

Методы оптимизации

Лекция 8: Введение в теорию двойственности

Александр Катруца

Физтех-школа прикладной математики и информатики
Московский физико-технический институт



2 ноября 2020 г.

На прошлой лекции

- ▶ Общие условия оптимальности
- ▶ Сопряжённые функции
- ▶ Свойства сопряжённых функций

План на эту лекцию

- ▶ Двойственная функция её свойства
- ▶ Двойственная задача
- ▶ Коническая двойственность и связь с сопряжёнными функциями

Равносильные преобразования задач

Равносильные преобразования задач

Запись через надграфик

$$\begin{array}{ll} \min_{\mathbf{x}} f_0(\mathbf{x}) & \\ \text{s.t. } g_i(\mathbf{x}) = 0, \ i = 1, \dots, m & \\ f_i(\mathbf{x}) \leq 0, \ i = 1, \dots, p & \end{array} \quad \Longrightarrow \quad \begin{array}{ll} \min_{\mathbf{x}, t} t & \\ \text{s.t. } g_i(\mathbf{x}) = 0, \ i = 1, \dots, m & \\ f_i(\mathbf{x}) \leq 0, \ i = 1, \dots, p & \\ f_0(\mathbf{x}) \leq t & \end{array}$$

Равносильные преобразования задач

Запись через надграфик

$$\begin{array}{ll} \min_{\mathbf{x}} f_0(\mathbf{x}) & \\ \text{s.t. } g_i(\mathbf{x}) = 0, i = 1, \dots, m & \\ f_i(\mathbf{x}) \leq 0, i = 1, \dots, p & \end{array} \implies \begin{array}{ll} \min_{\mathbf{x}, t} t & \\ \text{s.t. } g_i(\mathbf{x}) = 0, i = 1, \dots, m & \\ f_i(\mathbf{x}) \leq 0, i = 1, \dots, p & \\ f_0(\mathbf{x}) \leq t & \end{array}$$

Преобразования ограничений

- ▶ $\mathbf{g}(\mathbf{x}) \leq \mathbf{b} \rightarrow \mathbf{g}(\mathbf{x}) + \mathbf{y} = \mathbf{b}, \mathbf{y} \geq 0$
- ▶ $\mathbf{x} \rightarrow \mathbf{x}_1 - \mathbf{x}_2, \mathbf{x}_1 \geq 0, \mathbf{x}_2 \geq 0$
- ▶ Формирование блочных матриц

Равносильные преобразования задач

Запись через надграфик

$$\begin{array}{ll} \min_{\mathbf{x}} f_0(\mathbf{x}) & \\ \text{s.t. } g_i(\mathbf{x}) = 0, i = 1, \dots, m & \\ f_i(\mathbf{x}) \leq 0, i = 1, \dots, p & \end{array} \quad \Longrightarrow \quad \begin{array}{ll} \min_{\mathbf{x}, t} t & \\ \text{s.t. } g_i(\mathbf{x}) = 0, i = 1, \dots, m & \\ f_i(\mathbf{x}) \leq 0, i = 1, \dots, p & \\ f_0(\mathbf{x}) \leq t & \end{array}$$

Преобразования ограничений

- ▶ $\mathbf{g}(\mathbf{x}) \leq \mathbf{b} \rightarrow \mathbf{g}(\mathbf{x}) + \mathbf{y} = \mathbf{b}, \mathbf{y} \geq 0$
- ▶ $\mathbf{x} \rightarrow \mathbf{x}_1 - \mathbf{x}_2, \mathbf{x}_1 \geq 0, \mathbf{x}_2 \geq 0$
- ▶ Формирование блочных матриц

Перенос ограничений в целевую функцию

$$\min_{\mathbf{x} \in \mathcal{X}} f_0(\mathbf{x}) \Rightarrow \min_{\mathbf{x} \in \mathbb{R}^n} f_0(\mathbf{x}) + \mathbb{I}_{\mathcal{X}}(\mathbf{x}), \quad \mathbb{I}_{\mathcal{X}}(\mathbf{x}) = \begin{cases} 0, & \mathbf{x} \in \mathcal{X}, \\ +\infty, & \mathbf{x} \notin \mathcal{X}. \end{cases}$$

Задача оптимизации с функциональными ограничениями

$$\min_{\mathbf{x} \in \mathcal{D}} f_0(\mathbf{x})$$

$$\text{s.t. } g_i(\mathbf{x}) = 0, \quad i = 1, \dots, m$$

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

$$\text{dom } f_0 = \mathcal{D} \subseteq \mathbb{R}^n, \quad f_0(\mathbf{x}^*) = p^*$$

Задача оптимизации с функциональными ограничениями

$$\min_{\mathbf{x} \in \mathcal{D}} f_0(\mathbf{x})$$

$$\text{s.t. } g_i(\mathbf{x}) = 0, \quad i = 1, \dots, m$$

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

$$\text{dom } f_0 = \mathcal{D} \subseteq \mathbb{R}^n, \quad f_0(\mathbf{x}^*) = p^*$$

Основная цель

Сформулировать условия оптимальности для таких задач

Задача оптимизации с функциональными ограничениями

$$\begin{aligned} \min_{\mathbf{x} \in \mathcal{D}} f_0(\mathbf{x}) \\ \text{s.t. } g_i(\mathbf{x}) = 0, \quad i = 1, \dots, m \\ h_j(\mathbf{x}) \leq 0, \quad j = 1, \dots, p \end{aligned}$$

$$\text{dom } f_0 = \mathcal{D} \subseteq \mathbb{R}^n, \quad f_0(\mathbf{x}^*) = p^*$$

Основная цель

Сформулировать условия оптимальности для таких задач

Лагранжиан $L : \mathcal{D} \times \mathbb{R}^m \times \mathbb{R}^p \rightarrow \mathbb{R}$

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f_0(\mathbf{x}) + \sum_{i=1}^m \lambda_i g_i(\mathbf{x}) + \sum_{j=1}^p \mu_j h_j(\mathbf{x})$$

- ▶ λ_i – множители Лагранжа для ограничений $g_i(\mathbf{x}) = 0, \quad i = 1, \dots, m$
- ▶ μ_j – множители Лагранжа для ограничений $h_j(\mathbf{x}) \leq 0, \quad j = 1, \dots, p$

Двойственная функция

Определение

Функция $g : \mathbb{R}^m \times \mathbb{R}^p \rightarrow \mathbb{R}$ такая что

$$g(\boldsymbol{\lambda}, \boldsymbol{\mu}) = \inf_{\mathbf{x} \in \mathcal{D}} \left(f_0(\mathbf{x}) + \sum_{i=1}^m \lambda_i g_i(\mathbf{x}) + \sum_{j=1}^p \mu_j h_j(\mathbf{x}) \right)$$

называется *двойственной функцией*

Двойственная функция

Определение

Функция $g : \mathbb{R}^m \times \mathbb{R}^p \rightarrow \mathbb{R}$ такая что

$$g(\boldsymbol{\lambda}, \boldsymbol{\mu}) = \inf_{\mathbf{x} \in \mathcal{D}} \left(f_0(\mathbf{x}) + \sum_{i=1}^m \lambda_i g_i(\mathbf{x}) + \sum_{j=1}^p \mu_j h_j(\mathbf{x}) \right)$$

называется *двойственной функцией*

Свойства

- ▶ Всегда вогнута
- ▶ Может равняться $-\infty$ для некоторых $(\boldsymbol{\lambda}, \boldsymbol{\mu})$

