Методы оптимизации. Семинар 10. Оптимальность. Условия ККТ.

Корнилов Никита Максимович

Московский физико-технический институт

6 ноября 2025г

Прямая задача

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

Двойственная функция

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

Двойственная задача

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

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 $f_i(\overline{x}) < 0$.

f_i, h_j - линейные, f_0 - вып, $\exists$ допустимая точка

$\Rightarrow$ СД

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).

$$f_0(x^*) \overset{\text{сд}}{=} g(\lambda^*, \nu^*)$$

$L(x^*, \lambda^*, \nu^*)$

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

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

$= 0$ $\quad$ ⇒ допустимости

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

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

Или эквивалентно

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

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

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 x^*

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

неотрицательность

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

нежесткость

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

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

Стационарность

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.

без условий

лок min $\Rightarrow$ $\nabla f(x^*) = 0$

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.

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

f - вып, $\nabla f = 0 \Rightarrow$ лок min

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

bem

$\nabla L(x^*,\lambda^*,\nu^*) = 0 \Rightarrow$ und min b $x^* \Rightarrow = g(\lambda^*,\nu^*)$

$$L(x^*,\lambda^*,\nu^*) = f_0(x^*) + \underbrace{\sum_{i=1}^{n} \lambda_i^* f_i(x^*)}_{= 0} + \underbrace{\sum_{j=1}^{m} h_j(x^*)\nu^*}_{= 0}$$

RH

$$= f_0(x^*) = g(\lambda^*,\nu^*) \Rightarrow \text{CH}$$

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

Example

Найдите глобальный минимум задачи

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

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

$\min \frac{x^T P x}{2} + q^T x + r$ – выпукл $\Rightarrow$ СД по Слейтеру

s.t. $Ax = b$ – лин

- $L(x, \lambda, \nu) = \frac{x^T P x}{2} + q^T x + \cancel{r} + \nu^T (Ax - b)$

- ККТ: $Ax = b$ – доп по x $\Rightarrow x^*, \nu^*$

$\nabla_x L(x, \lambda, \nu) = Px + q + A^T \nu = 0$

$$\begin{array}{cc} & \begin{array}{cc} d & m \end{array} \\ \begin{array}{c} m \\ d \end{array} & \begin{pmatrix} A & 0 \\ P & A^T \end{pmatrix} \end{array} \begin{pmatrix} x \\ \nu \end{pmatrix} = \begin{pmatrix} b \\ -q \end{pmatrix}$$

квадратная (невырожденная $\Rightarrow$ решаем СЛУ)

Проекция на l_2 шар

Example

Найдите глобальный минимум задачи

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

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

вып условия $\to$ СД

- $L(x,\lambda) = \|x-s\|_2^2 + \lambda(\|x\|_2^2 - 1)$
- ККТ 1) $\|x\|_2^2 \le 1$ – доп по x

2) $\lambda \ge 0$ – доп по λ

3) $\lambda(\|x\|_2^2 - 1) = 0$ – АН

4) $\nabla_x L(x,\lambda) = 0 \Rightarrow 2(x-s) + 2\lambda x = 0$

$$\Downarrow$$

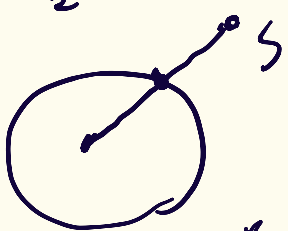
$$x = \frac{s}{1+\lambda}$$

- $\lambda > 0 \Rightarrow$ АН $\Rightarrow \|x\|_2^2 = 1 \qquad \frac{\|s\|_2^2}{(1+\lambda)^2} = 1$

$\boxed{(1+\lambda) = \frac{1}{\|s\|}} \Rightarrow \quad x = \frac{s}{(1+\lambda)} = \frac{s}{\|s\|}$

$\|s\|_2 \le 1 \Rightarrow x = 1 - \frac{1}{\|s\|} \le 0 \Rightarrow$ не может по 2)

$\Rightarrow \|s\|_2 > 1 \Rightarrow x^* = \frac{s}{\|s\|_2}$



- $\lambda = 0 \Rightarrow x^* = s \Rightarrow \|x\|_2^2 = \|s\|_2^2 \le 1.$

Остальные выполнены

$$x^* = \begin{cases} s, & \|s\|_2 \le 1 \\ \frac{s}{\|s\|_2}, & \|s\|_2 > 1 \end{cases}$$

Несколько условий неравенств

Example

Найдите глобальный минимум задачи

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

$\min x+3y$

$x-y \geq 0 \iff y-x \leq 0$

$(x-1)^2+(y-1)^2 \leq 9$

$L(x,\lambda_1,\lambda_2) = x+3y + \lambda_1(y-x) + \lambda_2((x-1)^2+(y-1)^2-9)$

ККТ

1) $x-y \geq 0$

2) $(x-1)^2+(y-1)^2 \leq 9$

3) $\lambda_1, \lambda_2 \geq 0$

4) $\lambda_1(y-x)=0$

5) $\lambda_2((x-1)^2+(y-1)^2-9)=0$

6) $\nabla_x L=0 \quad 1-\lambda_1+2\lambda_2(x-1)=0$

7) $\nabla_y L=0 \quad 3+\lambda_1+2\lambda_2(y-1)=0$

• $\lambda_2=0$ 6) и 7) $\iff \begin{matrix} 1-\lambda_1=0 \\ 3+\lambda_1=0 \end{matrix} \Rightarrow$ противоречие

