4 Выпуклые задачи

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. 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 $a_1, \dots, a_m \in X$. 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* $\operatorname{conv}\{a_1, \dots, a_m\} := \left\{a = \sum\limits_{i=1}^{m} t_i a_i \,\middle|\, t_i \ge 0,\ \sum\limits_{i=1}^{m} t_i = 1\right\}$ — 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*
  $\operatorname{aff}\{a_1,\dots,a_m\} := \left\{ a = \sum\limits_{i=1}^{m} t_i a_i \,\middle|\, \sum\limits_{i=1}^{m} t_i = 1 \right\}.$

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*:

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

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

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

Теорема 1. *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* $f(x) = \langle a, x \rangle + b$, $a \in \mathbb{R}^n$, $b \in \mathbb{R}$, в многомерном случае ($f\colon \mathbb{R}^n \to \mathbb{R}$).

Аффинная функция является выпуклой по неравенству Йенсена.

Выпуклость функций нескольких переменных можно определять также из следующего многомерного обобщения теоремы 1.

Теорема 2. *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* $Q(x) = \langle Ax, x \rangle$, где 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*

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

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 $A \subset X$

$sA(x^*) = \max\limits_{x \in A} \langle x^*, x \rangle$.

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^*:

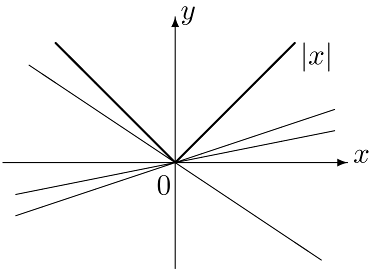
$$\partial f(\hat{x}) = \{x^* \in X^* \mid \langle x^*, x - \hat{x} \rangle \le f(x) - f(\hat{x}) \quad \forall\, x \in X\}.$$

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. $f(x) = |x|$.



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

Рис. 6

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

Теорема Моро–Рокафеллара. [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 Дубовицкого–Милютина. [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) $p(x_1 + x_2) \leq p(x_1) + p(x_2)$ 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–Рокафеллара. 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 Дубовицкого–Милютина. *Пусть p_1, p_2 — непрерывные сублинейные функции. Тогда*

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

4.2 Теоремы отделимости

При выводе необходимых условий экстремума (принципа Лагранжа) в выпуклых задачах и в задачах с равенствами и неравенствами мы будем использовать свойство отделимости непересекающихся выпуклых множеств.

Определение 1. Множества A и B из пространства X называются *отделимыми*, если существует линейный непрерывный функционал $\lambda \in X^*$, $\lambda \neq 0$, для которого

$$\inf_{a \in A} \langle \lambda, a \rangle \geq \sup_{b \in B} \langle \lambda, b \rangle.$$

Из определения следует, что множества являются отделимыми, если можно провести гиперплоскость $H = \{x \in X \mid \langle \lambda, x \rangle = c\}$, $c \in \mathbb{R}$, так что одно из множеств лежит в одном замкнутом полупространстве $H_+ = \{x \in X \mid \langle \lambda, x \rangle \geq c\}$, а другое — в другом замкнутом полупространстве $H_- = \{x \in X \mid \langle \lambda, x \rangle \leq c\}$.

Определение 2. Множества A и B называются *строго отделимыми*, если существует линейный непрерывный функционал $\lambda \in X^*$, для которого

$$\inf_{a \in A} \langle \lambda, a \rangle > \sup_{b \in B} \langle \lambda, b \rangle.$$

Приведем результаты об отделимости в конечномерном случае и в случае линейных нормированных пространств.

Теорема 1 (отделимости множеств в конечномерном пространстве). *Пусть A и B — непустые выпуклые множества в $\mathbb{R}^n$, $A \cap B = \emptyset$. Тогда множества A и B отделимы.*

Теорема 1′ (отделимости множеств в нормированном пространстве). *Пусть A и B — непустые выпуклые множества в X, $\operatorname{int} A \neq \emptyset$, $\operatorname{int} A \cap B = \emptyset$. Тогда множества A и B отделимы.*

В теореме отделимости в конечномерном пространстве $\operatorname{int} A$ может быть пуста.

