ОПТИМИЗАЦИЯ

ЧАСТЬ I

ВЫПУКЛЫЙ АНАЛИЗ

К. Ю. ОСИПЕНКО

Содержание

Лекция 1

1. Введение

Выпуклый анализ — раздел математики, в котором изучаются выпуклые множества, выпуклые функции и выпуклые экстремальные задачи. Выпуклый анализ имеет весьма разнообразные приложения в технике и экономике. Многие приложения связаны с задачами оптимизации, которые при наличии выпуклости могут решаться более просто.

Пристальное внимание к выпуклости было привлечено в сороковых годах прошлого столетия. Оно было связано с появлением такого раздела, как линейное программирование. Первые нетривиальные задачи в этом направлении были решены Леонидом Витальевичем Канторовмчем. В 1939 году вышла его книга "Математические методы организации и планирования производства". Затем в пятидесятых годах мощное развитие линейного программирования, а потом и выпуклой оптимизации началось в США. Это было вызвано проблемами экономики и различными приложениями в военных областях. В 1975 году Канторовичу и Купмансу (американскому экономисту) была присуждена Нобелевская премия по экономике за вклад в теорию оптимального распределения ресурсов.

Следует упомянуть американского математика-прикладника Данцига, которому принадлежит замечательный алгоритм решения задач линейного программирования, называемый симплекс-метод.

В 1970 году американский математик Рокафеллар опубликовал монографию, которая называлась "Выпуклый анализ". Тогда впервые и появилось это словосочетание. Сам Рокафеллар пишет, что это название предложил ему профессор Принстонского университета Таккер. С тех пор термин "выпуклый анализ" стал общеупотребительным.

2. Выпуклые множества

Пусть X — линейное пространство. Если $x,y\in X$, то множество

$$[x,y]=\{\,z\in X: z=(1-\alpha)x+\alpha y,\ 0\le\alpha\le 1\,\}$$

называется отрезком (соединяющим точки x и y).

Непустое множество A называется выпуклым, если для любых $x,y\in A$ $[x,y]\in A$. Пустое множество считается выпуклым по определению.

Пусть $x_1,\dots,x_n\in X$. Точка

$$x=\sum_{j=1}^{n}\alpha_j x_j,\quad \alpha_j\ge 0,\quad \sum_{j=1}^{n}\alpha_j=1, \tag{1}$$

называется выпуклой комбинацией точек $x_1, \dots, x_n$.

Предложение 1. *Если A — выпуклое множество, то любая выпуклая комбинация точек $x_1, \dots, x_n \in A$ принадлежит A.*

Доказательство. Доказательство проведем индукцией по числу точек. При $n = 1$ утверждение очевидно. Предположим, что утверждение верно для любой выпуклой комбинации из не более, чем $n - 1$ точек. Пусть имеется точка x вида (1). Если $\alpha_1 = 1$, то $x = x_1 \in A$. Если $\alpha_1 < 1$, положим

$$\widetilde{\alpha}_j = \frac{\alpha_j}{1 - \alpha_1}, \quad j = 2, \dots, n.$$

В силу того, что

$$\sum_{j=2}^{n} \widetilde{\alpha}_j = 1,$$

из предположения индукции вытекает, что

$$\widetilde{x} = \sum_{j=2}^{n} \widetilde{\alpha}_j x_j \in A.$$

Тогда из выпуклости A следует, что

$$x = \sum_{j=1}^{n} \alpha_j x_j = \alpha_1 x_1 + (1 - \alpha_1) \sum_{j=2}^{n} \frac{\alpha_j}{1 - \alpha_1} x_j = \alpha_1 x_1 + (1 - \alpha_1)\widetilde{x} \in A.$$

□

Выпуклой оболочкой $\operatorname{co} A$ множества $A \subset X$ называется множество всех выпуклых комбинаций его точек

$$\operatorname{co} A = \left\{ x \in X : x = \sum_{j=1}^{n} \alpha_j x_j,\ x_1, \dots, x_n \in A,\ \alpha_j \geq 0, \right.$$
$$\left. \sum_{j=1}^{n} \alpha_j = 1,\ n \in \mathbb{N} \right\}.$$

Среди всех выпуклых множеств, содержащих данное множество $A \subset X$, существует наименьшее, т.е. то, которое содержится в любом выпуклом множестве, содержащем A. Действительно, семейство всех выпуклых множеств, содержащих A, не пусто (в него входит, например, все пространство X). Пересечение всех множеств этого семейства будет выпуклым (т.к. пересечение любого семейства выпуклых множеств — выпукло), содержащим A и наименьшим. Докажем, что это множество совпадает с $\operatorname{co} A$.

Предложение 2. *Множество* $\operatorname{co} A$ *является наименьшим выпуклым множеством, содержащим A.*

Доказательство. Из предложения 1 вытекает, что если некоторое выпуклое множество содержит A, то оно содержит любую выпуклую комбинацию точек из A, а значит, оно содержит $\operatorname{co} A$. Следовательно, пересечение всех выпуклых множеств содержит $\operatorname{co} A$. Осталось показать что $\operatorname{co} A$ выпукло, в этом случае оно совпадет с пересечением всех выпуклых множеств, содержащих A. Рассмотрим две произвольные точки $\operatorname{co} A$

$$x = \sum_{j=1}^{n} \alpha_j x_j, \quad y = \sum_{j=1}^{m} \beta_j y_j.$$

Для любого $\alpha \in [0,1]$ имеем

$$(1-\alpha)x + \alpha y = \sum_{j=1}^{n} (1-\alpha)\alpha_j x_j + \sum_{j=1}^{m} \alpha\beta_j y_j.$$

Так как

$$\sum_{j=1}^{n} (1-\alpha)\alpha_j + \sum_{j=1}^{m} \alpha\beta_j = (1-\alpha) + \alpha = 1,$$

то мы получили выпуклую комбинацию точек $x_1, \dots, x_n, y_1, \dots, y_m$, которая принадлежит $\operatorname{co} A$. □

Следствие 1. *Множество A является выпуклым в том и только в том случае, если $A = \operatorname{co} A$.*

Предложение 3. *Если A и B — выпуклые множества, то*

$$\operatorname{co}(A \cup B) = \{\, x \in X : x = (1-\alpha)a + \alpha b,\ a \in A,\ b \in B,\ \alpha \in [0,1] \,\}.$$

Доказательство. Из определения выпуклой оболочки вытекает, что $(1-\alpha)a + \alpha b \in \operatorname{co}(A \cup B)$ для всех $a \in A$, $b \in B$ и $\alpha \in [0,1]$. Пусть $x \in \operatorname{co}(A \cup B)$. Тогда

$$x = \sum_{j=1}^{n} \alpha_j x_j + \sum_{j=1}^{m} \beta_j y_j,$$

где $x_1, \dots, x_n \in A$, $y_1, \dots, y_m \in B$, $\alpha_1, \dots, \alpha_n \ge 0$, $\beta_1, \dots, \beta_m \ge 0$ и

$$\sum_{j=1}^{n} \alpha_j + \sum_{j=1}^{m} \beta_j = 1.$$

Положим

$$\alpha = \sum_{j=1}^{m} \beta_j.$$

Предположим, что $0 < \alpha < 1$. Тогда $x = (1-\alpha)a + \alpha b$, где

$$a = \sum_{j=1}^{n} \frac{\alpha_j}{1-\alpha} x_j, \quad b = \sum_{j=1}^{m} \frac{\beta_j}{\alpha} y_j.$$

В силу того, что
$$\sum_{j=1}^{n} \frac{\alpha_j}{1-\alpha} = \sum_{j=1}^{m} \frac{\beta_j}{\alpha} = 1,$$
$a \in A$, а $b \in B$. При $\alpha = 1$ или $\alpha = 0$ можно непосредственно убедиться, что справедливо то же представление. □

3. Аффинные подпространства

Подмножество $Y \subset X$ называется аффинным подпространством X, если для любых $x, y \in Y$ точка $(1-\alpha)x + \alpha y \in Y$ при всех $\alpha \in \mathbb{R}$.

Аффинной комбинацией точек $x_1, \dots, x_n$ называется точка

$$x = \sum_{j=1}^{n} \alpha_j x_j, \quad \sum_{j=1}^{n} \alpha_j = 1, \quad \alpha_1, \dots, \alpha_n \in \mathbb{R}. \tag{2}$$

Предложение 4. *Если Y — аффинное подпространство, то любая аффинная комбинация точек $x_1, \dots, x_n \in Y$ принадлежит Y.*

Доказательство. Доказательство проведем индукцией по числу точек. При $n = 1$ утверждение очевидно. Предположим, что утверждение верно для любой аффинной комбинации из не более, чем $n-1$ точек. Пусть имеется точка x вида (2) и $n > 1$. Среди $\alpha_1, \dots, \alpha_n$ найдется отличное от единицы. Без ограничения общности можно считать, что $\alpha_1 \neq 1$. Положим
$$\widetilde{\alpha}_j = \frac{\alpha_j}{1-\alpha_1}, \quad j = 2, \dots, n.$$
В силу того, что
$$\sum_{j=2}^{n} \widetilde{\alpha}_j = 1,$$
из предположения индукции вытекает, что
$$\widetilde{x} = \sum_{j=2}^{n} \widetilde{\alpha}_j x_j \in Y.$$

Тогда из того, что Y аффинное подпространство следует, что
$$x = \sum_{j=1}^{n} \alpha_j x_j = \alpha_1 x_1 + (1-\alpha_1)\widetilde{x} \in Y.$$
□

Множество всех аффинных комбинаций точек из множества A называется аффинной оболочкой $\operatorname{aff} A$ множества A
$$\operatorname{aff} A = \left\{ x \in X : x = \sum_{j=1}^{n} \alpha_j x_j,\ x_1, \dots, x_n \in A,\ \sum_{j=1}^{n} \alpha_j = 1,\ n \in \mathbb{N} \right\}.$$

Среди всех аффинных подпространств, содержащих данное множество $A \subset X$, существует наименьшее, т.е. то, которое содержится в любом аффинном подпространстве, содержащем A. Действительно, семейство всех аффинных подпространств, содержащих A, не пусто (в него входит, например, все пространство X). Пересечение всех множеств этого семейства будет аффинным подпространством (т.к. пересечение любого семейства аффинных подпространств — аффинное подпространство), содержащим A и наименьшим. Докажем, что это множество совпадает с $\operatorname{aff} A$.

