

Тема 5

Правило множителей Лагранжа. Двойственность.

1. Правило множителей Лагранжа для выпуклых задач.

В этом пункте рассматриваются условия оптимальности для задач выпуклого программирования. Под выпуклым программированием понимается раздел теории экстремальных задач, в котором изучаются задачи минимизации выпуклых функционалов на выпуклых множествах. Мы будем рассматривать следующую постановку:

$$J(u) \rightarrow \inf_{u \in U}, \quad U = \{u \in U_0 \mid g_i(u) \leq 0, i = 1, \dots, m\}, \quad (5.1)$$

где U_0 — выпуклое множество из гильбертова пространства H , а функционалы $J(u)$, $g_1(u), \dots, g_m(u)$ выпуклы на U_0 . Такую задачу кратко принято называть выпуклой задачей. Основная идея метода множителей Лагранжа заключается в следующем: мы избавимся от неудобных для нас ограничений-неравенств и перейдем к другой задаче, где в качестве множества рассматривается "хорошее" множество U_0 . Не вдаваясь в суть этого перехода, приведем необходимые факты.

Введём функцию

$$L(u, \lambda) = J(u) + \sum_{i=1}^m \lambda_i g_i(u),$$

называемую функцией Лагранжа задачи (5.1), где $u \in U_0$, а вектор $\lambda = (\lambda_0, \lambda_1, \dots, \lambda_m)$ принадлежит множеству

$$\Lambda_0 = \{\lambda = (\lambda_0, \lambda_1, \dots, \lambda_m) \in \mathbb{R}^{m+1} : \lambda_0 \geq 0, \lambda_1 \geq 0, \dots, \lambda_m \geq 0\}.$$

Необходимые и достаточные условия оптимальности для исходной задачи описывает

Теорема. (Куна-Таккера)

Пусть задача (5.1) выпукла. Тогда, если $u_* \in U_*$, то существует такой набор множителей Лагранжа $\lambda^* = (\lambda_0^*, \lambda_1^*, \dots, \lambda_m^*) \neq \theta_{\mathbb{R}^{m+1}}$, что

- 1) $\min_{u \in U_0} L(u, \lambda^*) = L(u_*, \lambda^*)$, (принцип минимума)
- 2) $\lambda_i^* \geq 0 \quad \forall i = 0, 1, \dots, m$, (неотрицательность множителей Лагранжа)
- 3) $\lambda_i^* \cdot g_i(u_*) = 0 \quad \forall i = 1, 2, \dots, m$. (условия дополняющей нежёсткости)

Если для некоторой пары (u_*, λ^*) выполнены соотношения (1), (2), (3), причем $u_* \in U$, $\lambda_0^* \neq 0$, то $u_* \in U_*$.

Случай, когда $\lambda_0^* \neq 0$, принято называть регулярным. К сожалению, не во всех выпуклых задачах есть регулярность, и, как следствие, теорема Куна-Таккера не всегда является достаточным условием оптимальности. Это можно проиллюстрировать на примере.

Пример 1

Рассмотрим задачу $J(u) = -u \rightarrow \inf_{u \in U}$, $U = \{u \in U_0 = \mathbb{R}^1 \mid g(u) = u^2 \leq 0\}$.

Здесь функционал $J(u)$ линейный и выпуклый, функционал $g(u)$ сильно выпуклый на U_0 . Ясно, что множество U состоит из единственной точки $u = 0$, которая и будет решением этой задачи. Записывая теорему Куна-Таккера, имеем

$$\begin{cases} -\lambda_0^* u + \lambda_1^* u^2 \geq -\lambda_0^* u_* + \lambda_1^* u_*^2 \quad \forall u \in \mathbb{R}^1, \\ \lambda_1^* u_*^2 = 0 \end{cases} \Rightarrow \{u_* = 0\} \Rightarrow -\lambda_0^* + \lambda_1^* u \geq 0 \quad \forall u > 0.$$

Переходя к пределу при $u \rightarrow 0 + 0$, имеем $-\lambda_0^* \geq 0$, что в сочетании с условием неотрицательности множителей Лагранжа даёт нам $\lambda_0^* = 0$.

