Методы оптимизации. Семинар 10. Оптимальность. Условия ККТ.

Корнилов Никита Максимович

Московский физико-технический институт

6 ноября 2025г

Прямая и двойственная задачи

Прямая задача

$$p^* = \min_{x \in \mathbb{R}^d} \; f_0(x)$$
$$\textbf{s.t.} \;\; f_i(x) \le 0, \;\; i = 1, \dots, n, \tag{1}$$
$$h_j(x) = 0, \;\; j = 1, \dots, m.$$

Двойственная функция

$$g(\lambda, \nu) = \inf_{x \in \mathbb{R}^d} \underbrace{\left\{ f_0(x) + \sum_{i=1}^{n} \lambda_i f_i(x) + \sum_{j=1}^{m} \nu_j h_i(x) \right\}}_{= L(x, \lambda, \nu)}. \tag{2}$$

Двойственная задача

$$d^* = \max_{\lambda, \nu} \; g(\lambda, \nu)$$
$$\textbf{s.t.} \;\; \lambda \succeq 0. \tag{3}$$

Условия Слейтера для сильной двойственности

Достаточным условием сильной двойственности $p^* = d^*$, например, является *ослабленное условие Слейтера*:

1. Функции f_0 и f_i являются выпуклыми, а h_j являются аффинными.
2. Существует такая допустимая точка $\overline{x}$, что все *неафинные* условия неравенства выполняются строго $f_i(\overline{x}) < 0$.

Условие дополняющей нежёсткости

Предположим, что выполняется сильная двойственность. Также x^* - прямая переменная, доставляющая оптимум задачи (1), а (λ^*, ν^*) - двойственная переменная, доставляющая оптимум задачи (3).

Условия дополняющей нежёсткости

$$
\begin{aligned}
f_0(x^*) &= g(\lambda^*, \nu^*) \\
&= \inf_x \left(f_0(x) + \sum_{i=1}^n \lambda_i^* f_i(x) + \sum_{j=1}^m \nu_j^* h_j(x) \right) \\
&\leq f_0(x^*) + \sum_{i=1}^n \underbrace{\lambda_i^* f_i(x^*)}_{\leq 0} + \sum_{j=1}^m \nu_j^* \underbrace{h_j(x^*)}_{=0} \leq f_0(x^*).
\end{aligned}
$$

Поэтому мы получаем для $f_i(x^*) \leq 0$

$$\lambda_i^* f_i(x^*) = 0, \;\; i = 1, \ldots, n.$$

Или эквивалентно

$$
\begin{aligned}
\lambda_i^* > 0 \Rightarrow f_i(x^*) = 0, \\
f_i(x^*) < 0 \Rightarrow \lambda_i^* = 0.
\end{aligned}
$$

Условие ККТ

Дополнительно предположим, что $f_0, f_1, \dots f_n, h_1, \dots h_m$ дифференцируемы в x^*. Так как x^* глобально минимизирует $L(x, \lambda^*, \nu^*)$, градиент L по x в точке x^* должен быть равен нулю

$$\nabla f_0(x^*) + \sum_{i=1}^{n} \lambda_i^* \nabla f_i(x^*) + \sum_{j=1}^{m} \nu_j^* \nabla h_j(x^*) = 0,$$

Если $f_0, f_1, \dots f_n, h_1, \dots h_m$ только субдифференцируемы в x^*, то

$$\partial f_0(x^*) + \sum_{i=1}^{n} \lambda_i^* \partial f_i(x^*) + \sum_{j=1}^{m} \nu_j^* \partial h_j(x^*) \ni 0.$$

Каруша-Куна-Такера:

Допустимость по x $\quad f_i(x^*) \leq 0, \; i = 1, \dots n,$

Допустимость по x $\quad h_j(x^*) = 0, \; j = 1, \dots m,$

Неотрицательность $\quad \lambda_i^* \geq 0, \; i = 1, \dots n,$

Дополняющая нежесткость $\quad \lambda_i^* f_i(x^*) = 0, \; i = 1, \dots n,$

Стационарность $$\nabla f_0(x^*) + \sum_{i=1}^{n} \lambda_i^* \nabla f_i(x^*) + \sum_{j=1}^{m} \nu_j^* \nabla h_j(x^*) = 0. \quad (4)$$

Необходимость и достаточность

Необходимые условия: для оптимального набора прямых переменных x^* и двойственных переменных (λ^*, ν^*) при сильной двойственности следуют условия ККТ (4).