Предложение 5. *Множество $\operatorname{aff} A$ является наименьшим аффинным подпространством, содержащим A.*

Доказательство. Из предложения 4 вытекает, что если некоторое аффинное подпространство содержит A, то оно содержит любую аффинную комбинацию точек из A, а значит, оно содержит $\operatorname{aff} A$. Следовательно, пересечение всех аффинных подпространств содержит $\operatorname{aff} A$. Осталось показать что $\operatorname{aff} A$ является аффинным подпространством, в этом случае оно совпадет с пересечением всех аффинных подпространств, содержащих A. Рассмотрим две произвольные точки $\operatorname{aff} A$

$$x = \sum_{j=1}^{n} \alpha_j x_j, \quad y = \sum_{j=1}^{m} \beta_j y_j.$$

Для любого $\alpha \in \mathbb{R}$ имеем

$$(1-\alpha)x + \alpha y = \sum_{j=1}^{n} (1-\alpha)\alpha_j x_j + \sum_{j=1}^{m} \alpha\beta_j y_j.$$

Так как

$$\sum_{j=1}^{n} (1-\alpha)\alpha_j + \sum_{j=1}^{m} \alpha\beta_j = (1-\alpha) + \alpha = 1,$$

то мы получили аффинную комбинацию точек $x_1, \dots, x_n, y_1, \dots, y_m$, которая принадлежит $\operatorname{aff} A$. □

Следствие 2. *Множество A является аффинным подпространством в том и только в том случае, если $A = \operatorname{aff} A$.*

Лекция 2

Предложение 6. *Каждое аффинное подпространство $Y \subset X$ представляется в виде $Y = a + L_Y$, где a — произвольная точка из Y, а L_Y — линейное подпространство из X, причем L_Y определено однозначно (не зависит от a).*

Доказательство. Пусть $a \in Y$. Положим

$$L_Y = \{\, x \in X : x = y - a, \quad y \in Y \,\}.$$

Докажем, что L_Y — линейное подпространство X. Пусть $x \in L_Y$. Тогда $x = y - a$, где $y \in Y$. Для любого $\lambda \in \mathbb{R}$ в силу того, что $\lambda y + (1-\lambda)a \in Y$, имеем

$$\lambda x = \lambda(y - a) = \lambda y + (1-\lambda)a - a \in L_Y.$$

Пусть теперь $x_1, x_2 \in L_Y$ и $x_1 = y_1 - a$, а $x_2 = y_2 - a$, где $y_1, y_2 \in Y$. Так как $(y_1 + y_2)/2 \in Y$, то

$$\frac{x_1 + x_2}{2} = \frac{y_1 + y_2}{2} - a \in L_Y.$$

Из доказанного выше, положив $\lambda = 2$, получаем, что $x_1 + x_2 \in L_Y$.

Положим

$$M_Y = \{\, x \in X : x = y - b, \quad y \in Y \,\},$$

где $b \in Y$. В силу того, что $b - a \in L_Y$, имеем

$$b + L_Y = a + (b - a) + L_Y = a + L_Y = Y.$$

Следовательно, $M_Y = L_Y$. □

Размерностью аффинного подпространства Y называется размерность соответствующего линейного подпространства L_Y:

$$\dim Y = \dim L_Y.$$

Размерностью произвольного множества A называется размерность его аффинной оболочки $\operatorname{aff} A$:

$$\dim A = \dim \operatorname{aff} A.$$

Поскольку $\operatorname{co} A \subset \operatorname{aff} A$, то

$$\dim A = \dim \operatorname{co} A = \dim \operatorname{aff} A.$$

4. Теорема Каратеодори

Теорема 1 (Каратеодори)**.** *Если* $\dim \operatorname{co} A = d$*, то любой элемент множества* $\operatorname{co} A$ *представляется в виде выпуклой комбинации не более, чем* $d+1$ *элемента множества* A*.*

Доказательство. Положим $Y = \operatorname{aff} A$. Тогда $\dim Y = d$. Пусть

$$x = \sum_{j=1}^{n} \alpha_j x_j, \quad \alpha_j \geq 0, \quad \sum_{j=1}^{n} \alpha_j = 1, \quad n \geq d+2, \quad x_j \in A.$$

Тогда элементы $x_2 - x_1, \ldots, x_n - x_1$ принадлежат линейному пространству L_Y (см. предложение 6), размерность которого d. Следовательно, они линейно зависимы. Тем самым существуют $\beta_2, \ldots, \beta_n$, не все равные нулю, для которых

$$\sum_{j=2}^{n} \beta_j (x_j - x_1) = 0.$$

Положим

$$\gamma_1 = -\sum_{j=2}^{n} \beta_j, \quad \gamma_j = \beta_j, \quad j = 2, \dots, n.$$

Тогда

$$\sum_{j=1}^{n} \gamma_j x_j = 0, \quad \sum_{j=1}^{n} \gamma_j = 0,$$

причем $\gamma_1, \dots, \gamma_n$ не все равны нулю. Следовательно, среди чисел $\gamma_1, \dots, \gamma_n$ существуют отрицательные числа. Пусть

$$a = \min \left\{ -\frac{\alpha_j}{\gamma_j} : \gamma_j < 0 \right\} = -\frac{\alpha_k}{\gamma_k}.$$

Тогда $\alpha_j + a\gamma_j \geq 0$ для всех $j = 1, \dots, n$. Кроме того,

$$x = \sum_{j=1}^{n} (\alpha_j + a\gamma_j) x_j, \quad \sum_{j=1}^{n} (\alpha_j + a\gamma_j) = 1.$$

Так как $\alpha_k + a\gamma_k = 0$, то x представляется в виде выпуклой комбинации из $n-1$ точки. Если $n-1 > d+1$, то продолжая этот процесс придем к тому, что x представляется в виде выпуклой комбинации не более, чем $d+1$ элемента множества A. □

Пусть X — линейное нормированное пространство.

Теорема 2. *Выпуклая оболочка компакта в конечномерном пространстве X является компактом.*

Доказательство. Пусть $\dim X = d$, $A \subset X$ — компакт. Докажем, что $\operatorname{co} A$ — компакт. Для этого достаточно доказать, что из любой последовательности $\{x_k\} \subset \operatorname{co} A$ можно выбрать подпоследовательность, сходящуюся к некоторой точке из $\operatorname{co} A$. По теореме Каратеодори

$$x_k = \sum_{j=1}^{d+1} \alpha_{kj} x_{kj}, \quad x_{kj} \in A, \quad \alpha_{kj} \geq 0, \quad \sum_{j=1}^{d+1} \alpha_{kj} = 1.$$

Так как A — компакт, то в последовательности $\{x_{k1}\}$ существует подпоследовательность, сходящаяся к некоторой точке $y_1 \in A$, а в этой подпоследовательности в $\{\alpha_{k1}\}$ существует подпоследовательность, сходящаяся к некоторому числу $\alpha_1 \in [0, 1]$. Не ограничивая общности, можно считать, что $x_{k1} \to y_1$ и $\alpha_{k1} \to \alpha_1$ при $k \to \infty$. Аналогично, переходя к подпоследовательностям, можно считать, что $x_{kj} \to y_j$ и $\alpha_{kj} \to \alpha_j$ при $k \to \infty$ для любого $j = 2, \dots, d+1$. При этом

$$x_k \to \sum_{j=1}^{d+1} \alpha_j y_j, \quad \alpha_j \geq 0, \quad \sum_{j=1}^{d+1} \alpha_j = 1.$$

Это означает, что предельная точка принадлежит $\operatorname{co} A$. □

Конической комбинацией точек $x_1, \dots, x_n$ называется точка

$$x = \sum_{j=1}^{n} \alpha_j x_j, \quad \alpha_1, \dots, \alpha_n \geq 0.$$

Множество всех конических комбинаций точек из множества A называется конической оболочкой $\operatorname{cone} A$ множества A

$$\operatorname{cone} A = \left\{ x \in X : x = \sum_{j=1}^{n} \alpha_j x_j,\ x_1, \dots, x_n \in A,\ \alpha_j \geq 0,\ n \in \mathbb{N} \right\}.$$

Напомним, что линейной оболочкой множества A называется множество всех линейных комбинаций элементов из множества A:

$$\operatorname{lin} A = \left\{ x \in X : x = \sum_{j=1}^{n} \alpha_j x_j,\ x_1, \dots, x_n \in A,\ \alpha_j \in \mathbb{R},\ n \in \mathbb{N} \right\}.$$

Теорема 3 (Каратеодори для конической оболочки)**.** *Если* $\dim \operatorname{lin} A = d$, *то любой элемент множества* $\operatorname{cone} A$ *представляется в виде конической комбинации не более, чем* d *линейно независимых элементов множества* A.

Доказательство. Пусть

$$x = \sum_{j=1}^{n} \alpha_j x_j, \quad \alpha_1, \dots, \alpha_n \geq 0.$$

Если $x_1, \dots, x_n$ линейно независимы, то $n \leq d$ и все доказано. Если элементы $x_1, \dots, x_n$ линейно зависимы, то найдутся не все равные нулю $\gamma_1, \dots, \gamma_n$ такие, что

$$\sum_{j=1}^{n} \gamma_j x_j = 0.$$

Без ограничения общности можно считать, что среди $\gamma_1, \dots, \gamma_n$ есть отрицательные числа (если они все положительные, то можно рассматривать $-\gamma_1, \dots, -\gamma_n$). Пусть

$$a = \min \left\{ -\frac{\alpha_j}{\gamma_j} : \gamma_j < 0 \right\} = -\frac{\alpha_k}{\gamma_k}.$$

Тогда $\alpha_j + a\gamma_j \geq 0$ для всех $j = 1, \dots, n$ и

$$x = \sum_{j=1}^{n} (\alpha_j + a\gamma_j) x_j.$$

Так как $\alpha_k + a\gamma_k = 0$, то x представляется в виде конической комбинации из $n - 1$ точки. Если эти элементы линейно зависимы, то продолжая этот процесс придем к тому, что x представляется в виде конической комбинации не более, чем d линейно независимых элементов множества A. $\square$

5. Теоремы Радона и Хелли

Теорема 4 (Радона). *Пусть X — линейное пространство и $\dim X = d$. Тогда любое конечное множество из X, содержащее не менее $d+2$ точек, может быть разбито на два непересекающихся множества, выпуклые оболочки которых пересекаются.*