Сформулируем так называемое условие регулярности Слейтера, являющееся достаточным условием регулярности в выпуклых задачах.

Теорема. Пусть задача (5.1) является выпуклой, и существует такая точка $u_0 \in U_0$, что $g_i(u_0) < 0$, $i = 1, 2, \dots, m$. Тогда в любом наборе множителей Лагранжа, соответствующем произвольному $u_* \in U_*$, обязательно $\lambda_0^* > 0$.

Наконец, в регулярном случае переформулируем теорему Куна-Таккера в терминах седловой точки.

Определение. Пусть $f(x, y) : X \times Y \rightarrow \mathbb{R}^1$. Точка (x_*, y^*) называется седловой точкой функционала $f(x, y)$ на $X \times Y$, если $f(x_*, y) \leq f(x_*, y^*) \leq f(x, y^*)$ для всех $x \in X$, $y \in Y$.

Теорема. (седловая форма теоремы Куна-Таккера)

Пусть выполнены все условия теоремы Куна-Таккера, а также выполнено условие регулярности Слейтера. Тогда для того, чтобы точка u_* являлась решением задачи (5.1), необходимо и достаточно, чтобы у классической функции Лагранжа (т.е. функции Лагранжа с $\lambda_0 = 1$) существовала седловая точка вида (u_*, λ^*) на множестве $U_0 \times \mathbb{R}_+^m$. ($\mathbb{R}_+^m = \{\lambda = (\lambda_1, \lambda_2, \dots, \lambda_m) \mid \lambda_i \geq 0, i = 1, 2, \dots, m\}$)

Пример 2

В пространстве ℓ^2 рассматривается задача минимизации

$$J(x) = (x_1 + 6)^2 + (x_2 - 2)^2 + (x_3 + 2)^2 + \sum_{k=4}^{+\infty} x_k^2 \rightarrow \inf_{x \in X}$$

$$X = \{x = (x_1, x_2, \dots, x_n, \dots) \in \ell_2 : \|x\|_{\ell_2} \leq 3, x_1 + x_2 \geq 0\}.$$

Забл. 1: в различных седловых точках $(x_*, y_*) \neq (x_*, y_*)$ значения ф-ции f совпадают:

$$\begin{cases} f(x_*, y) \leq f(x_*, y_*) \leq f(x, y_*) \\ f(x_*, y) \leq f(x_*, y_*) \leq f(x, y_*) \\ f(x_*, y_*) \leq f(x_*, y_*) \leq f(x_*, y_*) \\ f(x_*, y_*) \leq f(x_*, y_*) \leq f(x_*, y_*) \end{cases}$$

Решить её с помощью правила множителей Лагранжа.

Решение. Запишем нашу задачу в терминах норм и скалярных произведений, заменив условие $\|x\|_{\ell^2} \leq 3$ на эквивалентное ему условие $\|x\|_{\ell^2}^2 \leq 9$. Такая замена хороша тем, что норма дифференцируема не везде, а квадрат нормы – везде. Нетрудно заметить, что функционал $J(x)$ имеет вид $J(x) = \|x - z\|_{\ell^2}^2$, где $z = (-6, 2, -2, 0, \dots, 0, \dots)$, а ограничение $x_1 + x_2 \geq 0$ можно записать в форме $\langle c, x \rangle_{\ell^2} \leq 0$, где $c = (-1, -1, 0, \dots, 0, \dots)$. Отметим, что $\|z\|_{\ell^2} = 2\sqrt{11}$, $\|c\|_{\ell^2} = \sqrt{2}$, $\langle c, z \rangle_{\ell^2} = 4$. Итак,

$$J(x) = \|x - z\|_{\ell^2}^2, \quad g_1(x) = \|x\|_{\ell^2}^2 - 9, \quad g_2(x) = \langle c, x \rangle_{\ell^2},$$

