infty$ . Произведение  $+\infty$  и отрицательного вещественного числа, а также произведение  $+\infty$  и  $-\infty$  равно  $-\infty$ . Аналогично для  $-\infty$ . Произведение «бесконечности» и нуля не определено.

**Замечание 10.** Определение 6 автоматически накладывает условие на выпуклость (эффективной) области определения функции  $f$ . Действительно, пусть  $x, y \in \text{Dom } f$  и  $\lambda \in [0, 1]$ . Поскольку  $x, y \in \text{Dom } f$ , то  $f(x)$  и  $f(y)$  являются конечными. Отсюда следует, что правая часть в неравенстве (3) также конечна. Но это возможно лишь в том случае, когда и левая часть в неравенстве (3) конечна. Значит,  $f(\lambda x + (1 - \lambda)y) < +\infty$  и  $\lambda x + (1 - \lambda)y \in \text{Dom } f$ . Таким образом, расширенное определение выпуклости 6 эквивалентно введенному до этого обычному определению выпуклости 2 (в котором функция задана на множестве  $Q := \text{Dom } f$ ).

**Пример 16** (Индикатор выпуклого множества). Пусть  $U$  — векторное пространство, и пусть  $Q$  — выпуклое множество в  $U$ . Рассмотрим функцию  $\delta_Q : U \rightarrow \mathbb{R} \cup \{+\infty\}$ , заданную формулой

$$\delta_Q(x) := \begin{cases} 0, & \text{если } x \in Q, \\ +\infty, & \text{если } x \notin Q. \end{cases}$$

Эта функция называется *индикатором* множества  $Q$  и является выпуклой (почему?).

## 2.4 Дифференциальные критерии выпуклости

**Утверждение 5** (Условие выпуклости первого порядка). Пусть  $\text{Dom } f$  является открытым множеством, и функция  $f$  дифференцируема всюду на  $\text{Dom } f$ . Функция  $f$  является выпуклой тогда и только тогда, когда  $\text{Dom } f$  является выпуклым множеством и

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle$$

для всех  $x, y \in \text{Dom } f$ .

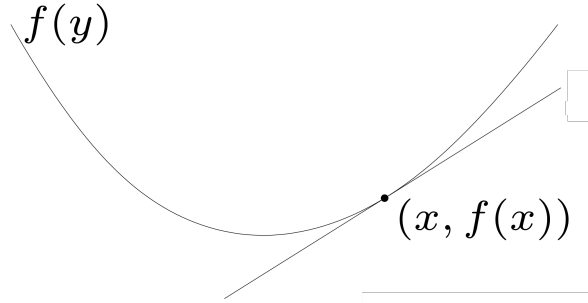


Рис. 14: Иллюстрация к условию выпуклости первого порядка. График функции лежит всюду выше касательной, проведенной к графику в любой точке  $x$ .

**Утверждение 6** (Дифференциальное условие оптимальности для выпуклой функции). Пусть  $f : U \rightarrow \mathbb{R} \cup \{+\infty\}$  — выпуклая функция, и пусть  $x^*$  — некоторая внутренняя точка множества  $\text{Dom } f$ . Точка  $x^*$  является глобальным минимумом функции  $f$  тогда и только тогда, когда  $\nabla f(x^*) = 0$ . Другими словами, любая стационарная точка автоматически является глобальным минимумом функции  $f$ .

*Доказательство.* Согласно условию оптимальности первого порядка, для всех  $x \in \text{Dom } f$  справедлива оценка

$$f(x) \geq f(x^*) + \langle \nabla f(x^*), x - x^* \rangle = f(x^*).$$

□

**Пример 17.** Это утверждение позволяет для выпуклых дифференцируемых функций не задумываться о том, достигается ли глобальный минимум или нет. Например, вспомним задачу регрессии наименьших квадратов  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ :

$$\min_{x \in \mathbb{R}^n} \left\{ f(x) := \frac{1}{2} \|Ax - b\|_2^2 \right\}.$$

Эта функция является всюду дифференцируемой ( $\text{Dom } f = \mathbb{R}^n$ ), и, значит, поиск ее минимума эквивалентен решению системы уравнений

$$\nabla f(x) = A^T(Ax - b) = 0.$$

Согласно установленному утверждению, решения задачи — это в точности все решения этой системы линейных уравнений, и только они. Таким образом, можно просто решать систему линейных уравнений и не переживать о том, что таким образом могут быть найдены какие-то стационарные точки, которые не являются глобальными решениями задачи. (Для невыпуклых функций так делать нельзя!)

