4 Выпуклый анализ

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

4.1 Элементы выпуклого анализа

4.1.1 Выпуклые множества

Введем некоторые понятия, которые используются в выпуклом анализе:

- *отрезок* $[a, b] =$
  $= \{x \in X \mid x = a + t(b - a) = (1 - t)a + tb,\ 0 \le t \le 1\}$;
- *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 $(1 - t)a + tb \in A\ \forall\ 0 \le t \le 1$;
- *конус* K $(K \ne \emptyset)$, если $\forall\ x \in K$ точка $tx \in K\ \forall\ t \ge 0$;
- *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.

Возьмем фиксированные точки $a_1, \ldots, a_m \in X$. Пусть $\sum\limits_{i=1}^{m} t_i a_i$ — комбинация этих точек, $t_i \in \mathbb{R}$. Дадим различные определения таких комбинаций и оболочек этих комбинаций:

- *выпуклая комбинация*, если $t_i \ge 0$, $\sum\limits_{i=1}^{m} t_i = 1$;
- *коническая комбинация*, если $t_i \ge 0$;
- *аффинная комбинация*, если $\sum\limits_{i=1}^{m} t_i = 1$;

- *выпуклая оболочка* $\operatorname{conv}\{a_1, \ldots, a_m\} := \Bigl\{ a = \sum_{i=1}^{m} t_i a_i \Bigm| t_i \geq 0,$ $\sum_{i=1}^{m} t_i = 1 \Bigr\}$ — совокупность всех выпуклых комбинаций;
- *коническая выпуклая оболочка* $\operatorname{conev}\{a_1, \ldots, a_m\} := \Bigl\{ a = \sum_{i=1}^{m} t_i a_i \Bigm| t_i \geq 0 \Bigr\};$
- *аффинная оболочка* $\operatorname{aff}\{a_1, \ldots, a_m\} := \Bigl\{ a = \sum_{i=1}^{m} t_i a_i \Bigm| \sum_{i=1}^{m} t_i = 1 \Bigr\}.$

4.1.2 Выпуклые функции

Пусть задана функция (функционал), отображающая линейное нормированное пространство в “расширенную” прямую:

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

С каждой такой функцией f связываются два множества:
$\operatorname{dom} f = \{x \in X \mid f(x) < +\infty\}$ — *эффективное множество*
и $\operatorname{epi} f = \{(a, x) \in \mathbb{R} \times X \mid a \geq f(x)\}$ — *надграфик* функции f.

- Функция f называется *выпуклой*, если надграфик f — выпуклое множество.
- Функция f называется *замкнутой*, если надграфик f — замкнутое множество.
- Функция f называется *собственной*, 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 функция выпукла тогда и только тогда, когда выполнено *неравенство Йенсена*:

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

Ясно, что сумма двух выпуклых функций является функцией выпуклой. Но суперпозиция двух выпуклых функций не всегда является выпуклой функцией. Приведите пример.

Выпуклость многих классических функций одной переменной сразу вытекает из следующей теоремы.

Теорема 1 (Р, с. 44)**.** *Пусть функция* $f\colon \mathbb{R} \to \mathbb{R}$ *дважды непрерывно дифференцируема* ($f \in D^2(\mathbb{R})$)*. Тогда она выпукла тогда и только тогда, когда ее вторая производная неотрицательна* ($f''(x) \geq 0\ \forall\ x \in \mathbb{R}$)*.*

Приведем несколько примеров выпуклых функций.

1. $f(x) = e^{ax}$, $a \in \mathbb{R}$.

2. $f(x) = |x|^p$, $p \geq 1$.

Выпуклость функций из примеров 1-2 сразу следует из теоремы 1 и определения выпуклой функции.

3. *Аффинная функция* (в многомерном случае $f\colon \mathbb{R}^n \to \mathbb{R}$, $f(x) = \langle a, x\rangle + b$, $a \in \mathbb{R}^n$, $b \in \mathbb{R}$).

Аффинная функция 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 можно определять также из следующего многомерного обобщения теоремы 1.

Теорема 2 (Р, с. 44)**.** *Пусть функция* $f\colon \mathbb{R}^n \to \mathbb{R}$ *дважды непрерывно дифференцируема* ($f \in D^2(\mathbb{R}^n)$)*. Тогда она выпукла тогда и только тогда, когда ее матрица вторых производных (гессиан) всюду неотрицательно определена*

$$\left(f''(x) = \left(\frac{\partial^2 f(x)}{\partial x_i \partial x_j} \right)_{i,j=1}^{n} \geq 0\ \forall\ x \in \mathbb{R}^n \right).$$

4. *Квадратичная функция* $Q(x) = \langle Ax, x\rangle$ (A — симметричная матрица), является выпуклой тогда и только тогда, когда матрица A неотрицательно определена.

