§4. Выпуклые анализ
4.1 Элементы выпуклого анализа.
4.1.1 Выпуклые множества

Пусть X — линейное нормированное пространство (для простоты понимания можно считать, что $X = \mathbb{R}^n$).

- отрезок $[a, b]$ $= \{x \in X \mid x = a + t(b - a) = (1 - t)a + tb,\ 0 \leq t \leq 1\}$;
- интервал $(a, b) = \{x \in X \mid x = (1 - t)a + tb,\ 0 < t < 1\}$;
- выпуклое множество A, если $\forall\, a, b \in A$ отрезок $[a, b] \subset A$, т. е. $\forall\ a, b \in A$ точка $(1 - t)a + tb \in A\ \forall\ 0 \leq t \leq 1$;
- конус K $(K \neq \emptyset)$, если $\forall\ x \in K$ точка $tx \in K\ \forall\ t \geq 0$;
- аффинное множество A, если $\forall\ a, b \in A$ точка $(1 - t)a + tb \in A\ \forall\ t \in \mathbb{R}$.

Очевидно: аффинное множество выпукло.

Пусть точки $a_1, \dots, a_m \in X$, $\sum\limits_{i=1}^{m} t_i a_i$ — комбинация, $t_i \in \mathbb{R}$.

- выпуклая комбинация, если $t_i \geq 0$, $\sum\limits_{i=1}^{m} t_i = 1$;
- коническая комбинация, если $t_i \geq 0$;
- аффинная комбинация, если $\sum\limits_{i=1}^{m} t_i = 1$;
- выпуклая оболочка
  $\operatorname{conv}\{a_1, \dots, a_m\} := \left\{ a = \sum\limits_{i=1}^{m} t_i a_i \;\middle|\; t_i \geq 0,\ \sum\limits_{i=1}^{m} t_i = 1 \right\}$ —
  совокупность всех выпуклых комбинаций;
- коническая выпуклая оболочка
  $\operatorname{conev}\{a_1, \dots, a_m\} := \left\{ a = \sum\limits_{i=1}^{m} t_i a_i \;\middle|\; t_i \geq 0 \right\}$;
- аффинная оболочка
  $\operatorname{aff}\{a_1, \dots, a_m\} := \left\{ a = \sum\limits_{i=1}^{m} t_i a_i \;\middle|\; \sum\limits_{i=1}^{m} t_i = 1 \right\}$.

$f\colon X \to \overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$.

- эффективное множество: $\operatorname{dom} f := \{x \in X \mid f(x) < +\infty\}$;
- надграфик функции: $\operatorname{epi} f := \{(a, x) \in \mathbb{R} \times X \mid a \geq f(x)\}$;
- выпуклая функция, если $\operatorname{epi} f$ — выпуклое множество;
- замкнутая функция, если $\operatorname{epi} f$ — замкнутое множество;
- собственная функция, если $f(x) > -\infty \ \forall\, x$ и $f \not\equiv +\infty$.

Мы будем изучать выпуклые собственные функции. Для краткости будем называть их просто “выпуклые” функции.

Из определения выпуклого множества сразу следует, что f выпукла $\Leftrightarrow$ выполнено неравенство Йенсена:

$$f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2) \quad \forall\, x_1, x_2 \in X,\ \forall\, t \in (0,1).$$

Сумма выпуклых функций является функцией выпуклой.

Суперпозиция не всегда является выпуклой функцией.

Выпуклость функций одной переменной.

Теорема (1)

$f\colon \mathbb{R} \to \mathbb{R}$, $f \in D^2(\mathbb{R}) \Rightarrow f$ выпукла $\Leftrightarrow f''(x) \geq 0 \ \forall\ x \in \mathbb{R}$.

Примеры выпуклых функций одной переменной:

- $f(x) = e^{ax}$, $a \in \mathbb{R}$;
- $f(x) = |x|^p$, $p \geq 1$;
- аффинная функция (в многомерном случае $f\colon \mathbb{R}^n \to \mathbb{R}$, $f(x) = \langle a, x\rangle + b$, $a \in \mathbb{R}^n$, $b \in \mathbb{R}$) является выпуклой по неравенству Йенсена.
- Модуль аффинной функции $|\langle a, x\rangle + b|$ также является выпуклой функцией, поскольку ее надграфик — пересечение полупространств — выпуклое множество.

Выпуклость функций нескольких переменных.

Теорема (2)

$f\colon \mathbb{R}^n \to \mathbb{R}$, $f \in D^2(\mathbb{R}^n) \Rightarrow f$ выпукла $\Leftrightarrow$
гессиан $f''(x) = \left(\dfrac{\partial^2 f(x)}{\partial x_i \partial x_j}\right)_{i,j=1}^{n} \geq 0 \ \forall\, x \in \mathbb{R}^n$.

