Методы оптимизации. Семинар 9. Двойственная задача

Корнилов Никита Максимович

Московский физико-технический институт

30 октября 2025г

Постановка прямой задачи оптимизации стандартной формы:

$$\begin{aligned} \min_x \; & f_0(x) \quad - f_0(x) \\ \textbf{s.t.} \; & f_i(x) \leq 0, \;\; i = 1, \ldots, n, \\ & h_j(x) = 0, \;\; j = 1, \ldots, m, \end{aligned} \tag{1}$$

с прямой переменной $x \in \mathbb{R}^d$.

$x \in \mathcal{X}$

Лагранжиан

Definition (Лагранжиан)

Лагранжиан $L : \mathbb{R}^d \times \mathbb{R}^n \times \mathbb{R}^m \to \overline{\mathbb{R}}$ для задачи (1) задается следующим образом:

$$L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^{n} \lambda_i f_i(x) + \sum_{j=1}^{m} \nu_j h_j(x). \tag{2}$$

$\lambda \in \mathbb{R}^n$ и $\nu \in \mathbb{R}^m$ мы будем называть двойственными переменными, в то время как $x \in \mathbb{R}^d$ — прямой.

Двойственная функция

Definition (Двойственная функция по Лагранжу)

Определим двойственную функцию по Лагранжу (или просто двойственную функцию) $g : \mathbb{R}^n \times \mathbb{R}^m \to \overline{\mathbb{R}}$ следующим образом:

$$g(\lambda, \nu) = \inf_{x \in \mathbb{R}^d} L(x, \lambda, \nu) = \inf_{x \in \mathbb{R}^d} \left(f_0(x) + \underbrace{\sum_{i=1}^{n} \lambda_i f_i(x) + \sum_{j=1}^{m} \nu_j h_j(x)}_{\text{линейна по } \lambda,\, \nu} \right). \tag{3}$$

- Если при (λ, ν) лагранжиан L является неограниченным снизу по переменной x, то значение $g(\lambda, \nu) = -\infty$.
- $g(\lambda, \nu)$ **всегда** является вогнутой по переменным (λ, ν).

Proposition

Пусть дано оптимальное значение задачи (1) p^* (может быть $-\infty$). Тогда, для любого $\lambda \succeq 0$ и любого ν выполняется

$$g(\lambda, \nu) \leq p^*. \tag{4}$$

Получили нижнюю оценку на оптимальное значение задачи (1).

$$g(\lambda, \nu) = \inf_x f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{j=1}^{m} \nu_j h_j(x) \leq$$

p^* достигается в x^* допустимая

$$\leq f_0(x^*) + \underbrace{\sum_{i=1}^{m} \lambda_i f_i(x_*)}_{\lambda_i \geq 0,\ f_i \leq 0} + \sum_{j=1}^{m} \nu_j \underbrace{h_j(x^*)}_{=0}$$

$$\leq f_0(x^*) = p^*$$

Двойственная задача

Нижняя оценка $g(\lambda, \nu)$ зависит напрямую от λ и ν. А какова **лучшая** оценка на p^* снизу?

может быть $+\infty$

$$
\begin{aligned}
d^* = \max_{\lambda,\nu}\, & g(\lambda, \nu), \quad \le p^* \\
\textbf{s.t. } & \lambda \succeq 0.
\end{aligned}
\tag{5}
$$

Такая задача называется **двойственной задачей** к задаче (1). Эта задача является задачей выпуклой оптимизации, так как максимизация вогнутой функции и линейные ограничения λ.

Сильная и слабая двойственности

Для оптимального значения двойственной задачи d^* всегда верно

$$d^* \le p^*.$$

Это свойство называется **слабой двойственностью**.
В частности, когда

$$d^* = p^*,$$

то выполняется свойство **сильной двойственности**.

Разрешимость и неограниченность

$d^* \leq p^*$ $\quad$ $p^* = -\infty$ $\quad$ $\min f_0(x)$

Proposition

Если прямая задача ***неограниченна снизу*** $(p^* = -\infty)$, *то двойственная задача* $g(\lambda, \nu) \equiv -\infty$.

Proposition

Если двойственная задача ***неограниченна сверху*** $(d^* = +\infty)$, *то прямая задача* ***не имеет допустимых прямых точек***.

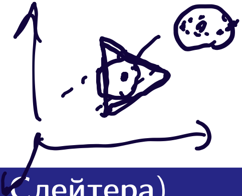
При выполнении **сильной двойственности** утверждения верны и в обратную сторону.