все эти функционалы являются выпуклыми на $U_0 = \ell^2$, задача выпуклая. Условие регулярности Слейтера тут также выполнено, поскольку, взяв $x_0 = (1, 0, 0, \dots, 0, \dots)$, мы имеем $g_1(x_0) = -8$, $g_2(x_0) = -1$. Это означает, что при выписывании функции Лагранжа мы сразу можем полагать $\lambda_0 > 0$, а именно $\lambda_0 = 1$,

$$\mathcal{L}(x, \lambda) = \|x - z\|_{\ell^2}^2 + \lambda_1(\|x\|_{\ell^2}^2 - 9) + \lambda_2 \langle c, x \rangle_{\ell^2}, \quad x \in \ell^2, \quad \lambda_{1,2} \geq 0.$$

Далее, поскольку $U_0 = \ell^2$, а функционалы $J(x)$, $g_1(x)$, $g_2(x)$ дифференцируемы на ℓ^2 , то принцип минимума необходимо даёт нам равенство нулевому элементу градиента функции Лагранжа, взятого в точке x_* . Поэтому нам надо рассмотреть такую систему:

$$\begin{cases} \mathcal{L}'_x(x_*, \lambda) = 2(x_* - z) + 2\lambda_1 x_* + \lambda_2 c = \Theta_{\ell^2}, \\ \lambda_1(\|x_*\|_{\ell^2}^2 - 9) = 0, \\ \lambda_2 \langle c, x_* \rangle_{\ell^2} = 0. \end{cases}$$

$\mathcal{L}(x, \lambda^*)$ - выш. как $\rightarrow$ и вост. сумма выш.

Рассмотрим условие дополняющей нежёсткости. Если $\lambda_1 = \lambda_2 = 0$, то из первого уравнения системы имеем $x_* = z$. Но точка z не попадает в множество X , поэтому такой вариант нам ничего не даёт.

Если $\lambda_1 = 0$, $\lambda_2 \neq 0$, то мы имеем

$$\begin{cases} 2(x_* - z) + \lambda_2 c = \Theta_{\ell^2}, \\ \langle c, x_* \rangle_{\ell^2} = 0 \end{cases} \Rightarrow \begin{cases} x_* = z - \frac{1}{2}\lambda_2 c, \\ \langle c, z \rangle_{\ell^2} - \frac{1}{2}\lambda_2 \langle c, c \rangle_{\ell^2} = 0 \end{cases} \Rightarrow \lambda_2 = \frac{2\langle c, z \rangle_{\ell^2}}{\|c\|_{\ell^2}^2} = 4,$$

$$x_* = z - 2c = (-4, 4, -2, 0, \dots, 0, \dots).$$

Однако, $\|x_*\|_{\ell^2}^2 = 36$, поэтому x_* не попадает в множество X и этот вариант также ничего нам не даёт.

Если $\lambda_1 \neq 0$, $\lambda_2 = 0$, то

$$\begin{cases} 2(x_* - z) + 2\lambda_1 x_* = \Theta_{\ell^2}, \\ \|x_*\|_{\ell^2}^2 = 9 \end{cases} \Rightarrow \begin{cases} x_* = \frac{1}{1+\lambda_1} z, \\ \frac{\|z\|_{\ell^2}^2}{(1+\lambda_1)^2} = 9 \end{cases} \Rightarrow 1 + \lambda_1 = \pm \frac{2\sqrt{11}}{3}, \quad \lambda_1 = -1 \pm \frac{2\sqrt{11}}{3}.$$

Поскольку $\lambda_1 \geq 0$, то

$$\lambda_1 = -1 + \frac{2\sqrt{11}}{3}, \quad x^* = \frac{3}{2\sqrt{11}} z, \quad \langle c, x_* \rangle_{\ell^2} = \frac{3}{2\sqrt{11}} \langle c, z \rangle_{\ell^2} = \frac{6}{\sqrt{11}} > 0,$$

полученная точка x_* не попадает в множество X . Этот вариант также нам не подходит. Наконец, если $\lambda_1 \neq 0$, $\lambda_2 \neq 0$, имеем

$$\begin{cases} 2(x_* - z) + 2\lambda_1 x_* + \lambda_2 c = \Theta_{\ell^2}, \\ \|x_*\|_{\ell^2}^2 = 9, \\ \langle c, x_* \rangle_{\ell^2} = 0. \end{cases} \Rightarrow \begin{cases} x_* = \frac{1}{1+\lambda_1} (z - \frac{1}{2}\lambda_2 c), \\ \frac{\|z - \frac{1}{2}\lambda_2 c\|_{\ell^2}^2}{(1+\lambda_1)^2} = 9, \\ \frac{1}{1+\lambda_1} (\langle c, z \rangle_{\ell^2} - \frac{1}{2}\lambda_2 \langle c, c \rangle_{\ell^2}) = 0. \end{cases}$$