Примеры выпуклых функций нескольких переменных:

- квадратичная функция $Q(x) = \langle Ax, x\rangle$ (A — симметричная матрица) выпукла $\Leftrightarrow A \geq 0$;
- функция нормы

$$f(x) = \|x\|_p = \|x\|_{l_p} := \begin{cases} \left(\sum\limits_{j=1}^{n} |x_j|^p\right)^{\frac{1}{p}}, & 1 \leq p < +\infty, \\ \max\{|x_1|, \ldots, |x_n|\}, & p = +\infty; \end{cases}$$

- индикаторная функция выпуклого множества $A \subset X$
$$\delta A(x) = \begin{cases} 0, & x \in A, \\ +\infty, & x \notin A; \end{cases}$$
- функция Минковского выпуклого множества $A \subset X$
$$\mu A(x) = \begin{cases} 0, & \alpha^{-1}x \in A \ \forall\ \alpha > 0, \\ +\infty, & \alpha^{-1}x \notin A \ \forall\ \alpha > 0, \\ \inf\{\alpha \mid \alpha > 0,\ \alpha^{-1}x \in A\}, & \text{иначе} \end{cases}$$
(функция Минковского означает наименьшее число, в которое надо уменьшить вектор x, чтобы он попал в множество A);
- опорная функция непустого множества $A \subset X$
$$sA(x^*) = \max_{x \in A} \langle x^*, x \rangle \qquad (sA \colon X^* \to \mathbb{R}).$$

Теорема

$f\colon \mathbb{R} \to \mathbb{R}$ — выпуклая ограниченная функция $\Rightarrow f(x) = \mathrm{const}$.

$\triangleleft$ Предположим противное, что $f(x) \neq \mathrm{const} \Rightarrow \exists\, x_1 < x_2 :$ $f(x_1) \neq f(x_2)$. f — выпукла $\Rightarrow \operatorname{epi} f$ — выпуклое множество $\Rightarrow$ отрезок $\big[(x_1, f(x_1)), (x_2, f(x_2))\big] \subset \operatorname{epi} f \Rightarrow$ на $[x_1, x_2]$ график f лежит ниже этого отрезка (или совпадает с ним). НеОО $f(x_1) < f(x_2)$. Пусть $l(x)$ — прямая, проходящая через точки $(x_1, f(x_1))$ и $(x_2, f(x_2))$. Покажем, что $f(x) \geq l(x)\ \forall\, x > x_2$. $(*)$ Если это не так, то $\exists\, x > x_2 : f(x) < l(x)$, т. е. точка $(x, f(x))$ лежит ниже прямой l. Так как f выпукла, то отрезок $\big[(x_1, f(x_1)), (x, f(x))\big] \subset \operatorname{epi} f$ и лежит ниже прямой l, но это не так в точке x_2 (ниже точки $(x_2, f(x_2))$ не может быть точек из $\operatorname{epi} f$). Значит, неравенство $(*)$ выполняется. Поскольку $l(x) \to +\infty$ при $x \to +\infty$, то $f(x) \to +\infty$ при $x \to +\infty$ — противоречие с ограниченностью функции f.

Значит, предположение противного неверно. $\triangleright$

Определение

Субдифференциал выпуклой функции f в точке $\hat{x}$:

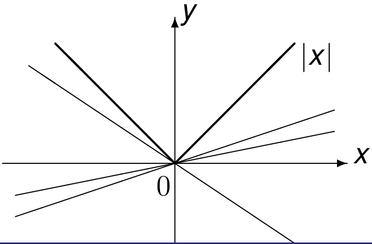
$$\partial f(\hat{x}) := \{x^* \in X^* \mid \langle x^*, x - \hat{x}\rangle \le f(x) - f(\hat{x}) \quad \forall\, x \in X\}.$$

Субдифференциал — выпуклое замкнутое множество.

$f \in D(\hat{x}) \Rightarrow \partial f(\hat{x}) = f'(\hat{x})$.

$f\colon \mathbb{R} \to \mathbb{R} \Rightarrow \partial f(\hat{x}) = \{k\colon$ прямые $y = kx + b$, проходящие через точку $(\hat{x}, f(\hat{x}))$, лежат под графиком $y = f(x)\}$.