$\cancel{\exists} x \quad f_i(x) \leq 0 \quad h_j(x) = 0$

Условие Слейтера

Рассмотрим задачу с выпуклыми $f_0, \ldots, f_m$ и линейными равенствами:

$$\begin{aligned} \min \ & f_0(x) \\ \text{s.t.} \ & f_i(x) \leq 0, \ i = 1, \ldots, m \\ & Ax = b. \end{aligned} \tag{6}$$



Proposition (**Условие сильной выпуклости Слейтера**)

*Будем говорить, что для задачи (6) выполняется условие Слейтера, если существует допустимая $x_0 \in$ **relint D**, такой что*

$$f_i(x_0) < 0, \ i = 1, \ldots, m, \ Ax = b.$$

$Ax_0 = b$

$f_i(x_0) \leq 0$

Ослабленное условие: $f_i(x) < 0$ только у не аффинных f_i.

Theorem (**Теорема Слейтера**)

Если для задачи (6) выполняется условие Слейтера, то тогда выполняется свойство сильной двойственности.

Другие условия сильной выпуклости

$$\min_x \ f_0(x)$$
$$\textbf{s.t.} \ f_i(x) \le 0, \ i = 1, \dots, m,$$
$$h_j(x) = 0, \ j = 1, \dots, n.$$

f_0 - невып

1. Функции ограничений f_i и h_j являются аффинными.
2. Для точки минимума x^* градиенты всех ограничений равенств и всех *активных* ограничений неравенств (выполняется равенство нулю) линейно независимы.

Пример

Example (Решение СЛАУ минимальной нормы)

$$
\begin{aligned}
&\min_{x}\ x^T x\\
&\textbf{s.t. } Ax = b.
\end{aligned}
$$

$$\min_x \langle x, x\rangle = \|x\|^2 \qquad A \in \mathbb{R}^{m\times n}$$

$$Ax = b \to m(\,)$$

- $L(x,\lambda,\nu) = f_0(x) + \sum_{i=1}^{n} \lambda_i f_i(x) + \sum_{j=1}^{m} \nu_j h_j(x) =$

$$= \langle x, x\rangle + \nu^T(Ax-b) = \uparrow\; L\text{ - вып по } x$$

$m(\,)$

f_0 - вып, f_i - вып; h_j ограничение линейно

- $g(\nu) = \inf_{x\in\mathbb{R}^n} L(x,\nu)$

$$\partial L(x,\nu) = 2\langle x, \partial x\rangle + \nu^T A\, \partial x = 0$$

$$2x + A^T\nu = 0 \Leftrightarrow x = -\frac{A^T\nu}{2} \quad \forall \nu \qquad \left(x^* = -\frac{A^T}{2}\nu^*\right)$$

По критерию глоб min:

$$\langle x, x\rangle + \nu^T(Ax-b) = \frac{\nu^T A A^T \nu}{4} - \frac{\nu^T A A^T \nu}{2} - \nu^T b = -\frac{\nu^T A A^T \nu}{4} - \nu^T b$$

- $\max_\nu g(\lambda) = \max_\nu \left\{ -\frac{\nu^T A A^T \nu}{4} - \nu^T b \right\} : \partial$

сильная двойственность достигается - ограничения линейные

$$d^* = +\infty \Rightarrow \nexists x \quad Ax = b$$

$$AA^T \geq 0 \quad \text{Если } AA^T > 0 \Rightarrow \underline{\nu^* = -2(AA^T)^{-1} b}$$

Примеры

Example (Задача линейного программирования)

Составьте двойственную задачу для

$$
\begin{aligned}
\min\ & c^T x \\
\textbf{s.t.}\ & Ax = b, \\
& x \succeq 0,
\end{aligned}
$$

где $A \in \mathbb{R}^{n \times m}$, $b \in \mathbb{R}^n, c \in \mathbb{R}^m$.

Общий алгоритм

Если функция f_0 и ограничения неравенства f_i выпуклые, а ограничения равенства h_j линейные, то лагранжиан $L(x, \lambda, \nu)$ выпуклый по x. Можно применять условие глобального минимума. Но нужно внимательно смотреть на (λ, ν), где инфинум не достигается.

В случае других f_0, f_i, h_j, надо смотреть каждый инфинум отдельно.

