Методы оптимизации. Семинар 10. Оптимальность. Условия ККТ.

Корнилов Никита Максимович

Московский физико-технический институт

5 ноября 2024г

Прямая и двойственная задачи

Прямая

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

Прямая и двойственная задачи

Прямая

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

Двойственная функция

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

Двойственная

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

Сильная двойственность

Напомним, что достаточным условием сильной двойственности, то есть $p^* = d^*$, например, является *ослабленное условие Слейтера*:

1. Функции f_0 и f_i являются выпуклыми, а h_j являются аффинными.
2. Существует такая допустимая точка x_0, что все *не афинные* условия неравенства выполняются строго $f_i(x_0) < 0$.

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

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

Поэтому, используя то, что $f_i(x^*) \leq 0$, мы получаем

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

Условия дополняющей нежёсткости

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

Поэтому, используя то, что $f_i(x^*) \le 0$, мы получаем

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

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

Условие ККТ

Давайте теперь еще предположим, что $f_0, f_1, \ldots f_m, h_1, \ldots h_n$ дифференцируемы в x^*. Тогда, так как x^* минимизирует $L(x, \lambda^*, \nu^*)$, градиент L по x в точке x^* должен быть равен нулю

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

Если $f_0, f_1, \ldots f_m, h_1, \ldots h_n$ только субдифференцируемы в x^*, то

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

Каруша-Куна-Такера:

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

Необходимое условие: для оптимального набора прямых и двойственных переменных при сильной двойственности следуют условия ККТ.

Достаточное условие: Когда f_i выпуклые, а h_j аффинные: для $\overline{x}, (\overline{\lambda}, \overline{\nu})$, которые удовлетворяют (4), выполняется следующее - эти точки доставляют оптимум прямой и двойственной задачи соответственно, и выполняется сильная двойственность.

Проверить это довольно легко – достаточно выписать равенство $g(\overline{\lambda}, \overline{\nu}) = L(\overline{x}, \overline{\lambda}, \overline{\nu}) = f_0(\overline{x})$.

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

Example

$$\min \frac{1}{2}x^T P x + q^T x + r$$
$$\text{s.t. } Ax = b,$$

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

Несколько условий неравенств

Example

Рассмотрим задачу минимизации

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

Проекция на l_2 шар

Example

Рассмотрим задачу минимизации

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

Проекция на l_1 шар

Example

Рассмотрим задачу минимизации

$$\min_x \|x - s\|_2^2$$
$$\text{s.t. } \|x\|_1 \le 1,$$

Example (Water-filling)

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

где $\alpha_i > 0$.

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.

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.

Example (Максимизация энтропии)

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

Примеры на решение

Example (Минимизация сепарабельной функции)

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

где f_i - строго выпуклые и дифференцируемые функции.

По условию Слейтера сильная выпуклость 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:

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

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

Условия регулярности

- Функции ограничений f_i и h_j являются аффинными;
- Для точки локального минимума x^* градиенты всех ограничений равенств и всех *активных* ограничений неравенств (выполняется равенство нулю) линейно независимы;
- Условие Слейтера.

Необходимые условия ККТ локального минимума

Theorem

Пусть x^* *является локальным минимумом прямой задачи. При этом 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.*

Example (Для точки минимума условия ККТ не выполнены)

$$\min_{x \in \mathbb{R}} x$$
$$\text{s.t. } x^2 \leq 0,$$

Second-Order Sufficient Condition (SOSC)

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

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

Definition

Достаточное условие второго порядка 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

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

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

Рассмотрим задачу оптимизации

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