Значит,

$$\lambda_2 = \frac{2\langle c, z \rangle_{\ell^2}}{\|c\|_{\ell^2}^2} = 4; \quad (1 + \lambda_1)^2 = \frac{\|z - 2c\|_{\ell^2}^2}{9} = 4 \Rightarrow \{\lambda_1 \geq 0\} \Rightarrow \lambda_1 = 1;$$

$$x_* = \frac{1}{2} (z - 2c) = (-2, 2, -1, 0, \dots, 0, \dots) \in X$$

Замечание 1:

Можно также отметить, что функционал $J(x)$ сильно выпуклый и дифференцируемый, множество X выпукло как пересечение выпуклых множеств (полупространства и шара), стало быть, по теореме Вейерштрасса для сильно выпуклых функционалов, рассматриваемая задача имеет единственное решение.

Замечание 2:

Данная задача есть задача проектирования точки z на конус.

Упражнение:

Проверить для x_* критерий оптимальности.

$$\langle J'(x_*), x - x_* \rangle_{\ell^2} = \langle 2(x_* - z), x - x_* \rangle_{\ell^2} = \langle (-z - 2c), x - \frac{1}{2}(z - 2c) \rangle_{\ell^2} =$$

$$= \frac{1}{2} \langle z - 2c, z + 2c \rangle_{\ell^2} - \langle z - 2c, x \rangle_{\ell^2} - \langle 4c, x \rangle_{\ell^2} \geq$$

$$\geq \frac{1}{2} (\|z\|_{\ell^2}^2 - 4\|c\|_{\ell^2}^2) - \|z - 2c\|_{\ell^2} \cdot \|x\|_{\ell^2} - \langle 4c, x \rangle_{\ell^2} = 18 - 6\|x\|_{\ell^2} - 4\langle c, x \rangle_{\ell^2} \geq 0$$

Если $x \in X$, то $\|x\|_{\ell^2} \leq 3$, $\langle c, x \rangle_{\ell^2} \leq 0$, и выражение в правой части полученного соотношения неотрицательно. Значит, $\langle J'(x_*), x - x_* \rangle_{\ell^2} \geq 0$ для всех $x \in X$, поэтому x_* минимизирует $J(x)$ на X .

процессор. 2,3,4
↓
V-слайд камня.

$\varphi(\lambda) = \mathcal{L}(u(\lambda), \lambda)$
 $\varphi^* = \varphi_*$, то $\varphi^* = \varphi(\lambda^*) = \sup_{\lambda \in T} \left(\inf_{u \in U} \mathcal{L}(u, \lambda) \right) = \varphi_* = \mathcal{L}_*$
 $\mathcal{L}(u(\lambda^*), \lambda^*)$
 и см. выше

сн. Утверждение: Рассм-ся $A \in \mathcal{L}(H \rightarrow H)$ H -малоб.
 А снро малоб., о.е. $\langle Ax, x \rangle_H \geq \mu \|x\|_H^2$ $\mu > 0$. - Дор. 6.3.1
 Тогда $\exists A^{-1} \in \mathcal{L}(H \rightarrow H)$: A^{-1} обратный, снр-на всем $u \in H$,
 снро малоб., о.е. $\langle A^{-1}x, x \rangle_H \geq \varepsilon \|x\|_H^2$ $\varepsilon > 0$.