Доказательство. Пусть даны точки $x_1, \dots, x_n$, $n \ge d+2$. Векторы $x_2 - x_1, \dots, x_n - x_1$ — линейно зависимы, т.к. их не менее $d+1$. Следовательно, существуют $\alpha_2, \dots, \alpha_n$, не все равные нулю, такие, что

$$\sum_{j=2}^{n} \alpha_j (x_j - x_1) = 0.$$

Положим

$$\beta_1 = -\sum_{j=2}^{n} \alpha_j, \quad \beta_j = \alpha_j, \quad j = 2, \dots, n.$$

Имеем

$$\sum_{j=1}^{n} \beta_j x_j = 0, \quad \sum_{j=1}^{n} \beta_j = 0,$$

причем $\beta_1, \dots, \beta_n$ не все равны нулю. Поменяем нумерацию и будем считать, что $\beta_1, \dots, \beta_m \ge 0$, а $\beta_{m+1}, \dots, \beta_n < 0$. Положим

$$\beta = \sum_{j=1}^{m} \beta_j = -\sum_{j=m+1}^{n} \beta_j, \quad A = \{x_1, \dots, x_m\}, \quad B = \{x_{m+1}, \dots, x_n\}.$$

Тогда

$$\sum_{j=1}^{m} \gamma_j x_j = \sum_{j=m+1}^{n} \gamma_j x_j, \tag{3}$$

где $\gamma_j = \beta_j/\beta$, $j = 1, \dots, m$, и $\gamma_j = -\beta_j/\beta$, $j = m+1, \dots, n$. Причем

$$\sum_{j=1}^{m} \gamma_j = \sum_{j=m+1}^{n} \gamma_j = 1.$$

Таким образом, выпуклая комбинация в левой части равенства (3) лежит в $\operatorname{co} A$, а выпуклая комбинация в правой части равенства (3) лежит в $\operatorname{co} B$. Тем самым $\operatorname{co} A \cap \operatorname{co} B \ne \emptyset$. $\square$

Лекция 3

Теорема 5 (Хелли).

1. *Если $\dim X = d$ и в X имеются $n \ge d+1$ выпуклых множеств, любые $d+1$ из которых имеют общую точку, то и все эти множества имеют общую точку.*

2. *Пусть X — линейное нормированное пространство, $\dim X = d$, $\mathcal{I}$ — некоторое бесконечное множество индексов и $\{A_\alpha\}_{\alpha \in \mathcal{I}}$ —*

семейство замкнутых выпуклых подмножеств X, по крайней мере одно из которых компактно. Тогда если любое подсемейство из $d+1$ множеств имеет общую точку, то и все семейство имеет общую точку.

Доказательство. 1. Будем доказывать утверждение индукцией по числу множеств n. При $n = d+1$ утверждение очевидно. Пусть $n \geq d+2$ и теорема доказана для любых $n-1$ выпуклых множеств. Обозначим данные множества через A_j, $j = 1, \dots, n$. Для каждого $k = 1, \dots, n$ возьмем произвольную точку

$$x_k \in \bigcap_{\substack{j=1 \\ j \neq k}}^{n} A_j$$

(по предположению индукции эти пересечения не пусты). Так как $n \geq d+2$, то к точкам $x_1, \dots, x_n$ применима теорема Радона. С возможной перенумерацией точек получаем, что существует натуральное число $m \leq n-1$, для которого множества $\operatorname{co}\{x_1, \dots, x_m\}$ и $\operatorname{co}\{x_{m+1}, \dots, x_n\}$ имеют общую точку x. Точка x_j при всех $j = 1, \dots, m$ принадлежит каждому из множеств $A_{m+1}, \dots, A_n$. В силу выпуклости этих множеств каждое из них содержит $\operatorname{co}\{x_1, \dots, x_m\}$. Следовательно, каждое из этих множеств содержит точку x. Аналогично доказывается, что каждое из множеств $A_1, \dots, A_m$ содержит точку x. Тем самым доказано, что у всех множеств $A_1, \dots, A_n$ имеется общая точка.

2. Пусть A_{α_0} — компактное множество. Из 1 следует, что любое конечное семейство из множества замкнутых подмножеств $B_\alpha = A_\alpha \cap A_{\alpha_0}$, $\alpha \in \mathcal{I}$, содержащихся в A_{α_0} имеет общую точку. Предположим, что все семейство не имеет общей точки. Тогда множества $\{X \setminus B_\alpha\}_{\alpha \in \mathcal{I}}$ — открытые и образуют покрытие компактного множества A_{α_0}. В силу компактности из этого покрытия можно выбрать конечное покрытие $\{X \setminus B_\alpha\}_{\alpha \in \mathcal{I}_n}$. Но тогда конечное семейство $\{B_\alpha\}_{\alpha \in \mathcal{I}_n}$ не может иметь общей точки. Полученное противоречие доказывает утверждение 2. □

6. Теоремы отделимости

Пусть A и B — непустые подмножества нормированного пространства X. Говорят, что ненулевой функционал $x^* \in X^*$ *отделяет множества A и B*, если

$$\sup_{x \in A} \langle x^*, x \rangle \leq \inf_{x \in B} \langle x^*, x \rangle.$$

Если неравенство строгое, то говорят, что x^* *строго отделяет A и B*.

Пусть число $\gamma \in \mathbb{R}$ таково, что

$$\sup_{x \in A} \langle x^*, x \rangle \leq \gamma \leq \inf_{x \in B} \langle x^*, x \rangle.$$

Тогда, геометрически, отделимость множеств A и B означает, что они расположены по разные стороны от гиперплоскости

$$\{\, x \in X : \langle x^*, x\rangle = \gamma \,\}.$$

Пусть X — нормированное пространство, $\widehat{x} \in X$ и $r > 0$. Положим

$$B_r(\widehat{x}) = \{\, x \in X : \|x - \widehat{x}\|_X < r \,\}.$$

Если A — некоторое множество из X, то точка $\widehat{x} \in A$ называется внутренней точкой A, если существует такое $\varepsilon > 0$, что $B_\varepsilon(\widehat{x}) \subset A$. Множество внутренних точек A обозначается через $\operatorname{int} A$.

Напомним формулировку первой теоремы отделимости (см. [4, стр. 243]).

Теорема 6 (Первая теорема отделимости)**.** *Пусть A и B — непустые выпуклые подмножества нормированного пространства X, причем $\operatorname{int} A \ne \emptyset$ и $B \cap \operatorname{int} A = \emptyset$. Тогда множества A и B отделимы.*

Отсюда следует

Теорема 7 (Вторая теорема отделимости)**.** *Пусть A — непустое замкнутое выпуклое подмножество нормированного пространства X и $\widehat{x} \notin A$. Тогда множества A и $\widehat{x}$ строго отделимы.*

Доказательство. Так как A замкнуто, то дополнение к A открыто и поэтому существует такое $r > 0$, что открытый шар $B_X(\widehat{x}, r)$ не пересекается с A. Тогда по первой теореме отделимости существует ненулевой функционал $x^* \in X^*$ такой, что

$$\sup_{x \in A} \langle x^*, x\rangle \le \inf_{x \in B_X(\widehat{x}, r)} \langle x^*, x\rangle.$$

Но

$$\inf_{x \in B_X(\widehat{x}, r)} \langle x^*, x\rangle < \langle x^*, \widehat{x}\rangle,$$

так как ненулевой линейный непрерывный функционал не может достигать точной нижней грани во внутренней точке. Следовательно, множества A и $\widehat{x}$ строго отделимы. □

7. Вторая теорема отделимости в конечномерном случае

В этом разделе мы сформулируем и докажем теоремы отделимости в конечномерном случае. Начнем со второй теоремы отделимости.

Теорема 8. *Пусть A — непустое замкнутое выпуклое подмножество $\mathbb{R}^d$ и $\widehat{x} \notin A$. Тогда множества A и $\widehat{x}$ строго отделимы.*

Доказательство. Пусть $x_0 \in A$. Положим

$$A_0 = \{\, x \in A : |x - \widehat{x}| \le |x_0 - \widehat{x}| \,\}.$$

Функция $f(x) = |x - \widehat{x}|$ является непрерывной на ограниченном замкнутом множестве A_0 и, следовательно, по теореме Вейерштрасса достигает своего минимума в некоторой точке $\widehat{a}$. Тем самым $\widehat{a}$ — ближайшая точка к $\widehat{x}$ из множества A_0, а значит, и из множества A. Положим $x' = (\widehat{a} - \widehat{x})^T$ и рассмотрим гиперплоскость $x' \cdot x = x' \cdot \widehat{a}$. Докажем, что эта гиперплоскость отделяет $\widehat{x}$ от A. Имеем

$$x' \cdot \widehat{x} = x' \cdot (\widehat{x} - \widehat{a} + \widehat{a}) = -|x'|^2 + x' \cdot \widehat{a} < x' \cdot \widehat{a}.$$

Остается доказать, что $x' \cdot x \ge x' \cdot \widehat{a}$ для всех $x \in A$. Предположим, что нашлась точка $a_0 \in A$, для которой $x' \cdot a_0 < x' \cdot \widehat{a}$. Так как A — выпуклое множество, $(1-t)\widehat{a} + ta_0 \in A$ при всех $t \in [0,1]$. Имеем

$$|(1-t)\widehat{a} + ta_0 - \widehat{x}|^2 = |(x')^T + t(a_0 - \widehat{a})|^2 = |x'|^2 + 2t\alpha + t^2|a_0 - \widehat{a}|^2,$$

где $\alpha = x' \cdot (a_0 - \widehat{a}) < 0$. Поэтому при достаточно малых t

$$|(1-t)\widehat{a} + ta_0 - \widehat{x}|^2 < |x'|^2 = |\widehat{a} - \widehat{x}|^2,$$

что противоречит тому, что $\widehat{a}$ — ближайшая точка к $\widehat{x}$ из точек множества A. $\square$

8. Аффинная независимость. Симплексы

Перед доказательством первой теоремы отделимости для конечномерного случая нам потребуется некоторые вспомогательные утверждения. Пусть X — линейное пространство. Векторы $x_1, \dots, x_{k+1} \in X$ называются аффинно независимыми, если из того, что

$$\sum_{j=1}^{k+1} \lambda_j x_j = 0 \quad \text{и} \quad \sum_{j=1}^{k+1} \lambda_j = 0, \tag{4}$$

следует, что $\lambda_1 = \ldots = \lambda_{k+1} = 0$.