Это сразу вытекает из теоремы 2.

Выпуклыми функциями многих переменных (функционалами) являются также следующие функции:

5. *Функция нормы*

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

6. *Индикаторная функция* выпуклого множества $A \subset X$

$$\delta A(x) = \begin{cases} 0, & x \in A, \\ +\infty, & x \notin A. \end{cases}$$

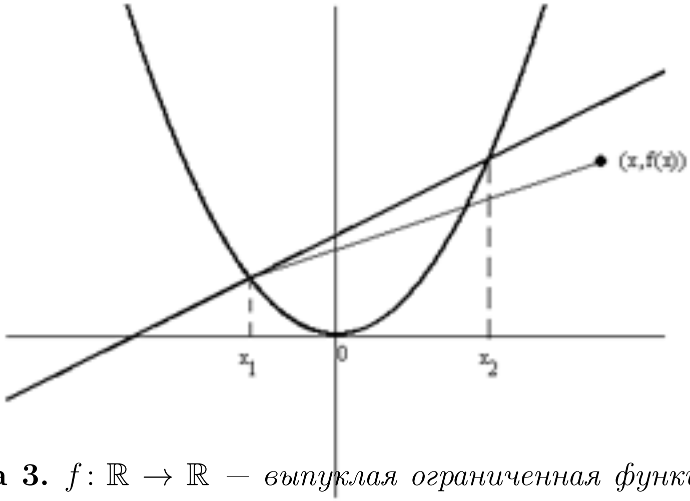
7. *Функция Минковского* выпуклого множества $A \subset X$

$$\mu A(x) = \begin{cases} 0, & \alpha^{-1}x \in A \ \forall\ \alpha > 0, \\ +\infty, & \alpha^{-1}x \notin A \ \forall\ \alpha > 0, \\ \inf\{\alpha \mid \alpha > 0,\ \alpha^{-1}x \in A\}, & \text{иначе.} \end{cases}$$

Функция Минковского означает наименьшее число, в которое надо уменьшить вектор x, чтобы он попал в множество A.

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

Доказательство. Предположим противное, что $f(x) \neq \mathrm{const}$. Тогда существуют точки x_1, x_2 такие, что $x_1 < x_2$ и $f(x_1) \neq f(x_2)$. Поскольку f — выпуклая ограниченная функция, то надграфик функции $\operatorname{epi} f$ — выпуклое множество, следовательно, отрезок $\big[(x_1, f(x_1)), (x_2, f(x_2))\big] \subset \operatorname{epi} f$. Значит, на промежутке $[x_1, x_2]$ график функции f лежит ниже этого отрезка (или совпадает с ним). Не ограничивая общности, будем считать, что $f(x_1) < f(x_2)$. Пусть $l(x)$ — прямая, проходящая через точки $(x_1, f(x_1))$ и $(x_2, f(x_2))$. Покажем, что

$$f(x) \geq l(x) \quad \forall\, x > x_2. \tag{$*$}$$

Если это не так, то существует точка $x > x_2$ такая, что $f(x) < l(x)$, т. е. точка $(x, f(x))$ лежит ниже прямой l. Так как функция 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 $\operatorname{epi} f$). Поэтому неравенство $(*)$ действительно выполняется.

Так как $l(x) \to +\infty$ при $x \to +\infty$, то $f(x) \to +\infty$ при $x \to +\infty$. Получаем противоречие с ограниченностью функции f. Значит, предположение противного неверно и $f(x) = \mathrm{const}$. $\square$

4.1.3 Субдифференциал выпуклой функции

Дадим определение важного понятия выпуклого анализа — субдифференциала функции, обобщающего для выпуклых функций понятие производной в гладком анализе.

Субдифференциалом выпуклой функции f в точке $\hat{x}$ называется следующее множество в сопряженном пространстве X^*:

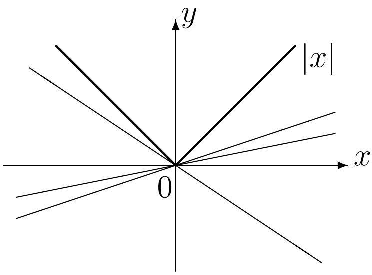
$$\partial f(\hat{x}) = \{x^* \in X^* \mid \langle x^*, x - \hat{x}\rangle \leq f(x) - f(\hat{x}) \quad \forall\, x \in X\}.$$

Напомним, что *сопряженным пространством* X^* называется пространство линейных непрерывных функционалов на X. В случае $X = \mathbb{R}^n$ сопряженное пространство $(\mathbb{R}^n)^* = \mathbb{R}^n$.

