Методы оптимизации. Семинар 9. Двойственная задача

Корнилов Никита Максимович

Московский физико-технический институт

29 октября 2024г

Начнем с рассмотрения постановки задачи оптимизации стандартной формы:

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

с прямой переменной $x \in \mathbb{R}^d$. Также мы скажем, что множество допустимых точек $\mathbf{D} = \cap_{i=0}^m \mathbf{dom}\ f_i \cap_{j=1}^n \mathbf{dom}\ h_j$ является непустым множеством.

Лагранжиан

Definition (Лагранжиан)

Лагранжиан L относительно задачи оптимизации (1) задается следующим образом:

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

Definition (Лагранжиан)

Лагранжиан L относительно задачи оптимизации (1) задается следующим образом:

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

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

Двойственная функция

Definition (Двойственная функция по Лагранжу)

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

$$g(\lambda, \nu) = \inf_{x \in \mathbf{D}} L(x, \lambda, \nu) = \inf_{x \in \mathbf{D}} \left(f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{j=1}^{n} \nu_i h_i(x) \right). \quad (3)$$

Двойственная функция

Definition (Двойственная функция по Лагранжу)

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

$$g(\lambda, \nu) = \inf_{x \in \mathbf{D}} L(x, \lambda, \nu) = \inf_{x \in \mathbf{D}} \left(f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{j=1}^{n} \nu_i h_i(x) \right). \quad (3)$$

- 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

Обозначим оптимальное значение задачи (1) как p^*. 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).

Двойственная задача

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

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

Двойственная задача

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

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

Такая задача называется **двойственной задачей** к задаче (1). Эта задача является задачей выпуклой оптимизации, так как максимизация вогнутой функции, а также функции, ограничивающие λ, тоже являются выпуклыми.

Классические примеры

Example (Задача линейного программирования)

Рассмотрим следующую задачу оптимизации

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

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

Это классика, это решать надо

Example (Задача разбиения)

Рассмотрим следующую задачу оптимизации

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

где $W \in \mathbb{S}^n_+$.

Связь с сопряженными функциями

$$f^*(y) = \sup_{x \in \mathbf{dom}\ f} (y^T x - f(x)).$$

Для нахождения связи рассмотрим следующую задачу

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

Связь с сопряженными функциями

$$f^*(y) = \sup_{x \in \mathbf{dom}\, f} (y^T x - f(x)).$$

Для нахождения связи рассмотрим следующую задачу

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

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

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

Example (Решение СЛАУ с наименьшей нормой общего вида)

Рассмотрим задачу

$$\min \|x\|$$
$$\textbf{s.t. } Ax = b,$$

где $\|\cdot\|$ - любая норма.

Example (Максимизация энтропии)

Рассмотрим следующую задачу

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

где **dom** $f = \mathbb{R}^d_{++}$.

Example (Кусочно-линейная оптимизация)

Рассмотрим следующую задачу оптимизации

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

Сильная и слабая двойственности

Обозначим оптимальное значение двойственной задачи относительно начальной как d^*. Поскольку $g(\lambda, \nu) \leq p^*$, то в силу произвольности выбора двойственных переменных выполняется

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

Заметим, что такое неравенство выполняется всегда. Такое свойство называется **слабой двойственностью**. В частности, когда

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

то будем говорить, что выполняется свойство **сильной двойственности**.

Разрешимость и неограниченность

Proposition

Если прямая задача ***неограниченна снизу*** *($p^* = -\infty$), то двойственная задача* ***не имеет допустимых двойственных точек****.*

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 функциями $f_0, \ldots, f_m$:

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

Proposition (Условие Слейтера)

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

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

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

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

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

Example (Решение СЛАУ минимальной нормы)

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

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.

В следующих сериях

- Условия оптимальности Каруша-Куна-Таккера.