Предложение 7. *Векторы $x_1, \dots, x_{k+1}$ аффинно независимы в том и только в том случае, если векторы $x_j - x_1$, $2 \le j \le k+1$, — линейно независимы.*

Доказательство. Пусть $x_1, \dots, x_{k+1}$ аффинно независимы и

$$\sum_{j=2}^{k+1} \lambda_j (x_j - x_1) = 0. \tag{5}$$

Тогда

$$\sum_{j=2}^{k+1} \lambda_j x_j - \left(\sum_{j=2}^{k+1} \lambda_j \right) x_1 = 0. \tag{6}$$

Из аффинной независимости $x_1, \dots, x_{k+1}$ вытекает, что $\lambda_2 = \dots = \lambda_{k+1} = 0$.

Пусть теперь векторы $x_j - x_1$, $2 \le j \le k+1$, — линейно независимы и выполняются условия (4). Тогда имеет место равенство (6), которое эквивалентно равенству (5). Из линейной независимости векторов $x_j - x_1$, $2 \le j \le k+1$, получаем, что $\lambda_2 = \dots = \lambda_{k+1} = 0$. Кроме того,

$$\lambda_1 = -\sum_{j=2}^{k+1} \lambda_j = 0.$$

□

Выпуклая оболочка аффинно независимых векторов $x_1, \dots, x_{k+1}$ называется *k-мерным симплексом*, а векторы $x_1, \dots, x_{k+1}$ — *вершинами симплекса*. Любой вектор из этого симплекса единственным образом представим в виде

$$x = \sum_{j=1}^{k+1} \lambda_j x_j, \quad \sum_{j=1}^{k+1} \lambda_j = 1, \quad \lambda_j \ge 0,\ j = 1, \dots, k+1.$$

Числа $\lambda_1, \dots, \lambda_{k+1}$ называются барицентрическими координатами вектора x.

Одномерные симплексы — это отрезки, двумерные — треугольники, трехмерные — тетраэдры.

Лекция 4

Предложение 8. *Пусть $A \subset X$ и $0 < \dim A < \infty$. Тогда максимальное число аффинно независимых векторов в множестве A равно $\dim A + 1$.*

Доказательство. Предположим, что $x_1, \dots, x_{k+1} \in A$, являются аффинно независимыми и максимальное число аффинно независимых векторов в A равно $k+1$. Если предположить, что существует вектор $x_0 \in A$ такой, что $x_0 \notin \operatorname{aff}\{x_1, \dots, x_{k+1}\}$, то векторы $x_0, x_1, \dots, x_{k+1}$ будут аффинно независимыми, что противоречит максимальности числа аффинно независимых векторов в A. Таким образом, $A \subset \operatorname{aff}\{x_1, \dots, x_{k+1}\}$. Следовательно,

$$\operatorname{aff} A = \operatorname{aff}\{x_1, \dots, x_{k+1}\}.$$

Поэтому

$$k = \dim \operatorname{aff}\{x_1, \dots, x_{k+1}\} = \dim \operatorname{aff} A = \dim A.$$

□

Следствие 3. *Если A — выпуклое множество, $\dim A = d$, $0 < d < \infty$, то A содержит d-мерный симплекс.*

Предложение 9. *Пусть $A \subset \mathbb{R}^d$ — непустое выпуклое множество и $\dim A = d$. Тогда $\operatorname{int} A \ne \emptyset$.*

Доказательство. Из следствия 3 вытекает, что A содержит d-мерный симплекс. Пусть $a_1, \dots, a_{d+1}$ — его вершины. Без ограничения общности можно считать, что $a_1 = 0$. Покажем, что любая точка этого симплекса с положительными барицентрическами координатами является внутренней точкой симплекса, а следовательно, и множества A. Пусть

$$x_0 = \sum_{k=1}^{d+1} \lambda_k a_k, \quad \sum_{k=1}^{d+1} \lambda_k = 1, \quad \lambda_k > 0, \quad k = 1, \dots, d+1.$$

Пусть $e_1, \dots, e_d$ — стандартный базис в $\mathbb{R}^d$. Так как $a_2, \dots, a_{d+1}$ — линейно независимы, векторы стандартного базиса можно представить в виде

$$e_j = \sum_{k=2}^{d+1} \alpha_{jk} a_k, \quad j = 1, \dots, d.$$

Выберем $\varepsilon > 0$ достаточно малым так, чтобы

$$\lambda_k - \varepsilon \sum_{j=1}^{d} |\alpha_{jk}| > 0, \quad k = 2, \dots, d+1,$$

и

$$\sum_{k=2}^{d+1} \left(\lambda_k + \varepsilon \sum_{j=1}^{d} |\alpha_{jk}| \right) < 1.$$

Пусть x — произвольный вектор из $B_\varepsilon(0)$. Тогда его можно записать в виде

$$x = \sum_{j=1}^{d} x_j e_j = \sum_{j=1}^{d} x_j \sum_{k=2}^{d+1} \alpha_{jk} a_k = \sum_{k=2}^{d+1} \sum_{j=1}^{d} x_j \alpha_{jk} a_k,$$

где $|x_j| < \varepsilon$. Таким образом,

$$x_0 + x = \sum_{k=2}^{d+1} \left(\lambda_k + \sum_{j=1}^{d} x_j \alpha_{jk} \right) a_k = \sum_{k=2}^{d+1} \gamma_k a_k.$$

В силу выбора ε, имеем

$$\gamma_k > 0, \quad k = 2, \dots, d+1, \quad \sum_{k=2}^{d+1} \gamma_k < 1.$$

Тем самым точка x_0 является внутренней точкой симплекса. $\square$

9. Относительная внутренность. Первая теорема отделимости в конечномерном случае

Если рассмотреть отрезок на плоскости, то все его точки не являются внутренним, хотя ясно, что точки, отличные от концевых, обладают некоторыми свойствами, присущими внутренним точкам. Это наблюдение приводит к следующему понятию.

Пусть X — линейное нормированное пространство и $A \subset X$. Точка $x \in A$ называется относительно внутренней точкой множества A, если существует $\varepsilon > 0$ такое, что $B_\varepsilon(x) \cap \operatorname{aff} A \subset A$. Множество всех относительно внутренних точек A называется относительной внутренностью A и обозначается $\operatorname{ri} A$.

Предложение 10. *Пусть $A \subset \mathbb{R}^d$ — непустое выпуклое множество. Тогда*

a) $\operatorname{ri} A \neq \emptyset$*;*
b) *если $x_1 \in \operatorname{ri} A$, а $x_2 \in \operatorname{cl} A$, то все точки интервала (x_1, x_2) принадлежат $\operatorname{ri} A$;*
c) $\operatorname{ri} A$ *и* $\operatorname{cl} A$ *— выпуклые множества и* $\operatorname{cl} \operatorname{ri} A = \operatorname{cl} A$.

Доказательство. a). Предположим сначала, что $\dim A = d$. Тогда утверждение непосредственно вытекает из предложения 9. Если $\dim A = d_1 < d$, можно вместо аффинного подпространства $\operatorname{aff} A$ рассматривать $\mathbb{R}^{d_1}$. Те же аргументы, которые приводились при доказательстве предложения 9, показывают, что A содержит внутренние точки относительно пространства $\mathbb{R}^{d_1}$. Это означает, что $\operatorname{ri} A \neq \emptyset$.

b). В силу сделанного выше замечания будем считать, что $\operatorname{aff} A = \mathbb{R}^d$. Тогда $\operatorname{ri} A = \operatorname{int} A$. Пусть $y \in (x_1, x_2)$. Тогда $y = (1-\alpha)x_1 + \alpha x_2$, $0 < \alpha < 1$. Так как $x_2 \in \operatorname{cl} A$, то $x_2 \in A + B_\varepsilon(0)$ для любого $\varepsilon > 0$. Тогда

$$y + B_\varepsilon(0) = (1-\alpha)x_1 + \alpha x_2 + B_\varepsilon(0) \subset (1-\alpha)x_1 + \alpha(A + B_\varepsilon(0)) + B_\varepsilon(0) = (1-\alpha)(x_1 + B_\gamma(0)) + \alpha A,$$

где $\gamma = \varepsilon(1+\alpha)/(1-\alpha)$. В силу того, что $x_1 \in \operatorname{int} A$ для достаточно малых $\varepsilon > 0$ $\; x_1 + B_\gamma(0) \subset A$. Поэтому

$$y + B_\varepsilon(0) \subset (1-\alpha)A + \alpha A = A.$$

c). Будем снова предполагать, что $\operatorname{aff} A = \mathbb{R}^d$. Выпуклость $\operatorname{ri} A$ непосредственно следует из b). Докажем выпуклость $\operatorname{cl} A$. Пусть $x_1, x_2 \in \operatorname{cl} A$, а $x = (1-\alpha)x_1 + \alpha x_2$. Для любого $\varepsilon > 0$

$$(x_j + B_\varepsilon(0)) \cap A \neq \emptyset, \quad j = 1, 2.$$

Пусть $\widehat{x}_j \in (x_j + B_\varepsilon(0)) \cap A$, $j = 1, 2$. Положим $\widehat{x} = (1-\alpha)\widehat{x}_1 + \alpha\widehat{x}_2$. Тогда $\widehat{x} \in A$. Кроме того,

$$\widehat{x} \in (1-\alpha)(x_1 + B_\varepsilon(0)) + \alpha(x_1 + B_\varepsilon(0)) = x + B_\varepsilon(0).$$

Тем самым в любой окрестности точки x есть точки из A. Следовательно, $x \in clA$.

Включение $\operatorname{cl} \operatorname{ri} A \subset \operatorname{cl} A$ очевидно. Пусть $x \in \operatorname{cl} A$. Из a) следует, что существует $x_1 \in \operatorname{ri} A$. Из b) вытекает, что $(x_1, x) \in \operatorname{ri} A$. Следовательно, $x \in \operatorname{cl} \operatorname{ri} A$. $\square$

Теорема 9. *Пусть* $A, B \subset \mathbb{R}^d$ *— непустые выпуклые множества и* $\operatorname{ri} A \cap B = \emptyset$. *Тогда* A *и* B *отделимы.*

Доказательство. Положим $C = \operatorname{ri} A - B$. Так как $\operatorname{ri} A$ — выпуклое множество (предложение 10, c)), то C — выпуклое множество и $0 \notin C$. Предположим $0 \notin \operatorname{cl} C$. Из предложения 10 (c)) вытекает, что $\operatorname{cl} C$ тоже выпуклое множество. Тогда по теореме 8 существует вектор $x' \in (\mathbb{R}^d)'$ такой, что $x' \cdot x < 0$ для всех $x \in \operatorname{cl} C$.