Теорема 2 (о строгой отделимости точки от множества в нормированном пространстве). *Пусть A — непустое выпуклое замкнутое множество в X, $b \notin A$. Тогда точку b можно строго отделить от множества A.*

4.3 Задачи без ограничений

Пусть $f\colon X \to \overline{\mathbb{R}}$ — выпуклая функция, отображающая нормированное пространство X в расширенную прямую. *Выпуклой задачей без ограничений* называется следующая экстремальная задача:

$$f(x) \to \min\,. \tag{P}$$

Теорема (аналог теоремы Ферма). *Для того чтобы точка $\hat{x}$ доставляла в выпуклой задаче без ограничений (P) абсолютный минимум ($\hat{x} \in absmin\, P$), необходимо и достаточно, чтобы выполнялось соотношение*

$$0 \in \partial f(\hat{x}).$$

Доказательство. $\hat{x} \in \operatorname{absmin} P \Leftrightarrow f(x) - f(\hat{x}) \geq 0 = \langle 0, x - \hat{x}\rangle\ \forall\, x \in X \Leftrightarrow 0 \in \partial f(\hat{x})$.

Поскольку из выпуклости функции f не следует, вообще говоря, выпуклость функции $-f$, то существенно, что рассматривается задача на минимум, а не на максимум.

4.4 Задачи с ограничением

Пусть $f\colon X \to \overline{\mathbb{R}}$ — выпуклая функция, отображающая нормированное пространство X в расширенную прямую, $D \subset X$ — выпуклое множество. *Выпуклой задачей с ограничением* (или просто *выпуклой задачей*) называется следующая экстремальная задача:

$$f(x) \to \min; \quad x \in D. \tag{P}$$

Теорема. *Пусть $\hat{x} \in locmin\, P$ — доставляет локальный минимум в выпуклой задаче (P). Тогда $\hat{x} \in absmin\, P$ — доставляет абсолютный минимум в этой задаче.*

Доказательство. Пусть $\hat{x} \in \operatorname{locmin} P$. Это означает, что существует окрестность U точки $\hat{x}$ такая, что

$$f(\hat{x}) \leq f(x) \quad \forall\, x \in U \cap D. \tag{$*$}$$

Возьмем произвольную точку $x \in D$. Тогда, поскольку x и $\hat{x}$ из D, то выпуклая комбинация $\bar{x} = (1-\alpha)\hat{x} + \alpha x \in U \cap D$ при достаточно малом $\alpha > 0$. Следовательно, $f(\hat{x}) \overset{(*)}{\leq} f(\bar{x}) \overset{\mathrm{def}\, \bar{x}}{=} f((1 -$

$\alpha)\hat{x} + \alpha x) \le$ (по неравенству Иенсена) $(1-\alpha)f(\hat{x}) + \alpha f(x)$, откуда $\alpha f(\hat{x}) \le \alpha f(x) \Leftrightarrow f(\hat{x}) \le f(x)\ \forall\ x \in D$. Значит, $\hat{x}$ доставляет абсолютный минимум в задаче (P). $\triangleright$

Из теоремы следует, что в выпуклой задаче локальный минимум является и абсолютным (глобальным). Поэтому в дальнейшем в выпуклых задачах, говоря “минимум”, имеем в виду абсолютный минимум.

4.5 Задача выпуклого программирования

Пусть $f_i\colon X \to \overline{\mathbb{R}}$, $i = 0, 1, \ldots, m$, — выпуклые функции, отображающие линейное нормированное пространство X в расширенную прямую, $A \subset X$ — выпуклое множество. *Задачей выпуклого программирования* называется следующая экстремальная задача:

$$f_0(x) \to \min; \quad f_i(x) \le 0,\ i = 1, \ldots, m,\ x \in A. \qquad (P)$$

Точка x называется *допустимой* в задаче (P), если $x \in A$ и выполняются все заданные ограничения типа неравенств.

Упражнение. Докажите, что задача выпуклого программирования является выпуклой задачей, т. е., что множество допустимых элементов в этой задаче является выпуклым множеством.

При проверке достаточных условий минимума в задаче выпуклого программирования будет использоваться некоторое условие регулярности множества допустимых элементов — условие Слейтера. Множество допустимых элементов в задаче (P) удовлетворяет *условию Слейтера*, если существует точка $\bar{x} \in A$, для которой $f_i(\bar{x}) < 0$, $i = 1, \ldots, m$.

