Методы оптимизации. Семинар 9. Двойственная задача

Корнилов Никита Максимович

Московский физико-технический институт

30 октября 2025г

Задача оптимизации с ограничениями

Постановка прямой задачи оптимизации стандартной формы:

$$\begin{aligned} \min_{x \in \mathbb{R}^d} \; & f_0(x) \\ \textbf{s.t. } & f_i(x) \leq 0, \; i = 1, \ldots, n, \\ & h_j(x) = 0, \; j = 1, \ldots, m, \end{aligned} \tag{1}$$

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

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:

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

- 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).

Нижняя оценка

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).

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

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

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

$$d^* \leq p^*.$$

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

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

то выполняется свойство **сильной двойственности**.

Разрешимость и неограниченность

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.

Условие Слейтера

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

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

Proposition (Условие сильной выпуклости Слейтера)

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

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

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

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

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)

$$\min_{x\in\mathbb{R}^d} \; x^T x$$
$$\textbf{s.t. } Ax = b,$$

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

Примеры

Example (Задача линейного программирования)

Составьте двойственную задачу для

$$\begin{aligned}\min_{x \in \mathbb{R}^n} \ & c^T x \\ \textbf{s.t.} \ & Ax = b, \\ & x \succeq 0,\end{aligned}$$

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

Общий алгоритм

1. Составляем лагранжиан (обратите внимание на знак $f_i(x) \leq 0$):

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

2. Ищем двойственную функцию $g(\lambda, \nu) = \inf_{x \in \mathbb{R}^d} L(x, \lambda, \nu)$. Если функция f_0 и неравенства f_i выпуклые, а равенства h_j линейные, то лагранжиан $L(x, \lambda, \nu)$ выпуклый по x. Можно применять условие глобального минимума $\nabla_x L(x, \lambda, \nu) = 0$. Но осторожно с (λ, ν), где инфинум не достигается.

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

3. Составляем двойственную задачу(**помним про** $\lambda \succeq 0$):

$$\max_{\lambda, \nu} g(\lambda, \nu),$$

$$\textbf{s.t. } \lambda \succeq 0.$$

Примеры

Example (Задача разбиения)

Составьте двойственную задачу для

$$\begin{aligned} \min_{x \in \mathbb{R}^m} \; & x^T W x \\ \textbf{s.t.} \; & x_j^2 = 1, \;\; j = 1, \dots m, \end{aligned}$$

где $W \in \mathbb{S}^m_+$.

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

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

$$g(\lambda, \nu) = -\lambda^T b - \nu^T d - f_0^*(-A^T\lambda - C^T\nu), \quad \lambda \succeq 0.$$

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

Example (Решение СЛАУ с наименьшей нормой общего вида)

Составьте двойственную задачу для

$$\min_{x \in \mathbb{R}^d} \|x\|$$
$$\textbf{s.t. } Cx = d,$$

где $\|\cdot\|$ - любая норма, $C \in \mathbb{R}^{m \times d}, d \in \mathbb{R}^m$.

Example (Максимизация энтропии)

Составьте двойственную задачу для

$$\min_{x \in \mathbb{R}^d} \sum_{i=1}^{d} x_i \log x_i$$
$$\textbf{s.t. } Ax \preceq b,$$
$$\mathbf{1}^T x = 1,$$

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

Example (Кусочно-линейная оптимизация)

Составьте двойственную задачу для

$$\min_{x \in \mathbb{R}^d} \max_{i=1,\ldots,m} (a_i^T x + b_i),$$

где $a_i \in \mathbb{R}^d$, $b_i \in \mathbb{R}, \forall i \in \overline{1, m}$.

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 останова 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.