Example (Задача разбиения)

Составьте двойственную задачу для

$$\min\ x^T W x$$
$$\textbf{s.t.}\ x_i^2 = 1, \quad i = 1, \dots n,$$

где $W \in \mathbb{S}^n_+$.

$$\min x^T W x \qquad W \succeq 0$$

$$x_i^2 = 1 \quad \forall i \in \overline{1,n}$$

$$L(x,\nu) = x^T W x + \sum_{i=1}^{n} \nu_i (x_i^2 - 1)$$

$$\underbrace{\sum_{i=1}^{n} \nu_i x_i^2}_{x^T \operatorname{diag}(\nu)\, x} \qquad \underbrace{-\sum_{i=1}^{n} \nu_i}_{=\langle \vec{1}, \nu \rangle}$$

$$= x^T (W + \operatorname{diag}(\nu)) x - \langle \vec{1}, \vec{\nu} \rangle$$

$$g(\nu) = \inf_x x^T [W + \operatorname{diag}(\nu)] x - \langle \vec{1}, \vec{\nu} \rangle$$

$$W + \operatorname{diag}(\nu) \succeq 0 \Rightarrow g(\nu) = -\langle \vec{1}, \nu \rangle$$

иначе $g(\nu) = -\infty$

$$\max_\nu g(\nu) = \boxed{\begin{aligned} &\max_\nu -\langle \vec{1}, \nu \rangle \\ &\text{s.t. } W + \operatorname{diag}(\nu) \succeq 0 \end{aligned}}$$

$$\nu_i \geq -\lambda_{min}(W).$$

Связь с сопряженными функциями

$$f_0^*(y) = \sup_{x \in \mathbb{R}^d} (y^T x - f_0(x)).$$

Для нахождения связи рассмотрим следующую задачу

$$\begin{aligned} \min_x \; & f_0(x) \\ \textbf{s.t. } & Ax \preceq b, \\ & Cx = d. \end{aligned}$$

$$g(\lambda, \nu) = -\lambda^T b - \nu^T d - f^*(-A^T \lambda - C^T \nu).$$

Для задач с линейными ограничениями, можно выписать двойственную задачу, зная лишь сопряженную функцию.

$$\min f_0(x)$$
$$Ax \le b \qquad Ax - b \le 0$$
$$Cx = d$$

$$L(x,\lambda,\nu) = f_0(x) + \lambda^T(Ax-b) + \nu^T(Cx-d)$$

$$\inf_x L(x,\lambda,\nu) = -\sup_x -L(x,\lambda,\nu) =$$

$$-\sup_x \{-f_0(x) - \lambda^T A x + \lambda^T b - \nu^T C x + \nu^T d\}$$

$$= -\sup_x \{\langle x, \underbrace{-A^T\lambda - C^T\nu}_{y}\rangle - f_0(x)\} + \lambda^T b + \nu^T d\}$$

$$= -f_0^*(-A^T\lambda - C^T\nu) - \lambda^T b - \nu^T d.$$

$$\max -f_0^*(-A^T\lambda - C^T\nu) - \lambda^T b - \nu^T d$$
$$\lambda \ge 0$$

$$\min_x \|x\| \quad \overset{\lambda=0}{\Longrightarrow} \quad \max_\nu -\nu^T d$$
$$Cx = d \qquad \|-C^T\nu\|_* \le 1$$

$$\min_x \sum_{i=1}^n x_i \log x_i \Longrightarrow \max_{\nu,\lambda} -\sum_{i=1}^n e^{(-a_i^T\lambda - \nu - 1)} - \lambda^T b - \nu$$
$$Ax \le b \qquad \lambda \ge 0$$
$$1^T x = 1 \qquad \nu \in R$$

Примеры на сопряженные функции

Example (Решение СЛАУ с наименьшей нормой общего вида)

Составьте двойственную задачу для

$$\min_{x} \|x\|$$
$$\text{s.t. } \cancel{Ax = b},$$
$$Cx = d$$

где $\|\cdot\|$ - любая норма.

Пример на сопряженные функции

Example (Максимизация энтропии)

Составьте двойственную задачу для

$$\begin{aligned} \min \ & \sum_{i=1}^{d} x_i \log x_i \\ \textbf{s.t. } & Ax \preceq b, \\ & \mathbf{1}^T x = 1. \end{aligned}$$

Example (Кусочно-линейная оптимизация)

Составьте двойственную задачу для

$$\min \max_{i=1,\dots,m} (a_i^T x + b_i).$$

На практике

На практике методы работают следующим образом - происходит инициализация x^0, λ^0, ν^0, и итеративным алгоритмом меняются сразу как прямые, так и двойственные методы. В качестве критерия остановы берут $f(x^k) - g(\lambda^k, \nu^k) \leq \epsilon$.
Поэтому, когда будет исследоваться график невязки между прямой и двойственной функцией, то станет понятно, выполняется сильная двойственность, или же нет: $f(x^k) - g(\lambda^k, \nu^k)$ должно стремиться к $p^* - d^*$ – так называемому **оптимальному двойственному зазору**, и если выполняется свойство сильной двойственности, то этот зазор на графике будет стремиться к нулю.