Пусть теперь $0 \in \operatorname{cl} C$. Возьмем произвольный вектор $x \in \operatorname{ri} C$ ($\operatorname{ri} C \neq \emptyset$ в силу предложения 10, a)). Тогда для любого $\lambda > 0$ вектор $-\lambda x \notin \operatorname{cl} C$. Действительно, в противном случае из предложения 10 (b)) вытекало бы, что

$$0 = \frac{\lambda}{\lambda+1} x + \frac{1}{\lambda+1}(-\lambda x) \in \operatorname{ri} C \subset C.$$

Тем самым существует последовательность векторов $x_k \notin \operatorname{cl} C$ такая, что $x_k \to 0$ при $k \to \infty$. По второй теореме отделимости эти векторы можно строго отделить от $\operatorname{cl} C$. Следовательно, существуют $x'_k \in (\mathbb{R}^d)'$, $x'_k \neq 0$, $k \in \mathbb{N}$, такие, что

$$x'_k \cdot x < x'_k \cdot x_k \tag{7}$$

для всех $x \in \operatorname{cl} C$ и всех $k \in \mathbb{N}$. Разделив (7) на $|x'_k|$, можно считать $|x'_k| = 1$. Из того, что сфера в конечномерном пространстве ограничена и замкнута, вытекает существование подпоследовательности в $\{x'_k\}$, сходящейся к некоторому вектору x', $|x'| = 1$. Переходя к пределу по этой подпоследовательности, получим из (7), что $x' \cdot x \le 0$ для всех $x \in \operatorname{cl} C$.

Итак, в обоих случаях доказано существование $x' \in (\mathbb{R}^d)'$ такого, что $x' \cdot x \le 0$ для всех $x \in \operatorname{cl} C$, а значит, и для всех $x = a - b$, где $a \in \operatorname{ri} A$, а $b \in B$. Таким образом,

$$\sup_{a \in \operatorname{ri} A} x' \cdot a \le \inf_{b \in B} x' \cdot b. \tag{8}$$

Неравенство (8) останется справедливым, если в левой части заменить $\operatorname{ri} A$ на $\operatorname{cl} \operatorname{ri} A$. А поскольку $A \subset \operatorname{cl} A = \operatorname{cl} \operatorname{ri} A$ (см. предложение 10, c)), то (8) верно и для A, что означает отделимость множеств A и B. □

Лекция 5

10. Выпуклые функции

Мы будем рассматривать функции, которые принимают не только конечные значения. С этой целью вводится понятие расширенной прямой $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$. Арифметические операции и неравенства для элементов расширенной прямой понимаются следующим образом: $-\infty < a < +\infty$, $a \in \mathbb{R}$, $a + (\pm\infty) = \pm\infty$ для всех

$a \in \mathbb{R}$,

$$a(\pm\infty) = \begin{cases} \pm\infty, & a > 0, \\ 0, & a = 0, \\ \mp\infty, & a < 0, \end{cases},$$

$+\infty + (+\infty) = +\infty$, $-\infty + (-\infty) = -\infty$.

Пусть X — линейное пространство и $f\colon X \to \overline{\mathbb{R}}$. Множества

$$\operatorname{dom} f = \{\, x \in X : f(x) < +\infty \,\},$$
$$\operatorname{epi} f = \{\, (x, \alpha) \in X \times \mathbb{R} : \alpha \geq f(x),\ x \in \operatorname{dom} f \,\}$$

называются соответственно эффективным множеством и надграфиком (или эпиграфом) функции f. Функцию f называют собственной, если $\operatorname{dom} f \neq \emptyset$ и $f(x) > -\infty$ при всех $x \in X$.

Функция $f\colon X \to \overline{\mathbb{R}}$ называется выпуклой, если ее надграфик выпуклое множество в $X \times \mathbb{R}$. Нетрудно проверить, что функция f выпукла тогда и только тогда, когда для любых $x_1, x_2 \in \operatorname{dom} f$ и любого $0 \leq \alpha \leq 1$ выполняется неравенство

$$f((1-\alpha)x_1 + \alpha x_2) \leq (1-\alpha) f(x_1) + \alpha f(x_2),$$

которое называется неравенством Йенссена.

Пусть $X = \mathbb{R}^d$ и $f\colon \mathbb{R}^d \to \mathbb{R}$. Через $f''(x)$ обозначим гессиан f в точке x

$$f''(x) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1^2}(x) & \dfrac{\partial^2 f}{\partial x_1 \partial x_2}(x) & \dots & \dfrac{\partial^2 f}{\partial x_1 \partial x_d}(x) \\ \dfrac{\partial^2 f}{\partial x_2 \partial x_1}(x) & \dfrac{\partial^2 f}{\partial x_2^2}(x) & \dots & \dfrac{\partial^2 f}{\partial x_2 \partial x_d}(x) \\ \dots & \dots & \dots & \dots \\ \dfrac{\partial^2 f}{\partial x_d \partial x_1}(x) & \dfrac{\partial^2 f}{\partial x_d \partial x_2}(x) & \dots & \dfrac{\partial^2 f}{\partial x_d^2}(x) \end{pmatrix}.$$

Теорема 10. *Если функция $f\colon \mathbb{R}^d \to \mathbb{R}$ дважды дифференцируема на $\mathbb{R}^d$, то она является выпуклой в том, и только в том случае, если ее гессиан в любой точке $x \in \mathbb{R}^d$ удовлетворяет условию*

$$h^T f''(x) h \geq 0$$

для всех $h \in \mathbb{R}^d$ (матрица гессиана в любой точке является неотрицательно определенной).

Доказательство. Пусть f — выпуклая функция. Предположим, что существует $x \in \mathbb{R}^d$ и $h \in \mathbb{R}^d$ такие, что $h^T f''(x) h < 0$. По формуле Тейлора для $t \in \mathbb{R}$ имеем

$$f(x + th) = f(x) + f'(x)ht + h^T f''(x) h \frac{t^2}{2} + o(t^2),$$
$$f(x - th) = f(x) - f'(x)ht + h^T f''(x) h \frac{t^2}{2} + o(t^2).$$

Отсюда, складывая эти равенства, получаем

$$f(x+th) - 2f(x) + f(x-th) = h^T f''(x)ht^2 + o(t^2).$$

Следовательно, при достаточно малых t

$$f(x+th) - 2f(x) + f(x-th) < 0,$$

т.е.

$$f\left(\frac{x-th}{2} + \frac{x+th}{2}\right) > \frac{1}{2}f(x-th) + \frac{1}{2}f(x+th),$$

что противоречит выпуклости f.

Пусть теперь $h^T f''(x)h \geq 0$ для всех $x \in \mathbb{R}^d$ и всех $h \in \mathbb{R}^d$. Для произвольных $x_1, x_2 \in \mathbb{R}^d$ рассмотрим функцию

$$F(t) = f(x_1 + t(x_2 - x_1)) - f(x_1) - t(f(x_2) - f(x_1)).$$

Имеем $F(0) = F(1) = 0$,

$$F''(t) = (x_2 - x_1)^T f''(x_1 + t(x_2 - x_1))(x_2 - x_1) \geq 0.$$

Предположим, что при некотором $t \in (0,1)$ $\ F(t) > 0$. Тогда найдется точка $t_0 \in (0,1)$, в которой функция F будет достигать максимального значения и, значит, $F'(t_0) = 0$. Поскольку $F''(t) \geq 0$ при всех $t \in [t_0, 1]$, то $F'(t) \geq 0$ при $t \in [t_0, 1]$. Тем самым функция F не убывает на отрезке $[t_0, 1]$, а значит, $F(1) > 0$. Полученное противоречие доказывает, что $F(t) \leq 0$ для всех $t \in [0,1]$. Таким образом, для всех $t \in [0,1]$ справедливо неравенство

$$f((1-t)x_1 + tx_2) \leq (1-t)f(x_1) + tf(x_2),$$

что и означает выпуклость функции f. □

11. Теорема Каруша–Куна–Таккера

Пусть $f_j\colon A \to \mathbb{R}$, $j = 0, 1, \ldots, m$, — функции, определенные на некотором множестве A. Рассмотрим экстремальную задачу

$$(9) \qquad f_0(x) \to \min, \quad f_j(x) \leq 0, \quad j = 1, \ldots, m, \quad x \in A.$$

Свяжем с задачей (9) следующую функцию Лагранжа

$$\mathcal{L}(x, \overline{\lambda}) = \sum_{j=0}^{m} \lambda_j f_j(x),$$

где $\overline{\lambda} = (\lambda_0, \lambda_1, \ldots, \lambda_m)$ — набор множителей Лагранжа. Всякий элемент, для которого выполнены условия экстремальной задачи ($f_j(x) \leq 0, \quad j = 1, \ldots, m, \quad x \in A$), будем называть допустимым.

Лемма 1. *Пусть $\widehat{x}$ — допустимый элемент в задаче* (9) *и существует набор множителей Лагранжа $\overline{\lambda} = (\lambda_0, \lambda_1, \ldots, \lambda_m)$ таких, что*

(*a*) $\min_{x \in A} \mathcal{L}(x, \overline{\lambda}) = \mathcal{L}(\widehat{x}, \overline{\lambda})$ *(условие минимума);*
(*b*) $\lambda_j \geq 0$, $j = 0, 1, \ldots, m$ *(условие неотрицательности);*

(c) $\lambda_j f_j(\widehat{x}) = 0$, $j = 1, \dots, m$ *(условие дополняющей нежесткости).*

Тогда, если $\lambda_0 > 0$, *то* $\widehat{x}$ *— экстремальный элемент в задаче* (9).

Доказательство. Для любого допустимого в задаче (9) элемента $x \in A$ имеем
$$\lambda_0 f_0(x) \ge \mathcal{L}(x, \overline{\lambda}) \ge \mathcal{L}(\widehat{x}, \overline{\lambda}) = \lambda_0 f_0(\widehat{x}).$$
Деля на λ_0, получаем требуемое. □

Пусть X — вещественное линейное пространство, $f_j \colon X \to \mathbb{R}$, $j = 0, 1, \dots, m$, — выпуклые функции и A — непустое выпуклое подмножество X. Задачу

$$f_0(x) \to \min, \quad f_j(x) \le 0, \quad j = 1, \dots, m, \quad x \in A, \tag{10}$$

называют выпуклой задачей или задачей выпуклого программирования.

