

Методы оптимизации. Семинар 10. Оптимальность. Условия ККТ.

Корнилов Никита Максимович

Московский физико-технический институт

6 ноября 2025г

Прямая и двойственная задачи

Прямая задача

$$\begin{aligned} p^* &= \min_{x \in \mathbb{R}^d} f_0(x) \\ \text{s.t. } f_i(x) &\leq 0, \quad i = 1, \dots, n, \\ h_j(x) &= 0, \quad j = 1, \dots, m. \end{aligned} \quad (1)$$

Двойственная функция

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

Двойственная задача

$$\begin{aligned} d^* &= \max_{\lambda, \nu} g(\lambda, \nu) \\ \text{s.t. } \lambda &\succeq 0. \end{aligned} \quad (3)$$

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 \bar{x} , что все *неаффинные* условия неравенства выполняются строго $f_i(\bar{x}) < 0$.

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).

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

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

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

Или эквивалентно

$$\lambda_i^* > 0 \Rightarrow f_i(x^*) = 0,$$

$$f_i(x^*) < 0 \Rightarrow \lambda_i^* = 0.$$

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

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

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

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

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

Необходимые условия ККТ глобального минимума

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

Example (Сильная двойственность существенна!)

$$\begin{aligned} \min_{x \in \mathbb{R}} \quad & x^3 + x \\ \text{s.t.} \quad & x^2 \leq 0. \end{aligned} \tag{5}$$

Для (5) единственная допустимая точка $x^* = 0$ оптимальна с $p^* = 0$. При этом $g(\lambda) = \inf_{x \in \mathbb{R}} \{x^3 + x + \lambda x^2\} \equiv -\infty$, а в двойственной задаче $d^* = \max_{\lambda \geq 0} g(\lambda) = -\infty$ сильная двойственность не достигается и любая $\lambda^* \geq 0$ является оптимальной.

Однако условие стационарности не выполнено ни для какого $\lambda^* \geq 0$

$$\nabla_x L(x^*, \lambda^*) = 3(x^*)^2 + 1 + 2\lambda^* x^* = 1 \neq 0.$$

Достаточные условия ККТ глобального минимума

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.

$$g(\lambda^*, \nu^*) = L(x^*, \lambda^*, \nu^*) = f_0(x^*).$$

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

Example

Найдите глобальный минимум задачи

$$\begin{aligned} \min_{x \in \mathbb{R}^d} \quad & \frac{1}{2} x^T P x + q^T x + r \\ \text{s.t.} \quad & Ax = b, \end{aligned}$$

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

Example

Найдите глобальный минимум задачи

$$\begin{aligned} \min_{x \in \mathbb{R}^d} & \|x - s\|_2^2 \\ \text{s.t. } & \|x\|_2^2 \leq 1. \end{aligned}$$

Example

Найдите глобальный минимум задачи

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

- 1 Проверить, что для задачи выполнено хотя бы одно условие регулярности. Тогда глобальный минимум будет среди точек ККТ. Для выпуклой задачи это условие Слейтера или существование любой точки ККТ.
- 2 Выписать Лагранжиан
$$L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^n \lambda_i f_i(x) + \sum_{j=1}^m \nu_j h_j(x).$$
- 3 Выписать условия ККТ (4).
- 4 Рассмотреть различные варианты $\lambda_i = 0$ и $\lambda_i > 0$ и выяснить, приводят ли они к противоречиям или нет.
- 5 Выбрать из точек ККТ глобальный минимум. Для выпуклой задачи это любая точка ККТ.

Example (Water-filling)

Найдите глобальный минимум задачи

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

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.

Алгоритм:

- 1 Составить лагранжиан $L(x, \lambda, \nu)$ и двойственную функцию $g(\lambda, \nu)$.
- 2 Найти оптимальные переменные (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.

Example (Максимизация энтропии)

Найдите глобальный минимум задачи

$$\begin{aligned} \min_{x \in \mathbb{R}^d} \quad & \sum_{i=1}^d x_i \log x_i \\ \text{s.t.} \quad & Ax \preceq b, \\ & \mathbf{1}^T x = 1. \end{aligned}$$

По ослабленному условию Слейтера сильная двойственность достигается.

Example (Минимизация сепарабельной функции)

Найдите глобальный минимум задачи

$$\begin{aligned} \min_x \quad & \sum_{i=1}^d f_i(x_i) \\ \text{s.t.} \quad & a^T x = b, \end{aligned}$$

где f_i - строго выпуклые и дифференцируемые функции.

По ослабленному условию Слейтера сильная двойственность достигается.

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.

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 или глобальным минимумом прямой задачи и пусть выполняется хотя бы одно из **условий регулярности**. Тогда существуют такие двойственные переменные (λ^*, ν^*) , такие что выполняются условия ККТ.*

В общем случае достаточно требовать не сильной двойственности для необходимых условий ККТ, а более слабые условия регулярности. При этом сильная двойственность может и не достигаться.

- 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*.

Условия регулярности существенны!

Для примера (5) не выполнено ни одно условие регулярности (задача невыпуклая, ограничение нелинейное, градиент ограничения в x^* нулевой и линейно зависимый). Поэтому для x^* не существует λ^* , т.ч. выполнялись ККТ.

Example

$$\begin{aligned} \min_{x \in \mathbb{R}} \quad & x^3 \\ \text{s.t.} \quad & x \geq 0. \end{aligned} \tag{6}$$

Для (6) глобальный минимум это $x^* = 0$ с $p^* = 0$.

При этом $g(\lambda) = \inf_{x \in \mathbb{R}} \{x^3 - \lambda x\} \equiv -\infty$, а в двойственной задаче $d^* = \max_{\lambda \geq 0} g(\lambda) = -\infty$ сильная двойственность не достигается и любая $\lambda^* \geq 0$ является оптимальной.

Здесь выполнено условие регулярности по линейности ограничений, и поэтому $\exists \lambda^* = 0$, т.ч. выполнено условие стационарности ККТ

$$\nabla_x L(x^*, \lambda^*) = 3(x^*)^2 - \lambda^* = 0 \Rightarrow \lambda^* = 0.$$

При других $\lambda^* > 0$ условия ККТ не выполняются.

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:

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

верно что

$$z^T \nabla_{xx}^2 L(x, \lambda, \nu) z > 0. \quad (7)$$

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

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