Нижняя оценка оптимального значения функции

Утверждение

Если $\mu \geq 0$, тогда $f_0(\mathbf{x}^*) = p^* \geq g(\lambda, \mu)$

Нижняя оценка оптимального значения функции

Утверждение

Если $\mu \geq 0$, тогда $f_0(\mathbf{x}^*) = p^* \geq g(\lambda, \mu)$

Доказательство

Нижняя оценка оптимального значения функции

Утверждение

Если $\mu \geq 0$, тогда $f_0(\mathbf{x}^*) = p^* \geq g(\boldsymbol{\lambda}, \mu)$

Доказательство

- ▶ Если $\hat{\mathbf{x}} \in \mathcal{D}$ и лежит в допустимом множестве, а также $\mu \geq 0$, тогда

$$f_0(\hat{\mathbf{x}}) \geq L(\hat{\mathbf{x}}, \boldsymbol{\lambda}, \mu) \geq \inf_{\mathbf{x} \in \mathcal{D}} L(\mathbf{x}, \boldsymbol{\lambda}, \mu) = g(\boldsymbol{\lambda}, \mu)$$

Нижняя оценка оптимального значения функции

Утверждение

Если $\mu \geq 0$, тогда $f_0(\mathbf{x}^*) = p^* \geq g(\boldsymbol{\lambda}, \mu)$

Доказательство

- ▶ Если $\hat{\mathbf{x}} \in \mathcal{D}$ и лежит в допустимом множестве, а также $\mu \geq 0$, тогда

$$f_0(\hat{\mathbf{x}}) \geq L(\hat{\mathbf{x}}, \boldsymbol{\lambda}, \mu) \geq \inf_{\mathbf{x} \in \mathcal{D}} L(\mathbf{x}, \boldsymbol{\lambda}, \mu) = g(\boldsymbol{\lambda}, \mu)$$

- ▶ Минимизируя обе части по всем допустимым $\hat{\mathbf{x}}$, получим

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

Двойственная задача

Определение

Двойственной задачей называется следующая задача

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

Двойственная задача

Определение

Двойственной задачей называется следующая задача

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

- Всегда выпуклая задача

Двойственная задача

Определение

Двойственной задачей называется следующая задача

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

- ▶ Всегда выпуклая задача
- ▶ Обозначим $d^* = g(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$

Двойственная задача

Определение

Двойственной задачей называется следующая задача

$$\begin{aligned} \max g(\boldsymbol{\lambda}, \boldsymbol{\mu}) \\ \text{s.t. } \boldsymbol{\mu} \geq 0 \end{aligned}$$

- ▶ Всегда выпуклая задача
- ▶ Обозначим $d^* = g(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$
- ▶ Лучшая нижняя оценка для p^* , которую может дать двойственная функция

Двойственная задача

Определение

Двойственной задачей называется следующая задача

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

- ▶ Всегда выпуклая задача
- ▶ Обозначим $d^* = g(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$
- ▶ Лучшая нижняя оценка для p^* , которую может дать двойственная функция
- ▶ Вектора $(\boldsymbol{\lambda}, \boldsymbol{\mu})$ называются допустимыми для двойственной задачи, если $\boldsymbol{\mu} \geq 0$ и $(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \text{dom } g$

Связь с сопряжённой функцией

Рассмотрим задачу

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

Тогда

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

Области определений двойственной и сопряжённой функций связаны:

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

Слабая и сильная двойственность

Слабая двойственность: $d^* \leq p^*$

Слабая и сильная двойственность

Слабая двойственность: $d^* \leq p^*$

- ▶ Всегда выполняется по построению двойственной задачи

Слабая и сильная двойственность

Слабая двойственность: $d^* \leq p^*$

- ▶ Всегда выполняется по построению двойственной задачи
- ▶ Нетривиальные нижние границы для (NP-)сложных задач

Слабая и сильная двойственность

Слабая двойственность: $d^* \leq p^*$

- ▶ Всегда выполняется по построению двойственной задачи
- ▶ Нетривиальные нижние границы для (NP-)сложных задач

Сильная двойственность: $d^* = p^*$

Слабая и сильная двойственность

Слабая двойственность: $d^* \leq p^*$

- ▶ Всегда выполняется по построению двойственной задачи
- ▶ Нетривиальные нижние границы для (NP-)сложных задач

Сильная двойственность: $d^* = p^*$

- ▶ В общем случае **НЕ** выполняется

Слабая и сильная двойственность

Слабая двойственность: $d^* \leq p^*$

- ▶ Всегда выполняется по построению двойственной задачи
- ▶ Нетривиальные нижние границы для (NP-)сложных задач

Сильная двойственность: $d^* = p^*$

- ▶ В общем случае **НЕ** выполняется
- ▶ *Обычно* выполнена для выпуклых задач

Слабая и сильная двойственность

Слабая двойственность: $d^* \leq p^*$

- ▶ Всегда выполняется по построению двойственной задачи
- ▶ Нетривиальные нижние границы для (NP-)сложных задач

Сильная двойственность: $d^* = p^*$

- ▶ В общем случае **НЕ** выполняется
- ▶ *Обычно* выполнена для выпуклых задач
- ▶ Условия регулярности ограничений

Слабая и сильная двойственность

Слабая двойственность: $d^* \leq p^*$

- ▶ Всегда выполняется по построению двойственной задачи
- ▶ Нетривиальные нижние границы для (NP-)сложных задач

Сильная двойственность: $d^* = p^*$

- ▶ В общем случае **НЕ** выполняется
- ▶ *Обычно* выполнена для выпуклых задач
- ▶ Условия регулярности ограничений
- ▶ Может выполняться и для **невыпуклых** задач¹

¹Beck, Amir, and Yonina C. Eldar. "Strong duality in nonconvex quadratic optimization with two quadratic constraints." SIAM Journal on Optimization 17.3 (2006): 844-860.

Слабая и сильная двойственность

Слабая двойственность: $d^* \leq p^*$

- ▶ Всегда выполняется по построению двойственной задачи
- ▶ Нетривиальные нижние границы для (NP-)сложных задач

Сильная двойственность: $d^* = p^*$

- ▶ В общем случае **НЕ** выполняется
- ▶ *Обычно* выполнена для выпуклых задач
- ▶ Условия регулярности ограничений
- ▶ Может выполняться и для **невыпуклых** задач¹
- ▶ Решение двойственной задачи даёт решение прямой задачи

¹Beck, Amir, and Yonina C. Eldar. "Strong duality in nonconvex quadratic optimization with two quadratic constraints." SIAM Journal on Optimization 17.3 (2006): 844-860.

Условие оптимальности и зазор двойственности

Зазор двойственности: $f_0(\mathbf{x}) - g(\boldsymbol{\lambda}, \boldsymbol{\mu})$

Условие оптимальности и зазор двойственности

Зазор двойственности: $f_0(\mathbf{x}) - g(\boldsymbol{\lambda}, \boldsymbol{\mu})$

- Оценка точности решения

Условие оптимальности и зазор двойственности

Зазор двойственности: $f_0(\mathbf{x}) - g(\boldsymbol{\lambda}, \boldsymbol{\mu})$

- ▶ Оценка точности решения
- ▶ Доказательство корректности и сходимости алгоритма

Условие оптимальности и зазор двойственности

Зазор двойственности: $f_0(\mathbf{x}) - g(\boldsymbol{\lambda}, \boldsymbol{\mu})$

- ▶ Оценка точности решения
- ▶ Доказательство корректности и сходимости алгоритма

Теорема

Если $\hat{\mathbf{x}}$ лежит в допустимом множестве и найдутся допустимые $(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ такие что

$$f_0(\hat{\mathbf{x}}) = g(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}),$$

то $\hat{\mathbf{x}}$ является решением задачи.