Пример. $f(x) = |x|$.



$$\partial|x| = \begin{cases} \operatorname{sign} x, & x \neq 0, \\ [-1, 1], & x = 0. \end{cases}$$

4.1.3 Субдифференциал выпуклой функции

Теорема (Моро–Рокафеллар)

f_1 и f_2 — выпуклые функции, $\exists\ x_0 : |f_1(x_0)| < \infty$, $f_2 \in C(x_0) \Rightarrow$
$$\partial(f_1 + f_2)(x) = \partial f_1(x) + \partial f_2(x)\ \forall\ x.$$

Теорема (Дубовицкий–Милютин)

f_1 и f_2 — выпуклые функции, $f_1, f_2 \in C(\hat{x})$, $f_1(\hat{x}) = f_2(\hat{x}) \Rightarrow$
$$\partial \max\{f_1, f_2\}(\hat{x}) = \mathrm{conv}\left(\partial f_1(\hat{x}) \cup \partial f_2(\hat{x})\right).$$

Сублинейные функции

Определение

p — сублинейная функция, если $\operatorname{epi} p$ — выпуклый конус с вершиной в нуле.

Из неравенства Йенсена следует, что собственная функция p является сублинейной $\Leftrightarrow$
a) $p(\lambda x) = \lambda p(x)\ \forall\ \lambda > 0,\ \forall\ x \in X$;
b) $p(x_1 + x_2) \le p(x_1) + p(x_2)\ \forall\ x_1, x_2 \in X$.
X — нормированное пр-во $\Rightarrow f(x) = \|x\|$ — сублинейная.

Теорема (Моро–Рокафеллар)

p_1 и p_2 — сублинейные функции, p_1 непрерывна, p_2 замкнута $\Rightarrow \partial(p_1 + p_2)(x) = \partial p_1(x) + \partial p_2(x)$.

Теорема (Дубовицкий–Милютин)

p_1 и p_2 — непрерывные сублинейные функции $\Rightarrow$
$$\partial \max\{p_1, p_2\}(0) = \operatorname{conv}\left(\partial p_1(0) \cup \partial p_2(0)\right).$$

Субдифференциал модуля функции

Пусть $f(x)$ — дифференцируемая функция n переменных. Рассмотрим функцию $|f(x)|$. Найдем ее субдифференциал.
Если $f(x) < 0$, то $|f(x)| = -f(x)$ и $\partial|f(x)| = -f'(x)$.
Если $f(x) > 0$, то $|f(x)| = f(x)$ и $\partial|f(x)| = f'(x)$.
Если $f(x) = 0$, то запишем функцию $|f(x)|$ в виде
$|f(x)| = \max\{-f(x), f(x)\} \Rightarrow$ по т. Дубовицкого–Милютина

$$\partial|f(x)| = \partial \max\{-f(x), f(x)\} = \operatorname{conv}\left\{\partial(-f)(x) \cup \partial f(x)\right\} =$$

$$= \operatorname{conv}\{-f'(x), f'(x)\} = [-f'(x);\ f'(x)].$$

Субдифференциал является отрезком в пространстве $\mathbb{R}^n$.
Таким образом,

$$\partial|f| = \begin{cases} -f'(x), & f(x) < 0, \\ \{\alpha f'(x),\ \alpha \in [-1;\ 1]\}, & f(x) = 0, \\ f'(x), & f(x) > 0. \end{cases}$$

Субдифференциал нормы. Пусть X — линейное нормированное пространство, $f(x) = \|x\|$. Найдем $\partial f(0)$. По определению субдифференциала

$$x^* \in \partial f(\hat{x}) \iff \langle x^*, x - \hat{x}\rangle \leq f(x) - f(\hat{x}) \ \forall\, x \in X.$$

У нас $\hat{x} = 0$, $f(\hat{x}) = \|0\| = 0$. Следовательно,

$$x^* \in \partial f(0) \iff \langle x^*, x\rangle \leq \|x\| \quad \forall\, x \in X. \qquad (*)$$

Пусть $x^* \in \partial f(0) \Rightarrow$ по лемме Банаха $\exists\, x \in X : \|x\| = 1$, $\langle x^*, x\rangle = \|x^*\|$. Подставляя x в неравенство $(*)$, получим

$$\|x^*\| = \langle x^*, x\rangle \overset{(*)}{\leq} \|x\| = 1 \Longrightarrow \|x^*\| \leq 1.$$

Значит, $\partial f(0) \subset B^* := \left\{x^* \in X^* \mid \|x^*\|_{X^*} \leq 1\right\}$ — единичный шар сопряженного пространства. Докажем, что $\partial f(0) \supset B^*$. Возьмем $x^* \in B^* \Rightarrow$ по неравенству Коши–Буняковского

$$\langle x^*, x\rangle \leq \|x\| \cdot \|x^*\| \overset{\|x^*\|\leq 1}{\leq} \|x\| \ \forall\, x \in X.$$

Неравенство $(*)$ выполняется, значит, $x^* \in \partial f(0)$.
Таким образом, $\partial f(0) = B^* = \{x^* \in X^* \mid \|x^*\|_{X^*} \leq 1\}$.