

Методы оптимизации. Семинар 9. Двойственность.

Александр Катруца

Московский физико-технический институт

27 октября 2020 г.

Напоминание

- ▶ Существование решения оптимизационной задачи
- ▶ Условия оптимальности для
 - ▶ общей задачи оптимизации
 - ▶ задачи безусловной оптимизации
 - ▶ задачи оптимизации с ограничениями типа равенств
 - ▶ задачи оптимизации с ограничениями типа равенств и неравенств

Обозначения

Задача

$$\begin{aligned} \min f(\mathbf{x}) &= p^* \\ \text{s.t. } g_i(\mathbf{x}) &= 0, \quad i = 1, \dots, m \\ h_j(\mathbf{x}) &\leq 0, \quad j = 1, \dots, p \end{aligned}$$

Лагранжиан

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f(\mathbf{x}) + \sum_{i=1}^m \lambda_i g_i(\mathbf{x}) + \sum_{j=1}^p \mu_j h_j(\mathbf{x})$$

Двойственные переменные

Вектора $\boldsymbol{\mu}$ и $\boldsymbol{\lambda}$ называются двойственными переменными.

Двойственная функция

Функция $g(\boldsymbol{\mu}, \boldsymbol{\lambda}) = \inf_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu})$ называется двойственной функцией Лагранжа.

Свойства двойственной функции

Вогнутость

Двойственная функция является **вогнутой** как инфимум аффинных функций по (μ, λ) вне зависимости от того, является ли исходная задача выпуклой.

Нижняя граница

Для любого λ и для $\mu \geq 0$ выполнено $g(\mu, \lambda) \leq p^*$.

Двойственная задача

$$\begin{aligned} \max g(\mu, \lambda) &= d^* \\ \text{s.t. } \mu &\geq 0 \end{aligned}$$

Зачем?

- ▶ Двойственная задача выпукла независимо от того, выпукла ли прямая
- ▶ Нижняя оценка **может** достигаться

Связь с сопряжённой функцией

Рассмотрим задачу

$$\begin{aligned} \min f_0(\mathbf{x}) \\ \text{s.t. } \mathbf{Ax} \leq \mathbf{b} \\ \mathbf{Cx} = \mathbf{d} \end{aligned}$$

Тогда

$$\begin{aligned} g(\boldsymbol{\lambda}, \boldsymbol{\mu}) &= \inf_{\mathbf{x}} (f_0(\mathbf{x}) + \boldsymbol{\lambda}^\top (\mathbf{Ax} - \mathbf{b}) + \boldsymbol{\mu}^\top (\mathbf{Cx} - \mathbf{d})) = \\ &= -\mathbf{b}^\top \boldsymbol{\lambda} - \boldsymbol{\mu}^\top \mathbf{d} + \inf_{\mathbf{x}} (f_0(\mathbf{x}) + (\mathbf{A}^\top \boldsymbol{\lambda} + \mathbf{C}^\top \boldsymbol{\mu})^\top \mathbf{x}) = \\ &= -\mathbf{b}^\top \boldsymbol{\lambda} - \boldsymbol{\mu}^\top \mathbf{d} - f_0^*(-\mathbf{A}^\top \boldsymbol{\lambda} - \mathbf{C}^\top \boldsymbol{\mu}) \end{aligned}$$

Области определений двойственной и сопряжённой функций связаны:

$$\text{dom } g = \{(\boldsymbol{\lambda}, \boldsymbol{\mu}) \mid -\mathbf{A}^\top \boldsymbol{\lambda} - \mathbf{C}^\top \boldsymbol{\mu} \in \text{dom } f_0^*\}$$

Примеры

Найти двойственную функцию:

- ▶ Решение СЛУ минимальной нормы

$$\begin{aligned} \min \|\mathbf{x}\|_2^2 \\ \text{s.t. } \mathbf{Ax} = \mathbf{b} \end{aligned}$$

- ▶ Линейное программирование

$$\begin{aligned} \min \mathbf{c}^\top \mathbf{x} \\ \text{s.t. } \mathbf{Ax} = \mathbf{b} \\ \mathbf{x} \geq 0 \end{aligned}$$

- ▶ Задача разбиения

$$\begin{aligned} \min \mathbf{x}^\top \mathbf{W} \mathbf{x} \\ \text{s.t. } x_i^2 = 1, \quad i = 1, \dots, n \end{aligned}$$

Слабая и сильная двойственность

Определение

Оптимальные значения целевой функции в прямой и двойственной задаче связаны соотношением

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

Если $d^* < p^*$, то свойство называют слабой двойственностью.

Если $d^* = p^*$, то — сильной двойственностью.

Замечание

Слабая двойственность есть всегда по построению двойственной задачи.

Вопросы

- ▶ При каких условиях выполняется сильная двойственность?
- ▶ Как использовать двойственность для проверки оптимальности?