Условие оптимальности и зазор двойственности

Зазор двойственности: $f_0(\mathbf{x}) - g(\boldsymbol{\lambda}, \boldsymbol{\mu})$

- ▶ Оценка точности решения
- ▶ Доказательство корректности и сходимости алгоритма

Теорема

Если $\hat{\mathbf{x}}$ лежит в допустимом множестве и найдутся допустимые $(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ такие что

$$f_0(\hat{\mathbf{x}}) = g(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}),$$

то $\hat{\mathbf{x}}$ является решением задачи.

Доказательство

Условие оптимальности и зазор двойственности

Зазор двойственности: $f_0(\mathbf{x}) - g(\boldsymbol{\lambda}, \boldsymbol{\mu})$

- ▶ Оценка точности решения
- ▶ Доказательство корректности и сходимости алгоритма

Теорема

Если $\hat{\mathbf{x}}$ лежит в допустимом множестве и найдутся допустимые $(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ такие что

$$f_0(\hat{\mathbf{x}}) = g(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}),$$

то $\hat{\mathbf{x}}$ является решением задачи.

Доказательство

- ▶ По построению двойственной задачи: $p^* \geq g(\boldsymbol{\lambda}, \boldsymbol{\mu})$ для всех допустимых $(\boldsymbol{\lambda}, \boldsymbol{\mu})$

Условие оптимальности и зазор двойственности

Зазор двойственности: $f_0(\mathbf{x}) - g(\boldsymbol{\lambda}, \boldsymbol{\mu})$

- ▶ Оценка точности решения
- ▶ Доказательство корректности и сходимости алгоритма

Теорема

Если $\hat{\mathbf{x}}$ лежит в допустимом множестве и найдутся допустимые $(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ такие что

$$f_0(\hat{\mathbf{x}}) = g(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}),$$

то $\hat{\mathbf{x}}$ является решением задачи.

Доказательство

- ▶ По построению двойственной задачи: $p^* \geq g(\boldsymbol{\lambda}, \boldsymbol{\mu})$ для всех допустимых $(\boldsymbol{\lambda}, \boldsymbol{\mu})$
- ▶ Тогда выполнено, что $f_0(\mathbf{x}) - p^* \leq f_0(\mathbf{x}) - g(\boldsymbol{\lambda}, \boldsymbol{\mu})$

Условие оптимальности и зазор двойственности

Зазор двойственности: $f_0(\mathbf{x}) - g(\boldsymbol{\lambda}, \boldsymbol{\mu})$

- ▶ Оценка точности решения
- ▶ Доказательство корректности и сходимости алгоритма

Теорема

Если $\hat{\mathbf{x}}$ лежит в допустимом множестве и найдутся допустимые $(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ такие что

$$f_0(\hat{\mathbf{x}}) = g(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}),$$

то $\hat{\mathbf{x}}$ является решением задачи.

Доказательство

- ▶ По построению двойственной задачи: $p^* \geq g(\boldsymbol{\lambda}, \boldsymbol{\mu})$ для всех допустимых $(\boldsymbol{\lambda}, \boldsymbol{\mu})$
- ▶ Тогда выполнено, что $f_0(\mathbf{x}) - p^* \leq f_0(\mathbf{x}) - g(\boldsymbol{\lambda}, \boldsymbol{\mu})$
- ▶ Если нашёлся допустимый $\hat{\mathbf{x}}$ и допустимые $(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$, такие что $f_0(\hat{\mathbf{x}}) = g(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$, то $f_0(\hat{\mathbf{x}}) - p^* \leq 0$

Условие оптимальности и зазор двойственности

Зазор двойственности: $f_0(\mathbf{x}) - g(\boldsymbol{\lambda}, \boldsymbol{\mu})$

- ▶ Оценка точности решения
- ▶ Доказательство корректности и сходимости алгоритма

Теорема

Если $\hat{\mathbf{x}}$ лежит в допустимом множестве и найдутся допустимые $(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ такие что

$$f_0(\hat{\mathbf{x}}) = g(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}),$$

то $\hat{\mathbf{x}}$ является решением задачи.

Доказательство

- ▶ По построению двойственной задачи: $p^* \geq g(\boldsymbol{\lambda}, \boldsymbol{\mu})$ для всех допустимых $(\boldsymbol{\lambda}, \boldsymbol{\mu})$
- ▶ Тогда выполнено, что $f_0(\mathbf{x}) - p^* \leq f_0(\mathbf{x}) - g(\boldsymbol{\lambda}, \boldsymbol{\mu})$
- ▶ Если нашёлся допустимый $\hat{\mathbf{x}}$ и допустимые $(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$, такие что $f_0(\hat{\mathbf{x}}) = g(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$, то $f_0(\hat{\mathbf{x}}) - p^* \leq 0$
- ▶ Так как p^* минимальное значение f_0 , то $p^* = f_0(\hat{\mathbf{x}})$

Седловая точка

Определение седловой точки

Точка \bar{x} называется седловой точкой дифференцируемой функции f если она стационарная точка, то есть $f'(\bar{x}) = 0$, но не является локальным экстремумом, то есть найдутся направления $\mathbf{d}_1, \mathbf{d}_2$ такие что $f(\bar{x} + \mathbf{d}_1) \leq f(\bar{x}) \leq f(\bar{x} + \mathbf{d}_2)$

Седловая точка

Определение седловой точки

Точка \bar{x} называется седловой точкой дифференцируемой функции f если она стационарная точка, то есть $f'(\bar{x}) = 0$, но не является локальным экстремумом, то есть найдутся направления $\mathbf{d}_1, \mathbf{d}_2$ такие что $f(\bar{x} + \mathbf{d}_1) \leq f(\bar{x}) \leq f(\bar{x} + \mathbf{d}_2)$

Классический пример

Функция $f(x, y) = x^2 - y^2$, для которой точка $(0, 0)$ является седловой.

Седловая точка

Определение седловой точки

Точка \bar{x} называется седловой точкой дифференцируемой функции f если она стационарная точка, то есть $f'(\bar{x}) = 0$, но не является локальным экстремумом, то есть найдутся направления $\mathbf{d}_1, \mathbf{d}_2$ такие что $f(\bar{x} + \mathbf{d}_1) \leq f(\bar{x}) \leq f(\bar{x} + \mathbf{d}_2)$

Классический пример

Функция $f(x, y) = x^2 - y^2$, для которой точка $(0, 0)$ является седловой.

Частный случай

Если функция f зависит от двух переменных $f(x, y)$, то точка (\bar{x}, \bar{y}) будет седловой, если $f(\bar{x}, y) \leq f(\bar{x}, \bar{y}) \leq f(x, \bar{y})$ для точек из области определения.