Можно требовать не сильной двойственности, а других условий регулярности, см. слайд ниже.

Достаточные условия: Когда f_0, f_i выпуклые, а h_j аффинные: для $x^*, (\lambda^*, \nu^*)$, которые удовлетворяют условиям ККТ (4), выполняется следующее - эти точки доставляют оптимум прямой и двойственной задачи соответственно, и выполняется сильная двойственность. $g(\lambda^*, \nu^*) = L(x^*, \lambda^*, \nu^*) = f_0(x^*)$.

Аналитическое решение задач

С помощью достаточного условия ККТ можно находить решение прямой и двойственной задач аналитически

Example

Найдите глобальный минимум задачи

$$\begin{aligned} \min \ & \frac{1}{2}x^T Px + q^T x + r \\ \textbf{s.t.} \ & Ax = b, \end{aligned}$$

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

Проекция на l_2 шар

Example

Найдите глобальный минимум задачи

$$\min_x \|x - s\|_2^2$$
$$\text{s.t. } \|x\|_2^2 \leq 1.$$

Несколько условий неравенств

Example

Найдите глобальный минимум задачи

$$\begin{aligned} \min_{x,y\in\mathbb{R}} \ & x + 3y \\ \text{s.t. } & x - y \geq 0, \\ & (x-1)^2 + (y-1)^2 \leq 9. \end{aligned}$$

Общий алгоритм

1. Выписать Лагранжиан $L(x, \lambda, \nu) = f_0(x) + \sum\limits_{i=1}^{n} \lambda_i f_i(x) + \sum\limits_{j=1}^{m} \nu_j h_i(x)$.
2. Выписать условия ККТ (4).
3. Рассмотреть различные варианты $\lambda_i = 0$ и $\lambda_i > 0$ и выяснить, приводят ли они к противоречиям или нет.

Example (Water-filling)

Найдите глобальный минимум задачи

$$\begin{aligned} \min_{x\in\mathbb{R}^d} \; & -\sum_{i=1}^{d} \log\left(\alpha_i + x_i\right) \\ \textbf{s.t. } & x \succeq 0, \\ & \mathbf{1}^T x = 1, \end{aligned}$$

где $\alpha_i > 0$.

Решение прямой задачи через двойственную

Рассмотрим решение прямой задачи x^* и двойственной задачи (λ^*, ν^*). При сильной двойственности верно

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

Этот инфимум достигается в точке x^*. Для строго выпуклого лагранжиана выпуклой задачи это единственный возможный минимум и найти его можно через критерий

$$\nabla f_0(x^*) + \sum_{i=1}^m \lambda_i^* \nabla f_i(x^*) + \sum_{j=1}^n \nu_j^* \nabla h_j(x^*) = 0.$$

Если минимум лагранжиана не достигается, то и в прямой задаче минимум не достигается.

Решение прямой задачи через двойственную

Алгоритм:

1. Составить лагранжиан $L(x, \lambda, \nu)$ и двойственную функцию $g(\lambda, \nu)$.
2. Найти оптимальные перемеренные (λ^*, ν^*) для двойственной задачи $\max_{\nu, \lambda \geq 0} g(\lambda, \nu)$.
3. Проверить сильную двойственность, например, по условию Слейтера.
4. Найти решения $\inf_x L(x, \lambda^*, \nu^*)$ и проверить, какое подходит под условия f_i, h_j и доставляет глобальный минимум прямой задаче.

Примеры на решение

Example (Максимизация энтропии)

Найдите глобальный минимум задачи

$$\min_{x\in\mathbb{R}^d} \sum_{i=1}^{d} x_i \log x_i$$

$$\textbf{s.t. } Ax \preceq b,$$

$$\mathbf{1}^T x = 1.$$

По ослабленному условию Слейтера сильная двойственность достигается.