**Утверждение 7** (Условие выпуклости второго порядка). Пусть  $\text{Dom } f$  является открытым множеством, и функция  $f$  дважды дифференцируема на  $\text{Dom } f$ . Функция  $f$  является выпуклой тогда и только тогда, когда  $\text{Dom } f$  является выпуклым множеством и

$$D^2 f(x)[h, h] =: \langle \nabla^2 f(x)h, h \rangle \geq 0$$

для всех  $x \in \text{Dom } f$  и всех  $h \in U$ . Если  $U = \mathbb{R}^n$ , то это эквивалентно положительной полуопределенности гессиана:

$$\nabla^2 f(x) \succeq 0.$$

для всех  $x \in \text{Dom } f$ .

**Пример 18** (Одномерные выпуклые функции). Следующие функции являются выпуклыми:

- (Экспонента)  $\exp(x)$  выпукла на  $\mathbb{R}$
- (Минус логарифм)  $-\ln x$  выпукла на  $\mathbb{R}_{++}$
- (Степенная функция)
  - $x^{2p}$  для  $p \in \{1, 2, \dots\}$  на  $\mathbb{R}$
  - $x^p$  для  $p \geq 1$  на  $\mathbb{R}_+$
  - $-x^p$  для  $0 \leq p \leq 1$  на  $\mathbb{R}_+$
  - $1/x^p$  для  $p > 0$  на  $\mathbb{R}_{++}$
- $x \ln x$  выпукла на  $\mathbb{R}_+$

Доказывается через условие второго порядка.

**Пример 19** (Квадратичная функция). Пусть  $A \in \mathbb{S}^n$ ,  $b \in \mathbb{R}^n$ . Квадратичная функция

$$f(x) := \frac{1}{2}x^T A x - b^T x$$

является выпуклой тогда и только тогда, когда  $A \succeq 0$  и вогнутой тогда и только тогда, когда  $A \preceq 0$ . Это следует из условия второго порядка:  $\nabla^2 f(x) = A$ .

**Пример 20** (Log-sum-exp). Функция  $f : \mathbb{R}^n \rightarrow \mathbb{R}$

$$f(x) := \ln \left( \sum_{i=1}^n e^{x_i} \right)$$

является выпуклой. Ее гессиан равен

$$\nabla^2 f(x) = \text{Diag}\{\pi(x)\} - \pi(x)\pi(x)^T,$$

где  $\pi(x) := \exp(x)/1_n^T \exp(x)$  (поэлементно). Гессиан оказывается положительно полуопределенным:

$$u^T \nabla^2 f(x) u = u^T \text{Diag}\{\pi\} u - (\pi^T u)^2 = \sum_{i=1}^n \pi_i u_i^2 - \left( \sum_{i=1}^n \pi_i u_i \right)^2 \geq 0.$$

Последнее неравенство следует из неравенства Коши-Буняковского и того факта, что  $\sum_{i=1}^n \pi_i = 1$ .

**Пример 21** (Минус логарифм определителя). Функция

$$f(X) := -\ln \text{Det}(X)$$

является выпуклой на  $S_{++}^n$ . Действительно, рассмотрим

$$D^2 f(X)[H, H] = \text{Tr}(X^{-1} H X^{-1} H).$$

Покажем, что  $D^2 f(X)[H, H] \geq 0$  для всех  $X \in S_{++}^n$  и всех  $H \in \mathbb{S}^n$ . Поскольку  $X$  является симметричной положительно определенной матрицей, можно рассмотреть ее корень  $X^{1/2}$ . Тогда

$$D^2 f(X)[H, H] = \text{Tr}(X^{-1/2} H X^{-1/2} X^{-1/2} H X^{-1/2}) = \text{Tr}([X^{-1/2} H X^{-1/2}]^2)$$

Матрица  $X^{-1/2} H X^{-1/2}$  является симметричной. Значит, ее квадрат гарантированно будет симметричной неотрицательно определенной матрицей. Поскольку след равен сумме собственных значений, и в данном случае все они неотрицательные, то  $D^2 f(X)[H, H] \geq 0$ .