Условие оптимальности и седловая точка функции Лагранжа

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f_0(\mathbf{x}) + \sum_{i=1}^m \lambda_i g_i(\mathbf{x}) + \sum_{j=1}^p \mu_j h_j(\mathbf{x})$$

Условие оптимальности и седловая точка функции Лагранжа

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f_0(\mathbf{x}) + \sum_{i=1}^m \lambda_i g_i(\mathbf{x}) + \sum_{j=1}^p \mu_j h_j(\mathbf{x})$$

Теорема

Пусть $\hat{\mathbf{x}} \in \mathcal{D}$ и $(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ допустимы. Тогда эквивалентны следующие условия

- 1) $\hat{\mathbf{x}}$ лежит в допустимом множестве и $f_0(\hat{\mathbf{x}}) = g(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$
- 2) для всех $\mathbf{x} \in \mathcal{D}$ и всех допустимых $(\boldsymbol{\lambda}, \boldsymbol{\mu})$ выполнено

$$L(\hat{\mathbf{x}}, \boldsymbol{\lambda}, \boldsymbol{\mu}) \leq L(\hat{\mathbf{x}}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}) \leq L(\mathbf{x}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}),$$

то есть точка $(\hat{\mathbf{x}}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ есть седловая точка функции Лагранжа

- 3) $\hat{\mathbf{x}}$ лежит в допустимом множестве, является точкой минимума функции $L(\mathbf{x}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ и выполнено $\hat{\mu}_j h_j(\hat{\mathbf{x}}) = 0$

Доказательство: 1) \rightarrow 2)

- Для допустимого $\hat{\mathbf{x}}$ и произвольной допустимой пары $(\boldsymbol{\lambda}, \boldsymbol{\mu})$ выполнено

$$L(\hat{\mathbf{x}}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f_0(\hat{\mathbf{x}}) + \sum_{i=1}^m \lambda_i g_i(\hat{\mathbf{x}}) + \sum_{j=1}^p \mu_j h_j(\hat{\mathbf{x}}) = f_0(\hat{\mathbf{x}}) + \sum_{j=1}^p \mu_j h_j(\hat{\mathbf{x}}) \leq f_0(\hat{\mathbf{x}}) \text{ так как } \mu_j \geq 0 \text{ и } h_j(\hat{\mathbf{x}}) \leq 0$$

Доказательство: 1) \rightarrow 2)

- ▶ Для допустимого $\hat{\mathbf{x}}$ и произвольной допустимой пары $(\boldsymbol{\lambda}, \boldsymbol{\mu})$ выполнено
$$L(\hat{\mathbf{x}}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f_0(\hat{\mathbf{x}}) + \sum_{i=1}^m \lambda_i g_i(\hat{\mathbf{x}}) + \sum_{j=1}^p \mu_j h_j(\hat{\mathbf{x}}) = f_0(\hat{\mathbf{x}}) + \sum_{j=1}^p \mu_j h_j(\hat{\mathbf{x}}) \leq f_0(\hat{\mathbf{x}})$$
 так как $\mu_j \geq 0$ и $h_j(\hat{\mathbf{x}}) \leq 0$
- ▶ Также $g(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}) = \inf_{\mathbf{u} \in \mathcal{D}} L(\mathbf{u}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}) \leq L(\mathbf{x}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ для любого $\mathbf{x} \in \mathcal{D}$

Доказательство: 1) \rightarrow 2)

- ▶ Для допустимого $\hat{\mathbf{x}}$ и произвольной допустимой пары $(\boldsymbol{\lambda}, \boldsymbol{\mu})$ выполнено
$$L(\hat{\mathbf{x}}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f_0(\hat{\mathbf{x}}) + \sum_{i=1}^m \lambda_i g_i(\hat{\mathbf{x}}) + \sum_{j=1}^p \mu_j h_j(\hat{\mathbf{x}}) = f_0(\hat{\mathbf{x}}) + \sum_{j=1}^p \mu_j h_j(\hat{\mathbf{x}}) \leq f_0(\hat{\mathbf{x}})$$
 так как $\mu_j \geq 0$ и $h_j(\hat{\mathbf{x}}) \leq 0$
- ▶ Также $g(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}) = \inf_{\mathbf{u} \in \mathcal{D}} L(\mathbf{u}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}) \leq L(\mathbf{x}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ для любого $\mathbf{x} \in \mathcal{D}$
- ▶ Если $f(\hat{\mathbf{x}}) = g(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$, то выполнено

$$L(\hat{\mathbf{x}}, \boldsymbol{\lambda}, \boldsymbol{\mu}) \leq f_0(\hat{\mathbf{x}}) = g(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}) \leq L(\mathbf{x}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$$

для всех $\mathbf{x} \in \mathcal{D}$ и допустимых пар $(\boldsymbol{\lambda}, \boldsymbol{\mu})$

Доказательство: 1) \rightarrow 2)

- ▶ Для допустимого $\hat{\mathbf{x}}$ и произвольной допустимой пары $(\boldsymbol{\lambda}, \boldsymbol{\mu})$ выполнено
$$L(\hat{\mathbf{x}}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f_0(\hat{\mathbf{x}}) + \sum_{i=1}^m \lambda_i g_i(\hat{\mathbf{x}}) + \sum_{j=1}^p \mu_j h_j(\hat{\mathbf{x}}) = f_0(\hat{\mathbf{x}}) + \sum_{j=1}^p \mu_j h_j(\hat{\mathbf{x}}) \leq f_0(\hat{\mathbf{x}})$$
 так как $\mu_j \geq 0$ и $h_j(\hat{\mathbf{x}}) \leq 0$
- ▶ Также $g(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}) = \inf_{\mathbf{u} \in \mathcal{D}} L(\mathbf{u}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}) \leq L(\mathbf{x}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ для любого $\mathbf{x} \in \mathcal{D}$
- ▶ Если $f(\hat{\mathbf{x}}) = g(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$, то выполнено

$$L(\hat{\mathbf{x}}, \boldsymbol{\lambda}, \boldsymbol{\mu}) \leq f_0(\hat{\mathbf{x}}) = g(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}) \leq L(\mathbf{x}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$$

для всех $\mathbf{x} \in \mathcal{D}$ и допустимых пар $(\boldsymbol{\lambda}, \boldsymbol{\mu})$

- ▶ Если $\mathbf{x} = \hat{\mathbf{x}}$ и $(\boldsymbol{\lambda}, \boldsymbol{\mu}) = (\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$, то $f_0(\hat{\mathbf{x}}) = L(\hat{\mathbf{x}}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$

Доказательство: 2) \rightarrow 3)

- Из правого неравенства в условии седловой точки следует, что \hat{x} точка минимума Лагранжиана

Доказательство: 2) \rightarrow 3)

- ▶ Из правого неравенства в условии седловой точки следует, что $\hat{\mathbf{x}}$ точка минимума Лагранжиана
- ▶ Из левого неравенства следует, что

$$f_0(\hat{\mathbf{x}}) + \sum_{i=1}^m \lambda_i g_i(\hat{\mathbf{x}}) + \sum_{j=1}^p \mu_j h_j(\hat{\mathbf{x}}) \leq f_0(\hat{\mathbf{x}}) + \sum_{i=1}^m \hat{\lambda}_i g_i(\hat{\mathbf{x}}) + \sum_{j=1}^p \hat{\mu}_j h_j(\hat{\mathbf{x}})$$

Доказательство: 2) \rightarrow 3)

- ▶ Из правого неравенства в условии седловой точки следует, что $\hat{\mathbf{x}}$ точка минимума Лагранжиана
- ▶ Из левого неравенства следует, что

$$f_0(\hat{\mathbf{x}}) + \sum_{i=1}^m \lambda_i g_i(\hat{\mathbf{x}}) + \sum_{j=1}^p \mu_j h_j(\hat{\mathbf{x}}) \leq f_0(\hat{\mathbf{x}}) + \sum_{i=1}^m \hat{\lambda}_i g_i(\hat{\mathbf{x}}) + \sum_{j=1}^p \hat{\mu}_j h_j(\hat{\mathbf{x}})$$

- ▶ Или $\sum_{i=1}^m (\lambda_i - \hat{\lambda}_i) g_i(\hat{\mathbf{x}}) + \sum_{j=1}^p (\mu_j - \hat{\mu}_j) h_j(\hat{\mathbf{x}}) \leq 0$ для всех допустимых $(\boldsymbol{\lambda}, \boldsymbol{\mu})$

Доказательство: 2) \rightarrow 3)