Лем. 1: $\exists \nu > 0$: $\|Ax\|_H \geq \nu \|x\|_H \quad \forall x \in H$

 Обратн., кстанн., неверно: в $\mathbb{R}^2$ $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$
 $Ax = \begin{pmatrix} x_2 \\ -x_1 \end{pmatrix}$, $\|Ax\| = \|x\|$, $\langle Ax, x \rangle = 0$ $\forall x$ $\implies A^{-1} = -A$

Дов-во от противного: пусть $\forall \nu_k > 0 \exists x_k \neq 0$: $\|Ax_k\|_H < \nu_k \|x_k\|_H \iff \|A \frac{x_k}{\|x_k\|}\| < \nu_k \xrightarrow[k \rightarrow \infty]{} 0$
 получим $\{z_k\}$: $\|Az_k\|_H \xrightarrow[k \rightarrow \infty]{} 0$
 $\{z_k\} \subset H, \|z_k\|_H = 1$

Но тогда $0 < \mu = \mu \|z_k\|_H^2 \leq |\langle Az_k, z_k \rangle| \leq \frac{\|Az_k\|_H \cdot \|z_k\|_H}{\|z_k\|_H} < \nu_k \xrightarrow[k \rightarrow \infty]{} 0$ успешно, з.д.д. 1
 Но тогда $\{z_k\}$ - фундамент.

Лем. 2: $AH \equiv \text{Im } A$ - замкнут

 Но снро $\text{Ker } A$ всегда замкнут:
 Введем $x_n \xrightarrow[n \rightarrow \infty]{} x_0$, $Ax_n = 0$
 $\|Ax_n - Ax_0\|_H = \|Ax_n\|_H = 0 \implies \|Ax_0\|_H = 0 \implies x_0 \in \text{Ker } A$

Д-во: рассм. $\forall y_n = Ax_n \xrightarrow[n \rightarrow \infty]{} y_0 \notin \text{Im } A$ о.е. надо, чтобы $\exists x_0 \in H$: $Ax_0 = y_0$

$\{y_n\}$ - фундамент., о.к. сн-ся $\implies \|y_n - y_k\|_H \xrightarrow[n, k \rightarrow \infty]{} 0$
 $\implies \forall \|x_n - x_k\|_H \leq \|A(x_n - x_k)\|_H = \|Ax_n - Ax_k\|_H = \|y_n - y_k\|_H \implies \|x_n - x_k\|_H \leq \frac{1}{\nu} \|y_n - y_k\|_H \xrightarrow[n, k \rightarrow \infty]{} 0$
 т.е. $\{x_n\}$ - тоже фундамент.

$\implies$ о.к. H -малое, $x_n \xrightarrow[n \rightarrow \infty]{} x_0$ о.е. $\|x_n - x_0\|_H \xrightarrow[n \rightarrow \infty]{} 0$

$\implies \|Ax_n - Ax_0\|_H \leq \|A\| \cdot \|x_n - x_0\|_H \xrightarrow[n \rightarrow \infty]{} 0$
 $\implies \begin{cases} y_n \xrightarrow[n \rightarrow \infty]{} y_0 \\ y_n \xrightarrow[n \rightarrow \infty]{} Ax_0 \end{cases} \implies y_0 = Ax_0 \in \text{Im } A$ з.д.д. 2

Лем. 3: $\text{Im } A = H$ $\iff \text{Im } A \subset H$!

Д-во от противного: Пусть $\text{Im } A \subset H$ $\implies \exists z \neq 0, z \perp \text{Im } A$

 Имеем $\forall y_1 = Ax_1 \in \text{Im } A$
 $y_2 = Ax_2 \in \text{Im } A$
 $\alpha y_1 + \beta y_2 = A(\alpha x_1 + \beta x_2) \in \text{Im } A$

$\implies \langle Az, z \rangle_H = 0$
 Но з.д.д. $\langle Az, z \rangle_H \geq \mu \|z\|_H^2 > 0$
успешно, з.д.д. 3

$\vec{z} \perp \text{Im } A$
 $\vec{z} \neq \vec{0}$

Если $x \in \text{Im } A$:
 $\langle y - \frac{\langle y, z \rangle}{\|z\|^2} z, z \rangle_H = 0$
 $\implies \langle y, z \rangle_H = 0$ $\forall y \in \text{Im } A$
 т.е. $\text{Im } A \perp \text{Ker } A$

—Доп. 63.2—

Но т.к. $\text{Tim } A = H$, то нар. A , действую-
 щий на H -базисах в H -базисе, по теореме
 Банаха {т.к. инв.}