• $\lambda_2>0, \lambda_1=0$

5) $\Rightarrow (x-1)^2+(y-1)^2=9$

6) $\Rightarrow (x-1) = -\frac{1}{2\lambda_2}$

7) $\Rightarrow (y-1) = -\frac{3}{2\lambda_2}$

$\Rightarrow \frac{1}{4\lambda_2^2}+\frac{9}{4\lambda_2^2}=9$

$\frac{10}{\lambda_2^2}=36 \Rightarrow \lambda_2=\frac{\sqrt{10}}{6} \geq 0$

5), 6), 7) - выполнено 3), 2)

$x = 1-\frac{1}{2\lambda_2} = 1-\frac{3}{\sqrt{10}} > 0$

$y = 1-\frac{3}{2\lambda_2} = 1-\frac{9}{\sqrt{10}}$

$x-y>0$

• $\lambda_2>0, \lambda_1>0$ $\quad y=x$, $(x-1)^2+(y-1)^2=9 \Rightarrow (x-1)^2=\frac{9}{2}$

$x=1\pm\frac{3}{\sqrt{2}}$

$$1 - \lambda_1 + 2\lambda_2(x-1) = 0$$
$$3 + \lambda_1 + 2\lambda_2(y-1) = 0$$

вычитаем

$$2 + 2\lambda_1 = 0 \Rightarrow \lambda_1 = -1 < 0 \Rightarrow$$ не подходит

противоречие

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 $\alpha_i > 0$.

$$-\min \sum_{i=1} \log(\alpha_i + x_i)$$

$1^T x = 1 \quad \nu - R$

$x \geq 0 \quad - \lambda - R^d \quad -x \leq 0$

$$L(x, \lambda, \nu) = -\sum_{i=1} \log(\alpha_i + x_i) + \nu(1^T x - 1) - \lambda^T x$$

ККТ. 1) $1^T x = 1$

2) $x \geq 0$

3) $\lambda \geq 0$

4) $\lambda_i x_i = 0 \quad \forall i$

5) $\frac{\partial L}{\partial x_i} = \frac{-1}{\alpha_i + x_i} + \nu - \lambda_i = 0$

$$\lambda_i = \nu - \frac{1}{\alpha_i + x_i} \geq 0$$

• $\nu < \frac{1}{\alpha_i}$ Может быть $x_i = 0 \Rightarrow \lambda_i = \nu - \frac{1}{\alpha_i} < 0$ ✗

$\alpha_i < \frac{1}{\nu}$

$x_i > 0 \overset{4)}{\Rightarrow} \lambda_i = 0 = \nu - \frac{1}{\alpha_i + x_i} = 0$

$$\nu = \frac{1}{\alpha_i + x_i} \Rightarrow x_i = \frac{1}{\nu} - \alpha_i > 0$$

• $\nu \geq \frac{1}{\alpha_i}$ Может ли $x_i > 0 \overset{4)}{\Rightarrow} \lambda_i = 0 \Rightarrow \nu = \frac{1}{\alpha_i + x_i}$

$x_i > 0$

$\frac{1}{\alpha_i}$

$x_i = 0$

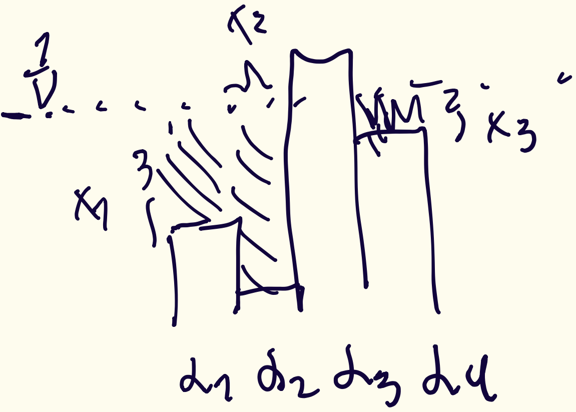
$$\lambda_i = \nu - \frac{1}{\alpha_i} \geq 0$$

$$x_i = \begin{cases} 0, & \nu \geq \frac{1}{\alpha_i} \\ \frac{1}{\nu} - \alpha_i, & \nu < \frac{1}{\alpha_i} \end{cases} = \max\{0, \frac{1}{\nu} - \alpha_i\}$$

Выразили $x_i, \lambda_i \geq 0$

$\nu, 0$

$$\sum_{i=1} x_i = 1 \Rightarrow \sum_{i=1} \max[0, \frac{1}{\nu} - \alpha_i] = 0$$



Ищем $\frac{1}{V}$

$\alpha_i > \frac{1}{V}$

$\alpha_i + x_i = \frac{1}{V}$

$\sum x_i = 1$

Решение прямой задачи через двойственную

Рассмотрим решение двойственной задачи (λ^*, ν^*). При сильной двойственности

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

Если у этой задаче единственный минимум (для выпуклой задачи это верно, когда лагранжиан строго выпуклый), то он обязательно достигается в точке x^* глобального минимума прямой задачи.

λ^*, ν^* – найти $\Leftrightarrow$ $x^* = \arg\min_x L(x, \lambda^*, \nu^*)$

Решение прямой задачи через двойственную

Для строго выпуклого лагранжиана смотрим условие:

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

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

Примеры на решение

Example (Максимизация энтропии)

Найдите глобальный минимум задачи

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

Примеры на решение

Example (Минимизация сепарабельной функции)

Найдите глобальный минимум задачи

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

где f_i - строго выпуклые и дифференцируемые функции.

По 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

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

$f^*(y) = \sup_x \{\langle x, y\rangle - f(x)\}$

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 $f_i'(x_i^*) = -\nu^* a_i$.

Необходимые условия ККТ локального минимума

loc min => $\partial E(x) = 0$

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

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

$\nabla f(x) = 0 \quad \nabla^2 f(x) \succeq 0 \Rightarrow$ loc min

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

$$\min_x \; -x$$
$$\text{s.t. } x^2 + y^2 \leq 1,$$
$$(x-1)^3 - y \leq 0.$$