- ▶ Из правого неравенства в условии седловой точки следует, что $\hat{\mathbf{x}}$ точка минимума Лагранжиана
- ▶ Из левого неравенства следует, что

$$f_0(\hat{\mathbf{x}}) + \sum_{i=1}^m \lambda_i g_i(\hat{\mathbf{x}}) + \sum_{j=1}^p \mu_j h_j(\hat{\mathbf{x}}) \leq f_0(\hat{\mathbf{x}}) + \sum_{i=1}^m \hat{\lambda}_i g_i(\hat{\mathbf{x}}) + \sum_{j=1}^p \hat{\mu}_j h_j(\hat{\mathbf{x}})$$

- ▶ Или $\sum_{i=1}^m (\lambda_i - \hat{\lambda}_i) g_i(\hat{\mathbf{x}}) + \sum_{j=1}^p (\mu_j - \hat{\mu}_j) h_j(\hat{\mathbf{x}}) \leq 0$ для всех допустимых $(\boldsymbol{\lambda}, \boldsymbol{\mu})$
- ▶ Пусть $\boldsymbol{\mu} = \hat{\boldsymbol{\mu}}$ и $\lambda_k = \hat{\lambda}_k \pm 1$, а остальные $\lambda_i = \hat{\lambda}_i$ для $i \neq k$, тогда $\pm g_k(\hat{\mathbf{x}}) \leq 0$, то есть $g_k(\hat{\mathbf{x}}) = 0$

Доказательство: 2) \rightarrow 3)

- ▶ Из правого неравенства в условии седловой точки следует, что $\hat{\mathbf{x}}$ точка минимума Лагранжиана
- ▶ Из левого неравенства следует, что

$$f_0(\hat{\mathbf{x}}) + \sum_{i=1}^m \lambda_i g_i(\hat{\mathbf{x}}) + \sum_{j=1}^p \mu_j h_j(\hat{\mathbf{x}}) \leq f_0(\hat{\mathbf{x}}) + \sum_{i=1}^m \hat{\lambda}_i g_i(\hat{\mathbf{x}}) + \sum_{j=1}^p \hat{\mu}_j h_j(\hat{\mathbf{x}})$$

- ▶ Или $\sum_{i=1}^m (\lambda_i - \hat{\lambda}_i) g_i(\hat{\mathbf{x}}) + \sum_{j=1}^p (\mu_j - \hat{\mu}_j) h_j(\hat{\mathbf{x}}) \leq 0$ для всех допустимых $(\boldsymbol{\lambda}, \boldsymbol{\mu})$
- ▶ Пусть $\boldsymbol{\mu} = \hat{\boldsymbol{\mu}}$ и $\lambda_k = \hat{\lambda}_k \pm 1$, а остальные $\lambda_i = \hat{\lambda}_i$ для $i \neq k$, тогда $\pm g_k(\hat{\mathbf{x}}) \leq 0$, то есть $g_k(\hat{\mathbf{x}}) = 0$
- ▶ Пусть $\boldsymbol{\lambda} = \hat{\boldsymbol{\lambda}}$, но $\mu_k = \hat{\mu}_k + 1$ и $\mu_j = \hat{\mu}_j$ для $j \neq k$, тогда $h_k(\hat{\mathbf{x}}) \leq 0$. Таким образом, $\hat{\mathbf{x}}$ лежит в допустимом множестве

Доказательство: 2) \rightarrow 3)

- ▶ Из правого неравенства в условии седловой точки следует, что $\hat{\mathbf{x}}$ точка минимума Лагранжиана
- ▶ Из левого неравенства следует, что

$$f_0(\hat{\mathbf{x}}) + \sum_{i=1}^m \lambda_i g_i(\hat{\mathbf{x}}) + \sum_{j=1}^p \mu_j h_j(\hat{\mathbf{x}}) \leq f_0(\hat{\mathbf{x}}) + \sum_{i=1}^m \hat{\lambda}_i g_i(\hat{\mathbf{x}}) + \sum_{j=1}^p \hat{\mu}_j h_j(\hat{\mathbf{x}})$$

- ▶ Или $\sum_{i=1}^m (\lambda_i - \hat{\lambda}_i) g_i(\hat{\mathbf{x}}) + \sum_{j=1}^p (\mu_j - \hat{\mu}_j) h_j(\hat{\mathbf{x}}) \leq 0$ для всех допустимых $(\boldsymbol{\lambda}, \boldsymbol{\mu})$
- ▶ Пусть $\boldsymbol{\mu} = \hat{\boldsymbol{\mu}}$ и $\lambda_k = \hat{\lambda}_k \pm 1$, а остальные $\lambda_i = \hat{\lambda}_i$ для $i \neq k$, тогда $\pm g_k(\hat{\mathbf{x}}) \leq 0$, то есть $g_k(\hat{\mathbf{x}}) = 0$
- ▶ Пусть $\boldsymbol{\lambda} = \hat{\boldsymbol{\lambda}}$, но $\mu_k = \hat{\mu}_k + 1$ и $\mu_j = \hat{\mu}_j$ для $j \neq k$, тогда $h_k(\hat{\mathbf{x}}) \leq 0$. Таким образом, $\hat{\mathbf{x}}$ лежит в допустимом множестве
- ▶ Аналогично, если $\boldsymbol{\lambda} = \hat{\boldsymbol{\lambda}}$, но $\mu_k = 2\hat{\mu}_k$ или $\mu_k = 0$, а $\mu_j = \hat{\mu}_j$ для $j \neq k$, тогда $\hat{\mu}_k h_k(\hat{\mathbf{x}}) = 0$

Доказательство: 3) \rightarrow 1)

- ▶ Так как $\hat{\mathbf{x}}$ точка минимума лагранжиана, то
$$g(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}) = \inf_{\mathbf{z} \in \mathcal{D}} L(\mathbf{z}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}) = L(\hat{\mathbf{x}}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}) =$$
$$f_0(\hat{\mathbf{x}}) + \sum_{i=1}^m \hat{\lambda}_i g_i(\hat{\mathbf{x}}) + \sum_{j=1}^p \hat{\mu}_j h_j(\hat{\mathbf{x}})$$

Доказательство: 3) \rightarrow 1)

- ▶ Так как $\hat{\mathbf{x}}$ точка минимума лагранжиана, то
$$g(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}) = \inf_{\mathbf{z} \in \mathcal{D}} L(\mathbf{z}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}) = L(\hat{\mathbf{x}}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}) =$$
$$f_0(\hat{\mathbf{x}}) + \sum_{i=1}^m \hat{\lambda}_i g_i(\hat{\mathbf{x}}) + \sum_{j=1}^p \hat{\mu}_j h_j(\hat{\mathbf{x}})$$
- ▶ $\hat{\mathbf{x}}$ лежит в допустимом множестве, поэтому $g_i(\hat{\mathbf{x}}) = 0$ для $i = 1, \dots, m$

Доказательство: 3) \rightarrow 1)

- ▶ Так как $\hat{\mathbf{x}}$ точка минимума лагранжиана, то
$$g(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}) = \inf_{\mathbf{z} \in \mathcal{D}} L(\mathbf{z}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}) = L(\hat{\mathbf{x}}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}) = f_0(\hat{\mathbf{x}}) + \sum_{i=1}^m \hat{\lambda}_i g_i(\hat{\mathbf{x}}) + \sum_{j=1}^p \hat{\mu}_j h_j(\hat{\mathbf{x}})$$
- ▶ $\hat{\mathbf{x}}$ лежит в допустимом множестве, поэтому $g_i(\hat{\mathbf{x}}) = 0$ для $i = 1, \dots, m$
- ▶ Так как выполнено $\hat{\mu}_j h_j(\hat{\mathbf{x}}) = 0$, где $j = 1, \dots, p$, то в итоге
$$g(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}) = f_0(\hat{\mathbf{x}})$$

Доказательство: 3) \rightarrow 1)