Имеет место следующая теорема.

Теорема 11 (Каруша–Куна–Таккера)**.**

1. *Если* $\widehat{x}$ *— минимум в задаче* (10), *то найдется такой ненулевой набор множителей Лагранжа* $\overline{\lambda} = (\lambda_0, \lambda_1, \dots, \lambda_m)$, *что выполнены условия* (a), (b) *и* (c) *из леммы* 1.

2. *Если существует допустимая в* (10) *точка* $\widehat{x}$ *и набор множителей Лагранжа* $\overline{\lambda} = (\lambda_0, \lambda_1, \dots, \lambda_m)$, *удовлетворяющие условиям* (a), (b) *и* (c) *и при этом* $\lambda_0 > 0$, *то* $\widehat{x}$ *— решение задачи* (10).

3. *Если при выполнении условий* 1 *найдется точка* $\overline{x} \in A$ *такая, что* $f_j(\overline{x}) < 0$, $1 \le j \le m$, *то* $\lambda_0 \ne 0$ *(условие Слейтера).*

Доказательство. 1. Пусть $\widehat{x}$ — решение задачи (10). Рассмотрим множество
$$M = \{\, \mu = (\mu_0, \mu_1, \dots, \mu_m)^T \in \mathbb{R}^{m+1} : \exists\, x \in A : f_0(x) - f_0(\widehat{x}) < \mu_0,\\ f_j(x) \le \mu_j,\ j = 1, \dots, m \,\}.$$
Непосредственная проверка показывает, что это множество выпукло. Кроме того, легко видеть, что оно содержит все векторы с положительными компонентами (надо взять $x = \widehat{x}$) и тем самым его внутренность не пуста. Наконец, $0 \notin M$, так как в противном случае нашелся бы элемент $\overline{x} \in A$ такой, что $f_j(\overline{x}) \le 0$, $j = 1, \dots, m$, и $f_0(\overline{x}) - f_0(\widehat{x}) < 0$, в противоречие с тем, что $\widehat{x}$ — минимум.

Согласно первой теореме отделимости найдется такой ненулевой функционал, т. е. вектор $\overline{\lambda} = (\lambda_0, \lambda_1, \dots, \lambda_m) \in (\mathbb{R}^{m+1})^*$, что

$$\sum_{j=0}^{m} \lambda_j \mu_j \ge 0 \tag{11}$$

для всех $\mu = (\mu_0, \mu_1, \dots, \mu_m)^T \in M$. Пусть $\delta > 0$. Подставляя в (11) векторы $(1, \delta, \dots, \delta)^T, \dots, (\delta, \dots, \delta, 1)^T$, а затем устремляя δ к нулю, получаем, что $\lambda_j \geq 0$, $j = 0, 1, \dots, m$, т. е. доказано утверждение (b) теоремы.

Теперь подставим в (11) векторы $(\delta, \dots, \delta, f_j(\widehat{x}), \delta, \dots, \delta)^T$, $j = 1, \dots, m$ (они принадлежат M, надо взять $x = \widehat{x}$) и снова, устремляя δ к нулю, получим, что $\lambda_j f_j(\widehat{x}) \geq 0$. Но $\lambda_j f_j(\widehat{x}) \leq 0$, так как $\lambda_j \geq 0$, а $f_j(\widehat{x}) \leq 0$ и поэтому $\lambda_j f_j(\widehat{x}) = 0$, $j = 1, \dots, m$, что доказывает утверждение (c).

Пусть $x \in A$. Ясно, что $(f_0(x) - f_0(\widehat{x}) + \delta, f_1(x), \dots, f_m(x))^T \in M$. Подставляя этот вектор в (11), приходим (в пределе при $\delta \to 0$) к неравенству $\sum_{j=0}^m \lambda_j f_j(x) \geq \lambda_0 f_0(\widehat{x})$. Добавляя справа нулевые слагаемые $\lambda_j f_j(\widehat{x})$, $j = 1, \dots, m$, получаем, что $\mathcal{L}(x, \overline{\lambda}) \geq \mathcal{L}(\widehat{x}, \overline{\lambda})$ и (a) доказано.

2. Второе утверждение теоремы непосредственно вытекает из леммы 1.

3. Докажем последнее утверждение теоремы. Если $\lambda_0 = 0$, то ненулевой множитель Лагранжа находится среди остальных и поэтому (с учетом (c)) $\mathcal{L}(\overline{x}, \overline{\lambda}) = \sum_{i=1}^m \lambda_i f_i(\overline{x}) < 0 = \sum_{i=1}^m \lambda_i f_i(\widehat{x}) = \mathcal{L}(\widehat{x}, \overline{\lambda})$, что противоречит (a). □

Лекция 6

12. Субдифференциал

Пусть X — линейное нормированное пространство, $f\colon X \to \overline{\mathbb{R}}$, $\widehat{x} \in X$ и функция f конечна в точке $\widehat{x}$. Субдифференциалом функции f в точке $\widehat{x}$ называется множество (возможно пустое)

$$\partial f(\widehat{x}) = \{\, x^* \in X^* : f(x) - f(\widehat{x}) \geq \langle x^*, x - \widehat{x} \rangle,\ \forall\, x \in X \,\}.$$

Предложение 11. *Субдифференциал функции является выпуклым множеством.*

Доказательство. Если субдифференциал — пустое множество, то утверждение вытекает из того, что пустое множество — выпукло. Пусть субдифференциал функции f в точке $\widehat{x}$ не пуст и $x_1^*, x_2^* \in \partial f(\widehat{x})$. Тогда для всех $x \in X$

$$f(x) - f(\widehat{x}) \geq \langle x_1^*, x - \widehat{x} \rangle, \quad f(x) - f(\widehat{x}) \geq \langle x_2^*, x - \widehat{x} \rangle.$$

Отсюда при всех $\alpha \in [0, 1]$ имеем

$$(1-\alpha)(f(x) - f(\widehat{x})) + \alpha(f(x) - f(\widehat{x})) \geq (1-\alpha)\langle x_1^*, x - \widehat{x} \rangle + \alpha \langle x_2^*, x - \widehat{x} \rangle.$$

Следовательно,

$$f(x) - f(\widehat{x}) \geq \langle (1-\alpha) x_1^* + \alpha x_2^*, x - \widehat{x} \rangle.$$

Это означает, что $(1-\alpha)x_1^* + \alpha x_2^* \in \partial f(\widehat{x})$. □

Для функции одной переменной субдифференциал $\partial f(\widehat{x})$ это совокупность угловых коэффициентов k, при которых прямые $y =$

$kx+b$, проходящие через точку $(\widehat{x}, f(\widehat{x}))$ лежат под графиком функции $y = f(x)$

Пример 12.1. Пусть $f(x) = |x|$. Тогда

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

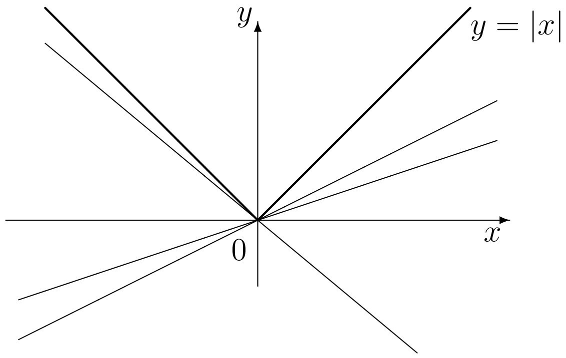

Рис. 1

На рис. 1 изображены несколько линий, проходящих через точку $(0, 0)$ с угловыми коэффициентами из множества $\partial|x|$ при $x = 0$.

Пример 12.2. Пусть X — линейное нормированное пространство и $f(x) = \|x\|$. Найдем сначала $\partial f(0)$. Поскольку $f(0) = 0$, то из определения субдифференциала вытекает, что

$$\partial f(0) = \{\, x^* \in X^* : \|x\| \geq \langle x^*, x\rangle,\ \forall\, x \in X \,\}.$$

Покажем, что единичный шар сопряженного пространства

$$BX^* = \{\, x^* \in X^* : \|x^*\| \leq 1 \,\}$$

содержится в $\partial f(0)$. Действительно, если $x^* \in BX^*$, то для всех $x \in X$

$$\langle x^*, x\rangle \leq \|x^*\|\|x\| \leq \|x\|.$$

Если теперь $x^* \in \partial f(0)$, то $x^* \in BX^*$, ибо если $\|x^*\| > 1$, то существует $x \in X$ такой, что $\|x\| \leq 1$ и $\langle x^*, x\rangle > 1 \geq \|x\|$. Итак, доказано, что $\partial f(0) = BX^*$.

Найдем теперь $\partial f(x_0)$ при $x_0 \neq 0$. Имеем

$$\partial f(x_0) = \{\, x^* \in X^* : \|x\| - \|x_0\| \geq \langle x^*, x - x_0\rangle,\ \forall\, x \in X \,\}.$$

Предположим, что $x^* \in \partial f(x_0)$. Подставив $x = 0$, получим, что $\langle x^*, x_0\rangle \geq \|x_0\|$. С другой стороны, если подставить $x = 2x_0$, то получим, что $\|x_0\| \geq \langle x^*, x_0\rangle$. Следовательно, $\langle x^*, x_0\rangle = \|x_0\|$. Тогда для всех $x \in X$ должно выполняться неравенство $\|x\| \geq \langle x^*, x\rangle$. Как было показано выше, это означает, что $x^* \in BX^*$. Нетрудно убедиться, что если $x^* \in BX^*$ и $\langle x^*, x_0\rangle = \|x_0\|$, то $x^* \in \partial f(x_0)$. Таким образом, при $x_0 \neq 0$

$$\partial f(x_0) = \{\, x^* \in BX^* : \langle x^*, x_0\rangle = \|x_0\|\,\}.$$

Пусть X — линейное нормированные пространства, а U — открытое подмножество X. Функция $f\colon U \to \mathbb{R}$ называется дифференцируемой в точке $\widehat{x} \in U$, если найдется такой линейный функционал $x^* \in X^*$, что для всех $h \in X$, для которых $\widehat{x} + h \in U$ справедливо представление

$$f(\widehat{x} + h) = f(\widehat{x}) + \langle x^*, h\rangle + r(h), \tag{12}$$

где $r(h) = o(\|h\|_X)$ ($|r(h)|/\|h\|_X \to 0$ при $h \to 0$). Линейный функционал x^* называется производной функции f в точке $\widehat{x}$ и обозначается $f'(\widehat{x})$.