Теорема Каруша-Куна–Таккера. 1. *Если $\hat{x} \in absmin\, P$ — решение задачи выпуклого программирования, то найдется ненулевой вектор множителей Лагранжа $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_m) \in \mathbb{R}^{m+1}$ такой, что для функции Лагранжа $\mathscr{L}(x) = \sum\limits_{i=0}^{m} \lambda_i f_i(x)$ выполняются условия:*

a) принцип минимума функции Лагранжа: $\min\limits_{x \in A} \mathscr{L}(x) = \mathscr{L}(\hat{x})$;

b) дополняющей нежесткости: $\lambda_i f_i(\hat{x}) = 0,\ i = 1, \ldots, m$;

c) неотрицательности: $\lambda_i \ge 0,\ i = 0, 1, \ldots, m$.

2. *Если для допустимой точки $\hat{x}$ выполнены условия а) – с) и $\lambda_0 \neq 0$, то $\hat{x} \in absmin\, P$.*

3. *Если для допустимой точки $\hat{x}$ выполнены условия а) – с) и множество допустимых элементов удовлетворяет условию Слейтера, то $\hat{x} \in absmin\, P$.*

Доказательство. Пусть $\hat{x} \in \operatorname{absmin} P$. Не ограничивая общности, считаем, что $f_0(\hat{x}) = 0$ — иначе введем новую функцию $\tilde{f}_0(x) = f_0(x) - f_0(\hat{x})$. Положим

$$B = \{b = (b_0, \ldots, b_m) \in \mathbb{R}^{m+1} \mid \exists\, x \in A : f_i(x) \leq b_i,\ i = 0, \ldots, m\}.$$

Покажем, что B — *непустое выпуклое множество.* Действительно, неотрицательный октант $\mathbb{R}_+^{m+1} \subset B$, т. е. любой вектор с неотрицательными компонентами принадлежит B, ибо в определении множества B можно положить $x = \hat{x}$. Докажем выпуклость множества B. Пусть точки b и b' принадлежат множеству B. Надо доказать, что точка $\alpha b + (1-\alpha)b' \in B$ для любого $\alpha \in (0, 1)$. Поскольку точки $b, b' \in B$, то по определению множества B существуют точки $x, x' \in A$ такие, что $f_i(x) \leq b_i$, $f_i(x') \leq b_i'$, $i = 0, 1, \ldots, m$. Положим $x_\alpha = \alpha x + (1-\alpha)x'$. Тогда $x_\alpha \in A$, поскольку множество A — выпукло, а ввиду выпуклости функций f_i, $i = 0, 1, \ldots, m$, по неравенству Йенсена

$$f_i(x_\alpha) = f_i\big(\alpha x + (1-\alpha)x'\big) \leq \alpha f_i(x) + (1-\alpha) f_i(x') \leq \alpha b_i + (1-\alpha)b_i',$$

т. е. точка $\alpha b + (1-\alpha)b' \in B$.

Докажем, что точка $0 \notin \operatorname{int} B$. Если бы $0 \in \operatorname{int} B$, то при достаточно малом $\varepsilon < 0$ точка $(\varepsilon, 0, \ldots, 0) \in B$. Тогда по определению множества B существует точка $\tilde{x} \in A$ такая, что $f_0(\tilde{x}) \leq \varepsilon < 0 = f_0(\hat{x})$, $f_i(\tilde{x}) \leq 0$, $i = 1, \ldots, m$. Значит, $\tilde{x}$ является допустимой точкой ($\tilde{x} \in D(P)$) и $f_0(\tilde{x}) < f_0(\hat{x})$, т. е. $\hat{x} \notin \operatorname{absmin} P$ — противоречие. Следовательно, $0 \notin \operatorname{int} B$.

По теореме отделимости в конечномерном пространстве множество B и точку 0 можно отделить, т. е. существует вектор $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_m) \neq 0$, для которого

$$\inf_{b \in B} \sum_{i=0}^{m} \lambda_i b_i \geq \langle \lambda, 0 \rangle = 0.$$

Таким образом,

$$\sum_{i=0}^{m} \lambda_i b_i \geq 0 \quad \forall\, b \in B\ . \tag{$*$}$$

Из неравенства ($*$) будут выведены условия неотрицательности, дополняющей нежесткости и принцип минимума.

1. *Условие неотрицательности.* Поскольку, как мы уже говорили, любой вектор с неотрицательными компонентами принадлежит B, то вектор $(0,\ldots,0,1,0,\ldots,0) \in B$, где единица стоит на i-ом месте (счет начинаем с нуля). Подставив эту точку в неравенство ($*$), получим, что $\lambda_i \geq 0$.