Критерий субоптимальности

По построению $p^* \geq g(\boldsymbol{\lambda}, \boldsymbol{\mu})$, поэтому

$$f_0(\mathbf{x}) - p^* \leq f_0(\mathbf{x}) - g(\boldsymbol{\lambda}, \boldsymbol{\mu}) = \varepsilon.$$

Определение

Разность $f_0(\mathbf{x}) - g(\boldsymbol{\lambda}, \boldsymbol{\mu})$ называется *двойственным зазором* и является оценкой сверху для разности текущего и оптимального значения функции.

Способы использования:

- ▶ критерий остановки в итерационном процессе
- ▶ теоретическая оценка сходимости алгоритма
- ▶ проверка оптимальности данной точки

Условия Слейтера

Теорема

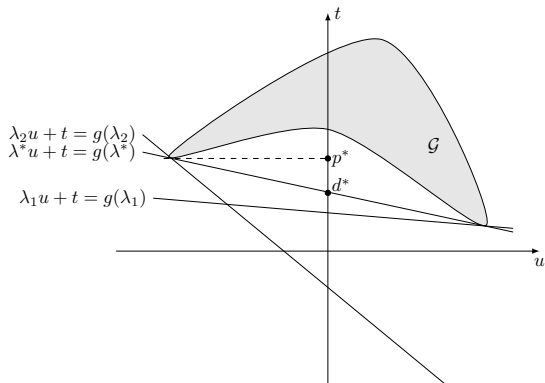
Если задача выпуклая и существует x , лежащий внутри допустимой области, т.е. ограничения типа неравенств выполнены как строгие неравенства, то выполнено свойство сильной двойственности.

- ▶ Решение СЛАУ наименьшей нормы
- ▶ Линейное программирование
- ▶ Квадратичное программирование с квадратичными ограничениями

Геометрическая интерпретация

$$\min_x f_0(x), \text{ где } f_1(x) \leq 0.$$

$$g(\lambda) = \inf_{(u,t) \in \mathcal{G}} (t + \lambda u) \quad \mathcal{G} = \{(f_1(x), f_0(x))\}$$



- ▶ $\lambda = 0$
- ▶ λ^* — оптимальное значение
- ▶ $\lambda > \lambda^*$

Условия дополняющей нежёсткости

Пусть \mathbf{x}^* и $(\boldsymbol{\mu}^*, \boldsymbol{\lambda}^*)$ решения прямой и двойственной задачи. То есть

$$\begin{aligned} f(\mathbf{x}^*) &= g(\boldsymbol{\mu}^*, \boldsymbol{\lambda}^*) = \inf_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) \leq \\ f(\mathbf{x}^*) &+ \sum_{i=1}^m \lambda_i^* g_i(\mathbf{x}^*) + \sum_{j=1}^p \mu_j^* h_j(\mathbf{x}^*) \leq \\ f(\mathbf{x}^*), \quad &\boldsymbol{\mu} \geq 0 \end{aligned}$$

Условия дополняющей нежёсткости

$$\mu_j^* h_j(\mathbf{x}^*) = 0, \quad j = 1, \dots, p$$

Для каждого неравенства

- ▶ либо множитель Лагранжа равен нулю
- ▶ либо оно активно.

Условия Каруша-Куна-Таккера

Из прошлого семинара известны необходимые условия ККТ:

1. $g_i(\mathbf{x}^*) = 0$ — допустимость в прямой задаче
2. $h_j(\mathbf{x}^*) \leq 0$ — допустимость в прямой задаче
3. $\mu_j^* \geq 0$ — допустимость в двойственной задаче
4. $\mu_j^* h_j(\mathbf{x}^*) = 0$ — условие дополняющей нежёсткости
5. $L'_x(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) = 0$ — стационарность лагранжиана по прямым переменным

Пример ($\mathbf{P} \in \mathbf{S}_+^n$)

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^n} \quad & \frac{1}{2} \mathbf{x}^\top \mathbf{P} \mathbf{x} + \mathbf{q}^\top \mathbf{x} + r \\ \text{s.t.} \quad & \mathbf{A} \mathbf{x} = \mathbf{b} \end{aligned}$$

Примеры

- ▶ Отрицательная энтропия при линейных ограничениях

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

- ▶ Сформулировать двойственную задачу и по её решению найти решение прямой задачи:

$$\begin{aligned} \min \quad & \frac{1}{2}x^2 + 2y^2 + \frac{1}{2}z^2 + x + y + 2z \\ \text{s.t.} \quad & x + 2y + z = 4 \end{aligned}$$

- ▶ Релаксация Лагранжа для задачи бинарного линейного программирования:

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^n} \quad & \mathbf{c}^\top \mathbf{x} \\ \text{s.t.} \quad & \mathbf{Ax} \leq \mathbf{b} \\ & x_i \in \{0, 1\}, \quad i = 1, \dots, n \end{aligned}$$

Главное

- ▶ Двойственная задача: что это такое и зачем она нужна?
- ▶ Сильная и слабая двойственность
- ▶ Условие Слейтера
- ▶ Геометрическая интерпретация
- ▶ Связь с сопряжённой функцией