Рассм. для $\forall y \in H \quad \langle A^{-1}y, y \rangle_H = \langle A^{-1}Ax, Ax \rangle_H = \langle Ax, x \rangle_H \geq \mu \|x\|_H^2$

Вашинкодн. стр.

$$\|Ax\|_H \leq \|A\| \cdot \|x\|_H \rightarrow \|x\| \geq \frac{\|Ax\|}{\|A\|}$$

($> 0, \forall k. A \neq \theta \in \mathcal{L}(H \rightarrow H)$)

$$\textcircled{\geq} \quad \mu \frac{\|Ax\|_H^2}{\|A\|^2} = \frac{\mu}{\|A\|^2} \cdot \|y\|_H^2 \rightarrow \text{z.v.g. } A^{-1} \text{ пропало на } \sup H$$

Замечание 1: Если $H \xrightarrow{A} H$ для $A \in \mathcal{Z}(H \rightarrow H)$, то $\exists A^{-1} \in \mathcal{Z}(H \rightarrow H)$ сг-го $\|AH \cdot HA^{-1}\| \geq 1$
Взаимноодн.

$$\underline{\text{Ex-60:}} \quad \forall x \in H \quad \|x\| = \|A y\| \leq \|A\| \cdot \|y\| \leq \|A\| \cdot \|A^{-1}\| \cdot \|x\| \rightarrow \text{2.6.21}$$

Замечание 2: Если дан $A \in \mathcal{L}(H \rightarrow H)$ имеет место $\langle Ax, x \rangle \geq \mu \|x\|^2$, то $\|A^{-1}\| \geq \frac{\mu}{\|A\|^2}$

One $y \neq 0$ $\mu \|y\|^2 \leq \langle Ay, y \rangle = \langle A^{-1}x, x \rangle \leq \|A^{-1}\| \cdot \|x\|^2 = \|A^{-1}\| \cdot \|Ay\|^2 \leq \|A^{-1}\| \cdot \|A\|^2 \cdot \|y\|^2 \rightarrow \text{2.o.d.}$ } exp. non-const. sup. }

Lemma 5: $A \in \mathcal{L}(H \rightarrow E)$ $A^{-1}x$ $\|A\| \|A^{-1}\| \|x\| \leq \|x\|$ $\|A\| \|A^{-1}\| \leq 1$

Lemma 5: $\forall A \perp T_m A^*$ $\exists X (H \rightarrow F)$ mod. $A^{-1}X$ $(\forall y \text{ Ker } A) (\forall y T_m A^*)$
 $\text{and } \forall x \in \text{Ker } A, \forall y \in F \quad 0 = \langle Ax, y \rangle = \langle x, A^*y \rangle$ $(2.7.2)$

Замечание 3: $\langle x, A+A^* \rangle_H = \langle A+A^*, x \rangle_H = \langle A, x \rangle_H + \langle A^*, x \rangle_H$
 всегда $\langle A, h \rangle_H = \langle \frac{A+A^*}{2}, h \rangle_H$, поэтому в к. 99 при $\langle A, h \rangle_H$ можно считать $A = \frac{A+A^*}{2}$

Замечание 4: Опр-е: $K \in \mathcal{L}(H \rightarrow H)$, H -линоб. $\{ \text{т.е. } K^* = -K \}$
 если $\langle Kx, y \rangle_H = -\langle x, Ky \rangle \quad \forall x, y \in H$
 K наз. кососимметрическим.

Exercice $A = A^*$ et $\langle Ah, h \rangle \geq \mu \|h\|^2$, on $\langle (A+K)h, h \rangle = \langle Ah, h \rangle + \langle Kh, h \rangle \geq \mu \|h\|^2$

$$\text{v.k. } \langle K_h, h \rangle = -\langle h, K_h \rangle = 0 \quad \forall h \in H$$

р.р. А+К ранее сфранко иконос. сур.

$$\langle (A+K)x, y \rangle = \langle x, (A+K)^{\dagger} y \rangle \quad \Rightarrow (A+K)^{\dagger} = A - K \neq A + K$$

$$\langle Ax, y \rangle + \langle Kx, y \rangle = \langle x, (A+K)y \rangle$$

р.е. А+К уже не вл. самим.