Условие дополняющей нежесткости. Нетрудно видеть, что точка $(0,\,\ldots,0,f_i(\hat{x}),0,\ldots,\,0) \in B$. Действительно, в определении множества B возьмем $x = \hat{x}$, тогда $x \in A$ и нужные неравенства выполняются: $f_j(\hat{x}) \leq 0$, $j \neq i$, $f_i(\hat{x}) \leq f_i(\hat{x})$. Подставив эту точку в неравенство ($*$), имеем $\lambda_i f_i(\hat{x}) \geq 0$. Если $\lambda_i f_i(\hat{x}) > 0$, то (так как по уже доказанному условию неотрицательности $\lambda_i \geq 0$) $\lambda_i > 0 \Rightarrow f_i(\hat{x}) > 0$ — это противоречит допустимости точки $\hat{x}$. Значит, $\lambda_i f_i(\hat{x}) = 0$.

Принцип минимума. Возьмем точку $x \in A$, тогда точка $(f_0(x),$ $f_1(x),\,\ldots,f_m(x)) \in B$. Поэтому по неравенству ($*$)

$$\mathscr{L}(x) = \sum_{i=0}^{m} \lambda_i f_i(x) \geq 0 = \sum_{i=0}^{m} \lambda_i f_i(\hat{x}) = \mathscr{L}(\hat{x}),$$

так как, не ограничивая общности, положили $f_0(\hat{x}) = 0$ и по уже доказанному условию дополняющей нежесткости $\lambda_i f_i(\hat{x}) = 0$, $i = 1,\ldots,m$. Таким образом, принцип минимума для функции Лагранжа доказан.

2. Пусть для допустимой точки $\hat{x}$ выполнены условия a)–c) и $\lambda_0 \neq 0$. Положим $\lambda_0 = 1$. Тогда для любой допустимой точки x будет выполняться неравенство

$$f_0(\hat{x}) \overset{b)}{=} f_0(\hat{x}) + \sum_{i=1}^{m} \lambda_i f_i(\hat{x}) \overset{\mathrm{def}}{=} \mathscr{L}(\hat{x}) \overset{a)}{\leq} \mathscr{L}(x) \overset{\mathrm{def}}{=} f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) \overset{c)}{\leq} f_0(x)$$

(в последней сумме все слагаемые $\lambda_i f_i(x)$ неположительны: $\lambda_i \geq 0$, $f_i(x) \leq 0$). Неравенство $f_0(\hat{x}) \leq f_0(x)$ для любой допустимой точки x означает, что $\hat{x} \in \operatorname{absmin} P$.

3. Пусть для допустимой точки $\hat{x}$ выполнены условия a)–c) и множество допустимых элементов удовлетворяет условию Слейтера в точке $\bar{x}$ ($\exists\ \bar{x} \in A : f_i(\bar{x}) < 0,\ i = 1, \ldots, m$). Предположим, что при этом $\lambda_0 = 0$. Так как вектор $\lambda \neq 0$, то в силу условия неотрицательности существует хотя бы одно $\lambda_j > 0$, $j \in \{1, \ldots, m\}$. Следовательно, имеем

$$\mathscr{L}(\bar{x}) = \sum_{i=1}^{m} \lambda_i f_i(\bar{x}) \overset{c)}{\leq} \lambda_j f_j(\bar{x}) < 0 \overset{b)}{=} \sum_{i=1}^{m} \lambda_i f_i(\hat{x}) = \mathscr{L}(\hat{x})$$

(отметим, что в первой сумме все слагаемые $\lambda_i f_i(\bar{x})$ в силу условия c) неположительны: $\lambda_i \geq 0$, а $f_i(\bar{x}) < 0$). Но полученное неравенство $\mathscr{L}(\bar{x}) < \mathscr{L}(\hat{x})$ противоречит выполненному условию a). Значит, наше предположение, что $\lambda_0 = 0$ неверно. Поэтому $\lambda_0 \neq 0$ и по доказанному п. 2 $\hat{x} \in \operatorname{absmin} P$.