Следующее предложение показывает, что субдифференциал достаточно естественное обобщение понятия производной на выпуклые функции.

Предложение 12. *Пусть X — линейное нормированное пространство и $f\colon X \to \overline{\mathbb{R}}$ — выпуклая функция, дифференцируемая в точке $\widehat{x}$. Тогда $\partial f(\widehat{x}) = \{f'(\widehat{x})\}$.*

Доказательство. Пусть $x \in X$. Для любого достаточно малого $\alpha > 0$ имеем по неравенству Йенссена

$$f((1-\alpha)\widehat{x} + \alpha x) \le (1-\alpha) f(\widehat{x}) + \alpha f(x),$$

откуда

$$f(\widehat{x} + \alpha(x - \widehat{x})) - f(\widehat{x}) \le \alpha(f(x) - f(\widehat{x})).$$

В силу дифференцируемости функции f в точке $\widehat{x}$ имеем

$$\alpha\langle f'(\widehat{x}), x - \widehat{x}\rangle + o(\alpha) \le \alpha(f(x) - f(\widehat{x})).$$

Сокращая на α и переходя к пределу при $\alpha \to 0$, получаем, что $f'(\widehat{x}) \in \partial f(\widehat{x})$.

Обратно, если $x^* \in \partial f(\widehat{x})$, то для любого $x \in X$ и любого $t > 0$ имеем $f(\widehat{x} + tx) - f(\widehat{x}) \ge t\langle x^*, x\rangle$. Следовательно,

$$t\langle f'(\widehat{x}), x\rangle + o(t) \ge t\langle x^*, x\rangle,$$

т. е. $\langle f'(\widehat{x}), x\rangle \ge \langle x^*, x\rangle$ для любого x и значит, $x^* = f'(\widehat{x})$. □

13. Теорема Ферма в субдифференциальной форме

Пусть X — линейное нормированное пространство и $f\colon X \to \overline{\mathbb{R}}$ — собственная функция. Рассмотрим задачу

$$f(x) \to \min, \quad x \in X. \tag{13}$$

Теорема 12 (Ферма в субдифференциальной форме)**.** *Точка $\widehat{x} \in \operatorname{dom} f$ является глобальным минимумом в задаче* (13) *тогда и только тогда, когда $0 \in \partial f(\widehat{x})$.*

Доказательство. Если $\widehat{x}$ — глобальный минимум, то $f(x)-f(\widehat{x}) \geq 0 = \langle 0, x-\widehat{x}\rangle$ для любого $x \in X$, т. е. $0 \in \partial f(\widehat{x})$. Если $0 \in \partial f(\widehat{x})$, то $f(x)-f(\widehat{x}) \geq \langle 0, x-\widehat{x}\rangle = 0$, т. е. $f(x) \geq f(\widehat{x})$ для любого $x \in X$. □

Из предложения 12 и теоремы 12 вытекает

Следствие 4. *Если в задаче* (13) f *— выпуклая функция, дифференцируемая в точке* $\widehat{x}$, *то* $\widehat{x}$ *— глобальный минимум в том и только в том случае, если* $f'(\widehat{x}) = 0$.

Предложение 13. *Если* X *— линейное нормированное пространство и у выпуклой собственной функции* $f\colon X \to \overline{\mathbb{R}}$ *в точке* $\widehat{x} \in \operatorname{dom} f$ *локальный минимум, то в точке* $\widehat{x}$ *и глобальный минимум.*

Доказательство. Пусть $\widehat{x}$ — локальный минимум, т. е. существует такая окрестность U точки $\widehat{x}$, что $f(\widehat{x}) \leq f(x)$ для всех $x \in U$. Пусть теперь x — произвольная точка из X. Для достаточно малых $0 < \alpha \leq 1$ точки $(1-\alpha)\widehat{x} + \alpha x$ принадлежат U и поэтому (по неравенству Йенссена) $f(\widehat{x}) \leq f((1-\alpha)\widehat{x}+\alpha x) \leq (1-\alpha)f(\widehat{x})+\alpha f(x)$, откуда следует, что $f(\widehat{x}) \leq f(x)$. □

Предложение 14. *Если* X *— линейное нормированное пространство,* $f\colon X \to \overline{\mathbb{R}}$ *— собственная выпуклая функция и для некоторого* $x^* \in X^*$ *в окрестности точки* $\widehat{x} \in \operatorname{dom} f$ *выполняется неравенство*

$$f(x) - f(\widehat{x}) \geq \langle x^*, x-\widehat{x}\rangle, \tag{14}$$

то $x^* \in \partial f(\widehat{x})$.

Доказательство. Функция $g(x) = f(x)-f(\widehat{x})-\langle x^*, x-\widehat{x}\rangle$ является выпуклой и в точке $\widehat{x}$ имеет локальный минимум, равный нулю. Из предложения 13 следует, что $\widehat{x}$ — глобальный минимум, т.е. для всех $x \in X$ выполняется неравенство (14). Это и означает, что $x^* \in \partial f(\widehat{x})$. □

14. Субдифференциальное исчисление. Теорема Моро–Рокафеллара

Теорема 13 (Моро–Рокафеллара)**.** *Если* $f_1, f_2\colon X \to \overline{\mathbb{R}}$ *— собственные выпуклые функции, хотя бы одна из которых непрерывна в точке* $\widehat{x}$, *то*

$$\partial(f_1+f_2)(\widehat{x}) = \partial f_1(\widehat{x}) + \partial f_2(\widehat{x}).$$

Лекция 7

Доказательство. Докажем, что

$$\partial f_1(\widehat{x}) + \partial f_2(\widehat{x}) \subset \partial(f_1+f_2)(\widehat{x}).$$

Пусть $x^* \in \partial f_1(\widehat{x}) + \partial f_2(\widehat{x})$. Это означает, что существуют $x_1^* \in \partial f_1(\widehat{x})$ и $x_2^* \in \partial f_2(\widehat{x})$ такие, что $x^* = x_1^* + x_2^*$. Имеем

$$\begin{aligned}(f_1+f_2)(x) - (f_1+f_2)(\widehat{x}) - \langle x^*, x-\widehat{x}\rangle = f_1(x) - f_1(\widehat{x}) - \langle x_1^*, x-\widehat{x}\rangle \\ + f_2(x) - f_2(\widehat{x}) - \langle x_2^*, x-\widehat{x}\rangle \ge 0.\end{aligned}$$

Следовательно, $x^* \in \partial(f_1+f_2)(\widehat{x})$.

Докажем теперь, что

$$\partial(f_1+f_2)(\widehat{x}) \subset \partial f_1(\widehat{x}) + \partial f_2(\widehat{x}).$$

Без ограничения общности можно читать, что $f_1(\widehat{x}) = f_2(\widehat{x}) = 0$. Если это не так, то можно рассмотреть функции

$$g_1(x) = f_1(x) - f_1(\widehat{x}), \quad g_2(x) = f_2(x) - f_2(\widehat{x}).$$

Докажем, что если

$$0 \in \partial(f_1+f_2)(\widehat{x}), \tag{15}$$

то найдется $\widetilde{x}^* \in \partial f_1(\widehat{x})$ такой, что $-\widetilde{x}^* \in \partial f_2(\widehat{x})$. Из (15) следует, что для всех $x \in X$

$$(f_1+f_2)(x) \ge (f_1+f_2)(\widehat{x}) = 0. \tag{16}$$

Будем считать, что f_2 непрерывна в $\widehat{x}$. Положим

$$A = \{\, (x,v) : x \in \operatorname{dom} f_1,\ v < -f_1(x) \,\},$$
$$B = \operatorname{epi} f_2 = \{\, (x,u) : x \in \operatorname{dom} f_2,\ u \ge f_2(x) \,\}.$$

Эти два множества выпуклы и непусты. Они не пересекаются, так как из того, что $(x,v) \in A$ следует, учитывая (16), что $v < -f_1(x) \le f_2(x)$. Тем самым $(x,v) \notin \operatorname{epi} f_2 = B$. Поскольку f_2 непрерывна в $\widehat{x}$, она ограничена в некоторой окрестности U этой точки. Положим

$$M = \sup_{x\in U} f_2(x).$$

Тогда открытое множество

$$V = \{\, (x,v) : x \in U,\ v > M \,\}$$

лежит в $\operatorname{epi} f_2$. Следовательно, $\operatorname{int} B \ne \emptyset$. По первой теореме отделимости (теорема 6, конечномерный случай — теорема 9) множества A и B можно отделить, т.е. существуют $x^* \in X^*$ и $\lambda \in \mathbb{R}$, не равные одновременно нулю, такие, что для любых $(x_1, v) \in A$ и $(x_2, u) \in B$ выполняется неравенство

$$\langle x^*, x_1\rangle + \lambda v \le \langle x^*, x_2\rangle + \lambda u. \tag{17}$$

Положим в (17) $x_1 = x_2 = \widehat{x}$, $v < 0$, а $u = f_2(x_2) = f_2(\widehat{x}) = 0$. Тогда

$$\langle x^*, \widehat{x}\rangle + \lambda v \le \langle x^*, \widehat{x}\rangle.$$

Следовательно, $\lambda v \le 0$, а так как $v < 0$, получаем, что $\lambda \ge 0$.

Если предположить, что $\lambda = 0$, то из (17) получаем

$$\sup_{x\in\operatorname{dom} f_1}\langle x^*, x\rangle \le \inf_{x\in\operatorname{dom} f_2}\langle x^*, x\rangle.$$

Поскольку $x^* \ne 0$, а линейная функция, отличная от тождественного нуля, не может принимать минимального значения во внутренней точке, то

$$\langle x^*, \widehat{x}\rangle \le \sup_{x\in\operatorname{dom} f_1}\langle x^*, x\rangle \le \inf_{x\in\operatorname{dom} f_2}\langle x^*, x\rangle < \langle x^*, \widehat{x}\rangle.$$

Полученное противоречие показывает, что $\lambda \ne 0$.