- ▶ Так как $\hat{\mathbf{x}}$ точка минимума лагранжиана, то
$$g(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}) = \inf_{\mathbf{z} \in \mathcal{D}} L(\mathbf{z}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}) = L(\hat{\mathbf{x}}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}) = f_0(\hat{\mathbf{x}}) + \sum_{i=1}^m \hat{\lambda}_i g_i(\hat{\mathbf{x}}) + \sum_{j=1}^p \hat{\mu}_j h_j(\hat{\mathbf{x}})$$
- ▶ $\hat{\mathbf{x}}$ лежит в допустимом множестве, поэтому $g_i(\hat{\mathbf{x}}) = 0$ для $i = 1, \dots, m$
- ▶ Так как выполнено $\hat{\mu}_j h_j(\hat{\mathbf{x}}) = 0$, где $j = 1, \dots, p$, то в итоге $g(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}) = f_0(\hat{\mathbf{x}})$
- ▶ Таким образом, выполнено условие оптимальности

Условия Каруша-Куна-Таккера (ККТ)

Следствие

Пусть \mathbf{x}^* решение задачи минимизации такое, что $\mathbf{x}^* \in \text{int}(\mathcal{D})$, f_0, g_i, h_j дифференцируемы в \mathbf{x}^* и выполнен критерий оптимальности $f_0(\mathbf{x}^*) = g(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ для некоторой допустимой пары $(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$, тогда выполнено

$$\begin{cases} L'_{\mathbf{x}}(\mathbf{x}^*, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}) = 0 \\ \hat{\mu}_k h_k(\mathbf{x}^*) = 0 \end{cases}$$

Условия ККТ

- ▶ $L'_{\mathbf{x}}(\mathbf{x}^*, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}) = 0$
- ▶ $\hat{\mu}_k h_k(\mathbf{x}^*) = 0$ — условия дополняющей нежёсткости
- ▶ $\hat{\boldsymbol{\mu}} \geq 0$
- ▶ $h_j(\mathbf{x}^*) \leq 0$
- ▶ $g_i(\mathbf{x}^*) = 0$

Анонс на следующую лекцию

- ▶ Следствие на предыдущем слайде требует наличия сильной двойственности
- ▶ Если задача выпукла и выполнено некоторое условие регулярности, то сильная двойственность выполняется. Этот факт будет доказан на следующей лекции.
- ▶ На следующей лекции обсудим, что меняется когда появляется выпуклость и докажем основную теорему для условий оптимальности

Переформулировка задачи

- ▶ Равносильные прямые задачи могут давать совершенно разные двойственные задачи

Переформулировка задачи

- ▶ Равносильные прямые задачи могут давать совершенно разные двойственные задачи
- ▶ Равносильное преобразование исходной задачи может дать более простую или полезную двойственную задачу

Переформулировка задачи

- ▶ Равносильные прямые задачи могут давать совершенно разные двойственные задачи
- ▶ Равносильное преобразование исходной задачи может дать более простую или полезную двойственную задачу

Стандартные приёмы

Переформулировка задачи

- ▶ Равносильные прямые задачи могут давать совершенно разные двойственные задачи
- ▶ Равносильное преобразование исходной задачи может дать более простую или полезную двойственную задачу

Стандартные приёмы

- ▶ Введение новых переменных

$$\min_{\mathbf{x}} \|\mathbf{Ax} - \mathbf{b}\| \rightarrow \begin{array}{l} \min_{(\mathbf{x}, \mathbf{y})} \|\mathbf{y}\| \\ \text{s.t. } \mathbf{Ax} - \mathbf{b} = \mathbf{y} \end{array}$$

Переформулировка задачи

- ▶ Равносильные прямые задачи могут давать совершенно разные двойственные задачи
- ▶ Равносильное преобразование исходной задачи может дать более простую или полезную двойственную задачу

Стандартные приёмы

- ▶ Введение новых переменных

$$\min_{\mathbf{x}} \|\mathbf{Ax} - \mathbf{b}\| \rightarrow \min_{(\mathbf{x}, \mathbf{y})} \|\mathbf{y}\|$$

s.t. $\mathbf{Ax} - \mathbf{b} = \mathbf{y}$

- ▶ Превращение явных ограничений в неявные

$$\begin{array}{ll} \min_{\mathbf{x} \in \mathbb{R}^n} \mathbf{c}^\top \mathbf{x} & \\ \text{s.t. } -1 \leq \mathbf{x} \leq 1 & \end{array} \rightarrow \begin{array}{ll} \min_{-1 \leq \mathbf{x} \leq 1} \mathbf{c}^\top \mathbf{x} & \\ \text{s.t. } \mathbf{Ax} = \mathbf{b} & \end{array}$$

Двойственная задача к задаче линейного программирования (LP)

- ▶ Исходная задача

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

Двойственная задача к задаче линейного программирования (LP)

- ▶ Исходная задача

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

- ▶ Лагранжиан:

$$L = \mathbf{c}^\top \mathbf{x} + \boldsymbol{\lambda}^\top (\mathbf{Ax} - \mathbf{b}) - \boldsymbol{\mu}^\top \mathbf{x} = (\mathbf{c} + \mathbf{A}^\top \boldsymbol{\lambda} - \boldsymbol{\mu})^\top \mathbf{x} - \boldsymbol{\lambda}^\top \mathbf{b}$$

Двойственная задача к задаче линейного программирования (LP)

- ▶ Исходная задача

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

- ▶ Лагранжиан:

$$L = \mathbf{c}^\top \mathbf{x} + \boldsymbol{\lambda}^\top (\mathbf{Ax} - \mathbf{b}) - \boldsymbol{\mu}^\top \mathbf{x} = (\mathbf{c} + \mathbf{A}^\top \boldsymbol{\lambda} - \boldsymbol{\mu})^\top \mathbf{x} - \boldsymbol{\lambda}^\top \mathbf{b}$$

- ▶ Двойственная функция

$$g(\boldsymbol{\lambda}, \boldsymbol{\mu}) = \begin{cases} -\boldsymbol{\lambda}^\top \mathbf{b}, & \mathbf{c} + \mathbf{A}^\top \boldsymbol{\lambda} - \boldsymbol{\mu} = 0, \\ -\infty, & \text{иначе.} \end{cases}$$

Двойственная задача к задаче линейного программирования (LP)

- ▶ Исходная задача

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

- ▶ Лагранжиан:

$$L = \mathbf{c}^\top \mathbf{x} + \boldsymbol{\lambda}^\top (\mathbf{Ax} - \mathbf{b}) - \boldsymbol{\mu}^\top \mathbf{x} = (\mathbf{c} + \mathbf{A}^\top \boldsymbol{\lambda} - \boldsymbol{\mu})^\top \mathbf{x} - \boldsymbol{\lambda}^\top \mathbf{b}$$

- ▶ Двойственная функция

$$g(\boldsymbol{\lambda}, \boldsymbol{\mu}) = \begin{cases} -\boldsymbol{\lambda}^\top \mathbf{b}, & \mathbf{c} + \mathbf{A}^\top \boldsymbol{\lambda} - \boldsymbol{\mu} = 0, \\ -\infty, & \text{иначе.} \end{cases}$$

- ▶ Двойственная задача

$$\begin{aligned} \min \boldsymbol{\lambda}^\top \mathbf{b} \\ \text{s.t. } \mathbf{A}^\top \boldsymbol{\lambda} + \mathbf{c} \geq 0 \end{aligned}$$

Связь между неограниченностью и неразрешимостью

Теорема

Если допустимое множество в прямой задаче LP пусто, то двойственная задача не ограничена.

Связь между неограниченностью и неразрешимостью

Теорема

Если допустимое множество в прямой задаче LP пусто, то двойственная задача не ограничена.