Теорема Каруша-Куна–Таккера полностью доказана. $\triangleright$

Из теоремы Каруша-Куна–Таккера можно вывести достаточные условия абсолютного минимума в выпуклой задаче с неравенствами без ограничения типа включения. Рассмотрим задачу

$$f_0(x) \to \min; \quad f_i(x) \leq 0,\ i = 1, \ldots, m. \qquad (P')$$

Теорема (достаточные условия абсолютного минимума в выпуклой задаче с неравенствами без ограничения типа включения). *Пусть $f_i\colon X \to \overline{\mathbb{R}}$, $i = 0, \ldots, m$, — выпуклые функции, $\hat{x}$ — допустимая точка в задаче (P') ($\hat{x} \in D_{P'}$), для функции Лагранжа $\mathscr{L}(x) = \sum\limits_{i=0}^{m} \lambda_i f_i(x)$ с множителем Лагранжа $\lambda_0 > 0$ в точке $\hat{x}$ выполняются условия:*

a) условие стационарности функции Лагранжа: $\quad 0 \in \partial\mathscr{L}(\hat{x})$;

b) дополняющей нежесткости: $\quad \lambda_i f_i(\hat{x}) = 0,\ i = 1, \ldots, m$;

c) неотрицательности: $\quad \lambda_i \geq 0,\ i = 1, \ldots, m$.

Тогда $\hat{x} \in \mathit{absmin}\, P'$.

Доказательство. Поскольку $f_i\colon X \to \overline{\mathbb{R}}$, $i = 0, \ldots, m$, — выпуклые функции, то функция Лагранжа $\mathscr{L}(x) = \sum\limits_{i=0}^{m} \lambda_i f_i(x)$ с неотрицательными множителями Лагранжа является выпуклой функцией. По аналогу теоремы Ферма для выпуклых функций условие $0 \in \partial\mathscr{L}(\hat{x})$ является необходимым и достаточным условием абсолютного минимума функции Лагранжа в точке $\hat{x}$. Значит, условие стационарности функции Лагранжа в настоящей теореме равносильно принципу минимума для функции Лагранжа в теореме Каруша-Куна–Таккера. Таким образом, для нашей задачи выполняются соотношения $a) - c)$ теоремы Каруша-Куна–Таккера с множителем Лагранжа $\lambda_0 \neq 0$. По п. 2 теоремы Каруша-Куна–Таккера следует, что $\hat{x} \in \operatorname{absmin} P'$. $\triangleright$

Легко видеть, что теорема остается верной и для задач с ограничениями типа равенств и неравенств, если функции, задающие равенства, являются аффинными. (В этом случае, как мы говорили об этом в п. 3. 2 равенство $f(x) = 0 \Leftrightarrow 0 \leq f(x) \leq 0$ заменяется двумя неравенствами $f(x) \leq 0$, $-f(x) \leq 0$. Из аффинности функции f следует выпуклость функций f и $-f$.)

4.6 Пример

$$f(x_1, x_2) = x_1^2 + x_1x_2 + x_2^2 + 3|x_1 + x_2 - 2| \to \min .$$

Решение. Функция $f(x)$ является выпуклой функцией как сумма двух выпуклых функций. Действительно, функция $g(x) = x_1^2 + x_1x_2 + x_2^2$ выпукла по теореме 2 п 4.1, так как по критерию Сильвестра матрица вторых производных

$$g''(x) = \left(\frac{\partial^2 g(x)}{\partial x_i \partial x_j}\right)_{i,j=1}^2 = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

положительно определена и не зависит от x. Очевидно, что функция $h(x_1, x_2) = |x_1 + x_2 - 2|$, являющаяся модулем линейной функции также является выпуклой функцией.

Необходимое и достаточное условие экстремума в выпуклой задаче без ограничений:

$$0 \in \partial f(\hat{x}) = \partial g(\hat{x}) + 3\partial h(\hat{x}) \tag{1}$$

(использовали теорему Моро-Рокафеллара о субдифференциале суммы функций). Найдем субдифференциалы функций g и h. Для дифференцируемой функции ее субдифференциал совпадает с производной. Поэтому $\partial g(\hat{x}) = (2\hat{x}_1 + \hat{x}_2, \hat{x}_1 + 2\hat{x}_2)$.