Следующее свойство позволяет с помощью производных доказать выпуклость на внутренности множества, а затем расширить это понятие на все множество — если функция непрерывна на множестве.

**Утверждение 8** (Полезное свойство расширения на замыкание). Пусть функция  $f$  является выпуклой всюду на внутренности  $\text{Dom } f$ , и непрерывной всюду на  $\text{Dom } f$ . Тогда  $f$  является выпуклой на всем  $\text{Dom } f$ .

## 2.5 Операции, сохраняющие выпуклость функций

**Утверждение 9** (Операции, сохраняющие выпуклость функций). *Следующие операции сохраняют выпуклость функций.*

1. (Положительная взвешенная сумма) Пусть  $U$  — вещественное векторное пространство и  $f_1, \dots, f_k : U \rightarrow \mathbb{R} \cup \{+\infty\}$  — функции. Пусть также  $w_1, \dots, w_k$  — положительные коэффициенты. Рассмотрим взвешенную сумму функций  $f_1, \dots, f_k$  с коэффициентами  $w_1, \dots, w_k$ , т. е. функцию  $\phi : U \rightarrow \mathbb{R} \cup \{+\infty\}$ , определенную по формуле

$$\phi(x) := \sum_{i=1}^k w_i f_i(x).$$

Если каждая из функций  $f_1, \dots, f_k$  является выпуклой, тогда и  $\phi$  будет выпуклой функцией.

2. (Аффинная подстановка аргумента) Пусть  $U$  и  $V$  — вещественные векторные пространства, и  $f : V \rightarrow \mathbb{R} \cup \{+\infty\}$  — функция. Пусть  $A : U \rightarrow V$  — аффинное преобразование из пространства  $U$  в пространство  $V$ , т. е. преобразование вида  $A(x) = Lx + a$ , где  $L : U \rightarrow V$  — линейное преобразование и  $a \in V$ . Рассмотрим функцию, получающуюся из функции  $f$  с помощью аффинной подстановки аргумента, т. е. функцию  $\phi : U \rightarrow \mathbb{R} \cup \{+\infty\}$ , определенную по формуле

$$\phi(x) := f(A(x)).$$

Если функция  $f$  выпуклая, тогда и функция  $\phi$  также будет выпуклой.

3. (Поточечный супремум) Пусть  $U$  — вещественное векторное пространство. Пусть  $A$  — произвольное (не обязательно конечное и не обязательно счетное) индексное множество, и пусть для каждого индекса  $\alpha \in A$  задана функция  $f_\alpha : U \rightarrow \mathbb{R} \cup \{+\infty\}$ . Рассмотрим поточечный супремум функций  $f_\alpha$ , т. е. функцию  $\phi : U \rightarrow \mathbb{R} \cup \{+\infty\}$ , определенную по формуле

$$\phi(x) := \sup_{\alpha \in A} f_\alpha(x).$$

Если каждая из функций  $f_\alpha$  является выпуклой, тогда и функция  $\phi$  также будет выпуклой.

4. (Монотонная суперпозиция) Пусть  $U$  — вещественное векторное пространство, и  $f_1, \dots, f_n : U \rightarrow \mathbb{R}$  — функции. Пусть также  $g : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  — функция. Рассмотрим функцию, являющуюся суперпозицией функции  $g$  и  $f_1, \dots, f_n$ , т. е. функцию  $\phi : U \rightarrow \mathbb{R} \cup \{+\infty\}$ , определенную по формуле

$$\phi(x) := g(f_1(x), \dots, f_n(x)).$$

Если каждая из функций  $f_1, \dots, f_n$  является выпуклой, а функция  $g$  является выпуклой и монотонно неубывающей, т. е.  $g(y) \leq g(y')$  для всех  $y \preceq y'$ , тогда функция  $\phi$  также будет выпуклой.

5. (Частичная минимизация) Пусть  $U$  и  $V$  — вещественные векторные пространства, и  $f : U \times V \rightarrow \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$  — функция. Рассмотрим функцию  $\phi : U \rightarrow \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$ , определенную по формуле

$$\phi(x) := \inf_{y \in V} f(x, y).$$

Если функция  $f$  является выпуклой (как функция одновременно двух переменных  $x$  и  $y$ ), тогда  $\phi$  является выпуклой функцией (при условии, что  $\phi$  ни в одной точке не принимает значение  $-\infty$ ).