Из определения сразу вытекает, что субдифференциал — выпуклое множество в X^*. Легко доказать, что оно замкнуто. Субдифференциал дифференцируемой функции совпадает с ее производной.

Для функции одной переменной субдифференциал $\partial f(\hat{x})$ — это совокупность угловых коэффициентов k, при которых прямые $y = kx + b$, проходящие через точку $(\hat{x}, f(\hat{x}))$, лежат под графиком функции $y = f(x)$.

Пример. $f(x) = |x|$.



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

Рис. 6

Для субдифференциала суммы функций имеет место теорема аналогичная теореме о производной суммы функций.

Теорема (Моро–Рокафеллар). *[16, с. 49] Пусть f_1 и f_2 — выпуклые функции на X. Существует точка x_0, в которой функция f_1 конечна ($|f_1(x_0)| < \infty$), а функция f_2 непрерывна ($f \in C(x_0)$). Тогда*

$$\partial(f_1 + f_2)(x) = \partial f_1(x) + \partial f_2(x) \quad \forall\ x.$$

При доказательстве необходимых условий экстремума в гладкой задаче с равенствами и неравенствами нам понадобится следующая теорема о субдифференциале максимума.

Теорема (Дубовицкий–Милютин). *[16, с. 51] Пусть f_1 и f_2 — выпуклые функции, непрерывные в точке $\hat{x}$, $f_1(\hat{x}) = f_2(\hat{x})$. Тогда*

$$\partial \max\{f_1, f_2\}(\hat{x}) = \operatorname{conv}\big(\partial f_1(\hat{x}) \cup \partial f_2(\hat{x})\big).$$

Важным примером выпуклой функции является сублинейная функция. Функция p называется *сублинейной*, если ее надграфик является выпуклым конусом с вершиной в нуле.

Из неравенства Йенсена следует, что собственная функция является сублинейной тогда и только тогда, когда
a) $p(\lambda x) = \lambda p(x)$ для любого $\lambda > 0$, для любого $x \in X$;
b) $p(x_1 + x_2) \le p(x_1) + p(x_2)$ для любых $x_1, x_2 \in X$.

Если X — нормированное пространство, то $x \to \|x\|$ — сублинейная функция.

Сформулируем теорему Моро–Рокафеллара и теорему Дубовицкого–Милютина для сублинейных функций.

Теорема (Моро–Рокафеллар). *Пусть p_1, p_2 — сублинейные функции, функция p_1 непрерывна, функция p_2 замкнута. Тогда в точке $\hat{x} = 0$*

$$\partial(p_1 + p_2) = \partial p_1 + \partial p_2.$$

Теорема (Дубовицкий–Милютин). *Пусть p_1, p_2 — непрерывные сублинейные функции. Тогда*

$$\partial \max\{p_1, p_2\}(0) = \operatorname{conv}(\partial p_1(0) \cup \partial p_2(0)).$$

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 $|f(x)| = \max\{-f(x), f(x)\}$. Тогда по теореме Дубовицкого–Милютина

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

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

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, функционал $f\colon X \to \mathbb{R}$, $f(x) = \|x\|$. Найдем субдифференциал нормы в нуле.

По определению субдифференциала

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

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

$$x^* \in \partial f(0) \iff \langle x^*, x\rangle \le \|x\| \quad \forall\; x \in X. \tag{$*$}$$

Пусть $x^* \in \partial f(0)$. По лемме Банаха существует вектор $x \in X$ такой, что $\|x\| = 1$, $\langle x^*, x\rangle = \|x^*\|$. Подставляя этот вектор x в неравенство $(*)$, получим

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

Значит, $\partial f(0) \subset B^* := \big\{x^* \in X^* \mid \|x^*\|_{X^*} \le 1\big\}$ — единичный шар сопряженного пространства. Докажем, что $\partial f(0) = B^*$.

Для этого надо доказать обратное включение $\partial f(0) \supset B^*$. Возьмем $x^* \in B^*$, тогда $|x^*| \le 1$. По неравенству Коши–Буняковского

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

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