Поскольку $h(x) = |x_1+x_2-2| = \begin{cases} -x_1-x_2+2, & x_1+x_2-2<0, \\ 0, & x_1+x_2-2=0, \\ x_1+x_2-2, & x_1+x_2-2>0, \end{cases}$

то $\partial h(\hat{x}) = \begin{cases} (-1,-1), & x_1+x_2-2<0, \\ (1,1), & x_1+x_2-2>0. \end{cases}$

Найдем $\partial h(\hat{x})$ в точке недифференцируемости функции h, т. е. при $\hat{x}_1+\hat{x}_2-2=0$. По определению субдифференциала

$$a=(a_1,a_2) \in \partial h(\hat{x}) \iff \langle a, x-\hat{x}\rangle \le h(x)-h(\hat{x}) \quad \forall\, x \in \mathbb{R}^2$$

$$\iff a_1(x_1-\hat{x}_1)+a_2(x_2-\hat{x}_2) \le |x_1+x_2-2|-|\hat{x}_1+\hat{x}_2-2|$$

$$\iff a_1x_1+a_2x_2 \le |x_1+x_2-2|+a_1\hat{x}_1+a_2\hat{x}_2 \quad \forall\, x_1,x_2. \qquad (2)$$

Полагая в неравенстве (2) $x_2=0$, получим, что $a_1x_1 \le |x_1-2| + a_1\hat{x}_1+a_2\hat{x}_2$ для любого x_1, откуда вытекает, что $|a_1| \le 1$. Полагая $x_1=-x_2$, получим, что $(a_2-a_1)x_2 \le 2+a_1\hat{x}_1+a_2\hat{x}_2$ для любого x_2, откуда вытекает, что $a_1=a_2$. Таким образом, $a=(\alpha,\alpha)$, $|\alpha|\le 1$. При таких значениях a в точках $\hat{x}_1,\hat{x}_2$ таких, что $\hat{x}_1+\hat{x}_2-2=0$ неравенство (2) будет выполняться:

$$\alpha(x_1+x_2) \le |x_1+x_2-2|+\alpha(\hat{x}_1+\hat{x}_2) \overset{\hat{x}_1+\hat{x}_2=2}{\iff}$$

$$\alpha(x_1+x_2-2) \le |x_1+x_2-2| \quad \forall\, x_1,x_2.$$

Таким образом,

$$\partial h(x) = \begin{cases} (-1,-1), & x_1+x_2-2<0, \\ (\alpha,\alpha),\ |\alpha|\le 1, & x_1+x_2-2=0, \\ (1,1), & x_1+x_2-2>0. \end{cases}$$

Значит,

$$\partial f(x) = \partial g(x) + 3\partial h(x) =$$

$$= (2x_1+x_2, x_1+2x_2) + 3\begin{cases} (-1,-1), & x_1+x_2-2<0, \\ (\alpha,\alpha),\ |\alpha|\le 1, & x_1+x_2-2=0, \\ (1,1), & x_1+x_2-2>0, \end{cases} =$$

$$= \begin{cases} (2x_1+x_2-3, x_1+2x_2-3), & x_1+x_2-2<0, \\ (2x_1+x_2+3\alpha, x_1+2x_2+3\alpha),\ |\alpha|\le 1, & x_1+x_2-2=0, \\ (2x_1+x_2+3, x_1+2x_2+3), & x_1+x_2-2>0. \end{cases}$$

Следовательно, соотношение (1) запишется в виде

$$\begin{cases} 2x_1 + x_2 - 3 = 0, \\ x_1 + 2x_2 - 3 = 0, \end{cases} \qquad \text{при} \quad x_1 + x_2 - 2 < 0, \tag{i}$$

$$\begin{cases} 2x_1 + x_2 + 3\alpha = 0, \\ x_1 + 2x_2 + 3\alpha = 0, \end{cases} \qquad \text{при} \quad x_1 + x_2 - 2 = 0, \quad (|\alpha| \le 1) \tag{ii}$$

$$\begin{cases} 2x_1 + x_2 + 3 = 0, \\ x_1 + 2x_2 + 3 = 0, \end{cases} \qquad \text{при} \quad x_1 + x_2 - 2 > 0. \tag{iii}$$

В случае (i) нет критических точек, так как из уравнений системы следует, что $x_1 + x_2 = 2$ — противоречие с условием $x_1 + x_2 - 2 < 0$.
В случае (iii) также нет критических точек, так как из уравнений системы следует, что $x_1 + x_2 = -2$ — противоречие с условием $x_1 + x_2 - 2 > 0$.
В случае (ii) система из трех уравнений с тремя неизвестными имеет единственное решение $\hat{x}_1 = \hat{x}_2 = 1$, $\alpha = -1$.

Ответ. $\hat{x} = (\hat{x}_1, \hat{x}_2) = (1, 1) \in \text{absmin}$, $S_{\text{absmin}} = 3$.