Доказательство

Связь между неограниченностью и неразрешимостью

Теорема

Если допустимое множество в прямой задаче LP пусто, то двойственная задача не ограничена.

Доказательство

- ▶ Допустимое множество в прямой задаче
 $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq 0\}$

Связь между неограниченностью и неразрешимостью

Теорема

Если допустимое множество в прямой задаче LP пусто, то двойственная задача не ограничена.

Доказательство

- ▶ Допустимое множество в прямой задаче $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq 0\}$
- ▶ Если оно пустое, то по лемме Фаркаша найдётся вектор \mathbf{p} такой что $\mathbf{p}^\top \mathbf{b} < 0$ и $\mathbf{p}^\top \mathbf{A} \geq 0$

Связь между неограниченностью и неразрешимостью

Теорема

Если допустимое множество в прямой задаче LP пусто, то двойственная задача не ограничена.

Доказательство

- ▶ Допустимое множество в прямой задаче $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq 0\}$
- ▶ Если оно пустое, то по лемме Фаркаша найдётся вектор \mathbf{p} такой что $\mathbf{p}^\top \mathbf{b} < 0$ и $\mathbf{p}^\top \mathbf{A} \geq 0$
- ▶ Двойственная задача

$$\begin{aligned} & \min \mathbf{\lambda}^\top \mathbf{b} \\ & \text{s.t. } \mathbf{A}^\top \mathbf{\lambda} + \mathbf{c} \geq 0 \end{aligned}$$

Связь между неограниченностью и неразрешимостью

Теорема

Если допустимое множество в прямой задаче LP пусто, то двойственная задача не ограничена.

Доказательство

- ▶ Допустимое множество в прямой задаче $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq 0\}$
- ▶ Если оно пустое, то по лемме Фаркаша найдётся вектор \mathbf{p} такой что $\mathbf{p}^\top \mathbf{b} < 0$ и $\mathbf{p}^\top \mathbf{A} \geq 0$
- ▶ Двойственная задача

$$\begin{aligned} \min \quad & \lambda^\top \mathbf{b} \\ \text{s.t.} \quad & \mathbf{A}^\top \lambda + \mathbf{c} \geq 0 \end{aligned}$$

- ▶ Пусть $\hat{\lambda} = \theta \mathbf{p}$, $\theta > 0$, тогда $\theta \mathbf{p}^\top \mathbf{b} \rightarrow -\infty$ и $\theta \mathbf{A}^\top \mathbf{p} + \mathbf{c} \geq 0$

Связь между неограниченностью и неразрешимостью

Теорема

Если допустимое множество в прямой задаче LP пусто, то двойственная задача не ограничена.

Доказательство

- ▶ Допустимое множество в прямой задаче $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq 0\}$
- ▶ Если оно пустое, то по лемме Фаркаша найдётся вектор \mathbf{p} такой что $\mathbf{p}^\top \mathbf{b} < 0$ и $\mathbf{p}^\top \mathbf{A} \geq 0$
- ▶ Двойственная задача

$$\begin{aligned} \min \quad & \lambda^\top \mathbf{b} \\ \text{s.t.} \quad & \mathbf{A}^\top \lambda + \mathbf{c} \geq 0 \end{aligned}$$

- ▶ Пусть $\hat{\lambda} = \theta \mathbf{p}$, $\theta > 0$, тогда $\theta \mathbf{p}^\top \mathbf{b} \rightarrow -\infty$ и $\theta \mathbf{A}^\top \mathbf{p} + \mathbf{c} \geq 0$
- ▶ Двойственная задача не ограничена

Зачем нужны конусы?

Обобщённое отношение частичного порядка

Пусть \mathcal{K} выпуклый, замкнутый конус, $\mathcal{K} \cap -\mathcal{K} = \{0\}$. Тогда $\mathbf{x} \leq_{\mathcal{K}} \mathbf{y} \Leftrightarrow \mathbf{y} - \mathbf{x} \in \mathcal{K}$ — отношение частичного порядка (докажите аксиомы!).

Зачем нужны конусы?

Обобщённое отношение частичного порядка

Пусть \mathcal{K} выпуклый, замкнутый конус, $\mathcal{K} \cap -\mathcal{K} = \{0\}$. Тогда $\mathbf{x} \leq_{\mathcal{K}} \mathbf{y} \Leftrightarrow \mathbf{y} - \mathbf{x} \in \mathcal{K}$ — отношение частичного порядка (докажите аксиомы!).

Пример: конус $\mathcal{K} = \mathbf{S}_+^n$

Пусть $\mathbf{X}, \mathbf{Y} \in \mathbf{S}^n$. Тогда $\mathbf{X} \leq_{\mathcal{K}} \mathbf{Y}$ означает, что $\mathbf{Y} - \mathbf{X} \in \mathbf{S}_+^n$

Зачем нужны конусы?

Обобщённое отношение частичного порядка

Пусть \mathcal{K} выпуклый, замкнутый конус, $\mathcal{K} \cap -\mathcal{K} = \{0\}$. Тогда $\mathbf{x} \leq_{\mathcal{K}} \mathbf{y} \Leftrightarrow \mathbf{y} - \mathbf{x} \in \mathcal{K}$ — отношение частичного порядка (докажите аксиомы!).

Пример: конус $\mathcal{K} = \mathbf{S}_+^n$

Пусть $\mathbf{X}, \mathbf{Y} \in \mathbf{S}^n$. Тогда $\mathbf{X} \leq_{\mathcal{K}} \mathbf{Y}$ означает, что $\mathbf{Y} - \mathbf{X} \in \mathbf{S}_+^n$

Задача линейного программирования

$$\begin{aligned} & \min_{\mathbf{x} \in \mathbb{R}^n} \mathbf{c}^\top \mathbf{x} \\ & \text{s.t. } \mathbf{A}\mathbf{x} = \mathbf{b} \\ & \quad x_i \geq 0 \end{aligned}$$

$$\begin{aligned} & \min_{\mathbf{x} \in \mathbb{R}^n} \mathbf{c}^\top \mathbf{x} \\ & \text{s.t. } \mathbf{A}\mathbf{x} = \mathbf{b} \\ & \quad \mathbf{x} \in \mathbb{R}_+^n \end{aligned}$$

Зачем нужны конусы?

Обобщённое отношение частичного порядка

Пусть \mathcal{K} выпуклый, замкнутый конус, $\mathcal{K} \cap -\mathcal{K} = \{0\}$. Тогда $\mathbf{x} \leq_{\mathcal{K}} \mathbf{y} \Leftrightarrow \mathbf{y} - \mathbf{x} \in \mathcal{K}$ — отношение частичного порядка (докажите аксиомы!).