Примеры на решение

Example (Минимизация сепарабельной функции)

Найдите глобальный минимум задачи

$$\min_x \sum_{i=1}^{d} f_i(x_i)$$
$$\textbf{s.t. } a^T x = b,$$

где f_i - строго выпуклые и дифференцируемые функции.

По ослабленному условию Слейтера сильная двойственность достигается.

Решение

Выпишем лагранжиан:

$$L(x,\nu) = \sum_{i=1}^{d} f_i(x_i) + \nu(a^T x - b) = -b\nu + \sum_{i=1}^{d}(f_i(x_i) + \nu a_i x_i),$$

который тоже является сепарабельным по компонентам вектора x. Тогда двойственная функция

$$\begin{aligned} g(\nu) &= -b\nu + \inf_x \left(\sum_{i=1}^{d}(f_i(x_i) + \nu a_i x_i) \right) \\ &= -b\nu + \sum_{i=1}^{d} \inf_{x_i}(f_i(x_i) + \nu a_i x_i) \\ &= -b\nu - \sum_{i=1}^{d} f_i^*(-\nu a_i), \end{aligned}$$

где f_i^* - сопряженные функция.

Решение

Тогда двойственная задача будет иметь следующий вид:

$$\max_{\nu} \; -b\nu - \sum_{i=1}^{d} f_i^*(-\nu a_i),$$

где ν – скаляр.

Для поиска оптимального значения одномерной задачи можно пользоваться уже известными вам методами, например, методом дихотомии или золотого сечения. В силу показанного ранее минимум прямой задачи совпадает с минимумом $\min_x L(x, \nu^*)$. Тогда, для поиска x^* можно взять градиент лагранжиана в ν^* по x и приравнять его к нулю: $\nabla_x L(x^*, \nu^*) = 0$, то есть решать уравнения $f_i'(x_i^*) = -\nu^* a_i$.

Theorem

Пусть x^* является локальным минимумом прямой задачи и пусть выполняется хотя бы одно из ***условий регулярности****. Тогда, если функции f_0, f_i, h_j дифференцируемы в точке x^*, то существуют такие двойственные переменные (λ^*, ν^*), что выполняются условия ККТ.*

В общем случае достаточно требовать не сильной двойственности для необходимых условий ККТ, а более слабые условия регулярности.

Условия регулярности

1. Ослабленное условие Слейтера.
2. Функции ограничений f_i и h_j являются аффинными, f_0 — любая.
3. Для точки x^* градиенты всех ограничений равенств и всех *активных* ограничений неравенств (выполняется равенство нулю) линейно независимы.
4. Ещё больше условий регулярности *по ссылке*.

Second-Order Sufficient Condition (SOSC)

Определим для набора переменных (x, λ, ν) следующие множества из активных неравенств:

$$
\begin{aligned}
I^0(x) &= \{i : f_i(x) = 0, \lambda_i = 0\}, \\
I^+(x) &= \{i : f_i(x) = 0, \lambda_i > 0\}.
\end{aligned}
$$

Definition

Достаточное условие второго порядка (SOSC) выполнено для набора переменных (x, λ, ν), если для любого вектора $z \neq 0$, такого что:

$$
\begin{aligned}
z^T \nabla_x f_i(x) &= 0, \quad i \in I^+(x), \\
z^T \nabla_x f_i(x) &\leq 0, \quad i \in I^0(x), \\
z^T \nabla_x h_j(x) &= 0, \quad j = 1, \ldots, m,
\end{aligned}
$$

верно что

$$
z^T \nabla^2_{xx} L(x, \lambda, \nu) z > 0. \tag{5}
$$

SOSC Достаточные условия ККТ

Theorem

Пусть функции f_0, f_i, h_j *являются дважды непрерывно дифференцируемыми. Тогда, если для набора переменных* (x^*, λ^*, ν^*) *выполнены все условия ККТ и SOSC, то* x^* *является точкой локального минимума прямой задачи.*

Example

Найдите все точки ККТ и проверьте условия SOSC для задачи

$$\begin{aligned}
&\min_x \ -x \\
&\text{s.t. } x^2 + y^2 \leq 1, \\
&\qquad (x-1)^3 - y \leq 0.
\end{aligned}$$