Пусть $x_1 = \widehat{x}$, $v < 0$, $x_2 \in X$, $u = f_2(x_2)$. Тогда из (17) получаем

$$\langle x^*, \widehat{x}\rangle + \lambda v \le \langle x^*, x_2\rangle + \lambda f_2(x_2).$$

Устремляя v к нулю, получаем

$$\langle x^*, \widehat{x}\rangle \le \langle x^*, x_2\rangle + \lambda f_2(x_2).$$

Отсюда

$$f_2(x_2) \ge \langle -\lambda^{-1}x^*, x_2 - \widehat{x}\rangle.$$

Следовательно, $-\lambda^{-1}x^* \in \partial f_2(\widehat{x})$.

Пусть теперь $x_2 = \widehat{x}$, $x_1 \in \operatorname{dom} f_1$. Тогда для любого $v < -f_1(x_1)$ выполняется неравенство

$$\langle x^*, x_1\rangle + \lambda v \le \langle x^*, \widehat{x}\rangle + \lambda f_2(\widehat{x}) = \langle x^*, \widehat{x}\rangle.$$

Устремляя v к $f_1(x_1)$, получаем

$$\langle x^*, x_1\rangle - \lambda f_1(x_1) \le \langle x^*, \widehat{x}\rangle,$$

откуда $\lambda^{-1}x^* \in \partial f_1(\widehat{x})$. Тем самым доказано, что $0 \in \partial f_1(\widehat{x}) + \partial f_2(\widehat{x})$.

Предположим, что $x^* \in \partial(f_1 + f_2)(\widehat{x})$. Положим

$$g(x) = f_1(x) - \langle x^*, x - \widehat{x}\rangle.$$

Нетрудно убедиться, что $0 \in \partial(g + f_2)(\widehat{x})$. По доказанному выше $0 \in \partial g(\widehat{x}) + \partial f_2(\widehat{x})$. Тем самым найдется $\widetilde{x}^* \in \partial g(\widehat{x})$ такой, что $-\widetilde{x}^* \in \partial f_2(\widehat{x})$. Включение $\widetilde{x}^* \in \partial g(\widehat{x})$ означает, что выполняется неравенство

$$f_1(x) - \langle x^*, x - \widehat{x}\rangle \ge \langle \widetilde{x}^*, x - \widehat{x}\rangle.$$

Следовательно, $x^* + \widetilde{x}^* \in \partial f_1(\widehat{x})$. Поэтому $x^* \in \partial f_1(\widehat{x}) + \partial f_2(\widehat{x})$. $\square$

Следствие 5. *Если f — выпуклая функция, непрерывная в точке $\widehat{x}$, то $\partial f(\widehat{x}) \ne \emptyset$.*

Доказательство. Отметим сначала, что функция f в силу непрерывности в точке $\widehat{x}$ является собственной (докажите!). Обозначим через $\delta\{\widehat{x}\}$ индикаторную функцию точки $\widehat{x}$

$$\delta\{\widehat{x}\}(x) = \begin{cases} 0, & x = \widehat{x}, \\ +\infty, & x \ne \widehat{x}. \end{cases}$$

Функция $\delta\{\widehat{x}\}$ является собственной выпуклой функцией и нетрудно убедиться, что $\partial\delta\{\widehat{x}\}(\widehat{x}) = X^*$. Рассмотрим функцию $f + \delta\{\widehat{x}\}$. Из теоремы Моро–Рокафеллара получаем

$$X^* = \partial(f + \delta\{\widehat{x}\})(\widehat{x}) = \partial f(\widehat{x}) + \partial\delta\{\widehat{x}\})(\widehat{x}) = \partial f(\widehat{x}) + X^*.$$

Отсюда следует, что $\partial f(\widehat{x}) \neq \emptyset$. □

15. Теорема Дубовицкого–Милютина

Теорема 14 (Дубовицкого–Милютина)**.** *Если* $f_1, \ldots, f_n \colon X \to \overline{\mathbb{R}}$ *— собственные выпуклые функции, непрерывные в точке* $\widehat{x}$, *и*

$$f(x) = \max\{f_1(x), \ldots, f_n(x)\},$$

то

$$\partial f(\widehat{x}) = \operatorname{co}\Bigg(\bigcup_{j \in I} \partial f_j(\widehat{x})\Bigg),$$

где $I = \{j : f_j(\widehat{x}) = f(\widehat{x})\}$.

Доказательство. Сначала рассмотрим случай, когда

$$f_1(\widehat{x}) = \ldots = f_n(\widehat{x}) = f(\widehat{x}).$$

Докажем, что

$$\operatorname{co}\Bigg(\bigcup_{j=1}^{n} \partial f_j(\widehat{x})\Bigg) \subset \partial f(\widehat{x}). \tag{18}$$

Пусть $x^* \in \partial f_j(\widehat{x})$ при некотором j, $1 \leq j \leq n$. Тогда для всех $x \in X$

$$f(x) \geq f_j(x) \geq f_j(\widehat{x}) + \langle x^*, x - \widehat{x}\rangle = f(\widehat{x}) + \langle x^*, x - \widehat{x}\rangle.$$

Следовательно, $x^* \in \partial f(\widehat{x})$. Тем самым $\partial f_j(\widehat{x}) \subset \partial f(\widehat{x})$. Отсюда

$$\bigcup_{j=1}^{n} \partial f_j(\widehat{x}) \subset \partial f(\widehat{x}).$$

В силу выпуклости субдифференциала (см. предложения 11 и 1) имеет место (18).

Пусть $x^* \in \partial f(\widehat{x})$. Положим

$$\varphi_j(x) = f_j(x) - f_j(\widehat{x}) - \langle x^*, x - \widehat{x}\rangle, \quad j = 1, \ldots, n.$$

Для каждого $k = 1, \ldots, n$ рассмотрим систему неравенств

$$\begin{cases} \varphi_k(x) < 0, \\ \varphi_j(x) \leq 0, \quad j \neq k. \end{cases} \tag{19}$$

Предположим, что для всех k эти системы совместны и обозначим через x_k какое-либо решение каждой из этой системы. Рассмотрим точку

$$\widetilde{x} = \frac{1}{n}\sum_{k=1}^{n} x_k.$$

Тогда при каждом $k=1,\dots,n$
$$\varphi_k(\widetilde{x})\le\frac{1}{n}\sum_{j=1}^n\varphi_k(x_j)\le\frac{1}{n}\varphi_k(x_k)<0.$$
Отсюда
$$f_k(\widetilde{x})-f(\widehat{x})-\langle x^*,\widetilde{x}-\widehat{x}\rangle<0,\quad k=1,\dots,n.$$
Тем самым
$$f(\widetilde{x})-f(\widehat{x})-\langle x^*,x-\widehat{x}\rangle<0,$$
что противоречит тому, что $x^*\in\partial f(\widehat{x})$. Следовательно, найдется k, $1\le k\le n$, для которого система (19) будет несовместна. Из непрерывности функций f_j, $j=1,\dots,n$, в точке $\widehat{x}$ следует, что существует окрестность U точки $\widehat{x}$, в которой все функции f_j конечны. Тогда для такого k экстремальная задача
$$\varphi_k(x)\to\min,\quad\varphi_j(x)\le0,\quad j\ne k,\quad x\in U$$
будет иметь решение $\widehat{x}$. По теореме Каруша–Куна–Таккера (теорема 11) найдутся $\lambda_1,\dots,\lambda_n\ge0$, не все равные нулю, для которых при всех $x\in U$
$$\sum_{j=1}^n\lambda_j\varphi_j(x)\ge\sum_{j=1}^n\lambda_j\varphi_j(\widehat{x})=0.$$
Из предложения 13 вытекает, что это неравенство выполняется для всех $x\in X$. Без ограничения общности можно считать, что
$$\sum_{j=1}^n\lambda_j=1.$$
Тогда получаем, что для всех $x\in X$
$$\sum_{j=1}^n\lambda_jf_j(x)-\sum_{j=1}^n\lambda_jf_j(\widehat{x})-\langle x^*,x-\widehat{x}\rangle\ge0.$$
Следовательно, пользуясь теоремой Моро–Рокафеллара, получаем
$$x^*\in\partial\biggl(\sum_{j=1}^n\lambda_jf_j\biggr)(\widehat{x})=\sum_{j=1}^n\partial(\lambda_jf_j)(\widehat{x})$$
$$=\sum_{j=1}^n\lambda_j\partial f_j(\widehat{x})\subset\operatorname{co}\biggl(\bigcup_{j=1}^n\partial f_j(\widehat{x})\biggr).$$

Рассмотрим теперь общий случай. Положим $J=\{j: f_j(\widehat{x})<f(\widehat{x})\}$ и $g(x)=\max_{j\in I}f_j(x)$. В силу непрерывности функций f_j, $j=1,\dots,n$, в точке $\widehat{x}$ существует окрестность V точки $\widehat{x}$, в которой $f_j(x)<f(x)$ для всех $j\in J$. Тогда в окрестности V функции f и g совпадают. Из предложения 14 получаем, что $\partial f(\widehat{x})=\partial g(\widehat{x})$, а для $\partial g(\widehat{x})$ нужное равенство уже доказано. □

Список литературы

[1] Алексеев В. М., Тихомиров В. М., Фомин С. В. Оптимальное управление. М.: ФИЗМАТЛИТ, 2007.

[2] Галеев Э. М. Оптимизация: Теория, примеры, задачи: Учебное пособие. М.: ЛИБРОКОМ, 2010.

[3] Жадан В. Г. Методы оптимизации. Часть I. Введение в выпуклый анализ и теорию оптимизации: учебное пособие. М.: МФТИ, 2014.

[4] Колмогоров А. Н., Фомин С. В. Элементы теории функций и фукционального анализа. М.: ФИЗМАТЛИТ, 2004.

[5] Магарил-Ильяев Г. Г., Тихомиров В. М. Выпуклый анализ и его приложения. М.: УРСС, 2003.

[6] Осипенко К. Ю. Вариационное исчисление и оптимальное управление. Курс лекций, 2015.
http://opu.math.msu.su/sites/default/files/main_courses/OsLect2015_0_0.pdf

[7] Осипенко К. Ю. Выпуклый анализ. Курс лекций, 2016.
http://opu.math.msu.su/sites/default/files/main_courses/CA_0.pdf

[8] Протасов В. Ю. Выпуклый анализ. Курс лекций, 2009.
http://opu.math.msu.su/sites/default/files/main_courses/convex-anal-new.pdf