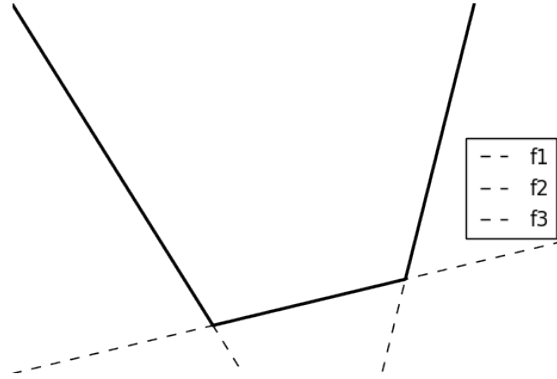


Рис. 15: Иллюстрация к примеру поточечный максимум. Так как линейная функция — выпуклая, максимум из линейных функций — выпуклая функция.

**Пример 22** (Взвешенная сумма и аффинная подстановка аргумента). Пусть  $a, b \in \mathbb{R}^k$  и  $c \in \mathbb{R}_+^k$ .  
Функция

$$f(x) := \sum_{i=1}^k c_i \exp(a_i^T x + b_i)$$

является выпуклой как взвешенная сумма экспонент с аффинной подстановкой аргумента.

**Пример 23** (Поточечный максимум). Пусть  $f_1, \dots, f_k$  — выпуклые функции. Тогда

$$\phi(x) := \max\{f_1(x), \dots, f_k(x)\}$$

будет выпуклой функцией, потому что максимум — это частный случай супремума.

**Пример 24** (Минимальное и максимальное собственные значения). Работаем в пространстве симметричных матриц  $\mathbb{S}^n$ . Функция  $\lambda_{\max}(X)$  является выпуклой, а функция  $\lambda_{\min}(X)$  является вогнутой. Это следует из представления

$$\lambda_{\max}(X) := \max\{u^T X u : u \in S_2^{n-1}\}, \quad \lambda_{\min}(X) := \min\{u^T X u : u \in S_2^{n-1}\},$$

где  $S_2^{n-1} := \{u \in \mathbb{R}^n : \|u\|_2 = 1\}$  — евклидова сфера в пространстве  $\mathbb{R}^n$ .

**Замечание 11.** Аналогично с сингулярными значениями:

$$\sigma_{\max}(X) := \max\{u^T X v : u \in S^{m-1}, v \in S_2^{n-1}\}, \quad \sigma_{\min}(X) := \min\{u^T X v : u \in S^{m-1}, v \in S_2^{n-1}\}.$$

**Замечание 12** (Опорная функция множества). Пусть  $M$  — произвольное (не обязательно выпуклое) непустое множество. Тогда *опорная функция* этого множества

$$S_M(y) := \sup_{x \in M} \langle y, x \rangle$$

является выпуклой как поточечный супремум от аффинных функций.

**Пример 25** (Норма в степени). Пусть  $\|\cdot\|$  — произвольная норма. Тогда функция

$$f(x) := \|x\|^p$$

является выпуклой при  $p \geq 1$  на всем пространстве. Здесь используется композиция выпуклой монотонно возрастающей одномерной функции  $x^p$  и выпуклой функции  $\|x\|$ .



**Пример 26** (Расстояние до выпуклого множества). Пусть  $Q$  — выпуклое множество,  $\|\cdot\|$  — произвольная норма, и пусть  $x \in U$ , Тогда функция

$$f(x) := \rho(x, Q) := \inf_{y \in Q} \|x - y\|$$

является выпуклой как частичная минимизация  $\|x - y\| + \delta_Q(y)$  по  $y$  (почему  $\|x - y\|$  выпукла совместно по  $(x, y)$ ?).

**Пример 27.** Пусть  $f$  — выпуклая функция. Тогда функция

$$\phi(x) := \inf_y \{f(y) : Ay = x\}$$

также выпуклая как частичная минимизация по  $y$  функции

$$g(x, y) := \begin{cases} f(y), & \text{если } Ay = x, \\ +\infty, & \text{иначе,} \end{cases}$$

которая является совместно выпуклой по  $(x, y)$ .