Пример: конус $\mathcal{K} = \mathbf{S}_+^n$

Пусть $\mathbf{X}, \mathbf{Y} \in \mathbf{S}^n$. Тогда $\mathbf{X} \leq_{\mathcal{K}} \mathbf{Y}$ означает, что $\mathbf{Y} - \mathbf{X} \in \mathbf{S}_+^n$

Задача линейного программирования

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^n} \quad & \mathbf{c}^\top \mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} = \mathbf{b} \\ & x_i \geq 0 \end{aligned}$$

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^n} \quad & \mathbf{c}^\top \mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \in \mathbb{R}_+^n \end{aligned}$$

Введение нелинейности

Использование декартового произведения трёх самосопряжённых конусов позволяет записать многие практически важные выпуклые задачи

Двойственность и обобщённые неравенства

Постановка задачи

$$\begin{aligned} & \min_{\mathbf{x} \in \mathcal{D}} f_0(\mathbf{x}) \\ \text{s.t. } & g_i(\mathbf{x}) = 0, \quad i = 1, \dots, m \\ & h_j(\mathbf{x}) \leq_{\kappa} 0, \quad j = 1, \dots, p \end{aligned}$$

Двойственность и обобщённые неравенства

Постановка задачи

$$\begin{aligned} & \min_{\mathbf{x} \in \mathcal{D}} f_0(\mathbf{x}) \\ & \text{s.t. } g_i(\mathbf{x}) = 0, \quad i = 1, \dots, m \\ & \quad h_j(\mathbf{x}) \leq_{\mathcal{K}} 0, \quad j = 1, \dots, p \end{aligned}$$

Лагранжиан

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f_0(\mathbf{x}) + \langle \boldsymbol{\lambda}, \mathbf{g}(\mathbf{x}) \rangle + \langle \boldsymbol{\mu}, \mathbf{h}(\mathbf{x}) \rangle$$

- ▶ $h_j(\mathbf{x}) \leq_{\mathcal{K}} 0 \Leftrightarrow -h_j(\mathbf{x}) \in \mathcal{K}$
- ▶ условие $\langle \boldsymbol{\mu}, \mathbf{h}(\mathbf{x}) \rangle \leq 0 \Leftrightarrow \langle \boldsymbol{\mu}, -\mathbf{h}(\mathbf{x}) \rangle \geq 0$ или $\langle \boldsymbol{\mu}, \mathbf{u} \rangle \geq 0$, где $\mathbf{u} \in \mathcal{K}$ выполнено, если $\boldsymbol{\mu} \in \mathcal{K}^*$ или $\boldsymbol{\mu} \geq_{\mathcal{K}^*} 0$

Двойственность и обобщённые неравенства

Постановка задачи

$$\begin{aligned} & \min_{\mathbf{x} \in \mathcal{D}} f_0(\mathbf{x}) \\ & \text{s.t. } g_i(\mathbf{x}) = 0, \quad i = 1, \dots, m \\ & \quad h_j(\mathbf{x}) \leq_{\kappa} 0, \quad j = 1, \dots, p \end{aligned}$$

Лагранжиан

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f_0(\mathbf{x}) + \langle \boldsymbol{\lambda}, \mathbf{g}(\mathbf{x}) \rangle + \langle \boldsymbol{\mu}, \mathbf{h}(\mathbf{x}) \rangle$$

- ▶ $h_j(\mathbf{x}) \leq_{\kappa} 0 \Leftrightarrow -h_j(\mathbf{x}) \in \mathcal{K}$
- ▶ условие $\langle \boldsymbol{\mu}, \mathbf{h}(\mathbf{x}) \rangle \leq 0 \Leftrightarrow \langle \boldsymbol{\mu}, -\mathbf{h}(\mathbf{x}) \rangle \geq 0$ или $\langle \boldsymbol{\mu}, \mathbf{u} \rangle \geq 0$, где $\mathbf{u} \in \mathcal{K}$ выполнено, если $\boldsymbol{\mu} \in \mathcal{K}^*$ или $\boldsymbol{\mu} \geq_{\mathcal{K}^*} 0$

Двойственная задача

$$\begin{aligned} & \max g(\boldsymbol{\lambda}, \boldsymbol{\mu}) \\ & \text{s.t. } \boldsymbol{\mu} \geq_{\mathcal{K}^*} 0 \end{aligned}$$

Двойственность для задачи конической оптимизации

- ▶ Стандартная форма

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

Двойственность для задачи конической оптимизации

- ▶ Стандартная форма

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

- ▶ Двойственная задача (аналогично LP)

$$\begin{aligned} \max \boldsymbol{\lambda}^\top \mathbf{b} \\ \text{s.t. } \mathbf{A}^\top \boldsymbol{\lambda} \leq_{\mathcal{K}^*} \mathbf{c} \end{aligned}$$

Двойственность для задачи конической оптимизации

- ▶ Стандартная форма

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

- ▶ Двойственная задача (аналогично LP)

$$\begin{aligned} \max \boldsymbol{\lambda}^\top \mathbf{b} \\ \text{s.t. } \mathbf{A}^\top \boldsymbol{\lambda} \leq_{\mathcal{K}^*} \mathbf{c} \end{aligned}$$

- ▶ Если конус \mathcal{K} самосопряжённый мы автоматически знаем, как выглядит двойственная задача!

Двойственная задача для SDP

- ▶ Исходная задача

$$\begin{aligned} \min_{\mathbf{X}} \quad & \text{trace}(\mathbf{C}\mathbf{X}) \\ \text{s.t.} \quad & \text{trace}(\mathbf{A}_i\mathbf{X}) = b_i \\ & \mathbf{X} \succeq 0 \end{aligned}$$

Двойственная задача для SDP

- ▶ Исходная задача

$$\begin{aligned} \min_{\mathbf{X}} \quad & \text{trace}(\mathbf{C}\mathbf{X}) \\ \text{s.t.} \quad & \text{trace}(\mathbf{A}_i\mathbf{X}) = b_i \\ & \mathbf{X} \succeq 0 \end{aligned}$$

- ▶ Двойственная задача

$$\begin{aligned} \max_{\boldsymbol{\lambda}} \quad & \boldsymbol{\lambda}^\top \mathbf{b} \\ \text{s.t.} \quad & \mathbf{C} - \sum_{i=1}^m \lambda_i \mathbf{A}_i \succeq 0 \end{aligned}$$

Пример выпуклой задачи с положительным зазором двойственности

Рассмотрим задачу

$$\begin{array}{ll} \min & x_2 \\ \text{s.t.} & \begin{bmatrix} x_2 + 1 & 0 & 0 \\ 0 & x_1 & x_2 \\ 0 & x_2 & 0 \end{bmatrix} \succeq 0 \end{array}$$

Пример выпуклой задачи с положительным зазором двойственности

Рассмотрим задачу

$$\begin{array}{ll}\min & x_2 \\ \text{s.t.} & \begin{bmatrix} x_2 + 1 & 0 & 0 \\ 0 & x_1 & x_2 \\ 0 & x_2 & 0 \end{bmatrix} \succeq 0\end{array}$$

► Допустимое множество:

$$x_2 \geq -1, x_1 \geq 0, -x_2^2 \geq 0, x_1(x_2 + 1) \geq 0, (x_2 + 1)(-x_2^2) \geq 0$$

Пример выпуклой задачи с положительным зазором двойственности

Рассмотрим задачу

$$\begin{array}{ll}\min & x_2 \\ \text{s.t.} & \begin{bmatrix} x_2 + 1 & 0 & 0 \\ 0 & x_1 & x_2 \\ 0 & x_2 & 0 \end{bmatrix} \succeq 0\end{array}$$

- ▶ Допустимое множество:

$$x_2 \geq -1, x_1 \geq 0, -x_2^2 \geq 0, x_1(x_2 + 1) \geq 0, (x_2 + 1)(-x_2^2) \geq 0$$

- ▶ $p^* = 0$

Пример выпуклой задачи с положительным зазором двойственности

Рассмотрим задачу

$$\begin{array}{ll}\min & x_2 \\ \text{s.t.} & \begin{bmatrix} x_2 + 1 & 0 & 0 \\ 0 & x_1 & x_2 \\ 0 & x_2 & 0 \end{bmatrix} \succeq 0\end{array}$$

► Допустимое множество:

$$x_2 \geq -1, x_1 \geq 0, -x_2^2 \geq 0, x_1(x_2 + 1) \geq 0, (x_2 + 1)(-x_2^2) \geq 0$$

► $p^* = 0$

Двойственная задача имеет вид

$$\begin{array}{ll}\min & -y_{11} \\ \text{s.t.} & \mathbf{Y} \succeq 0 \\ & y_{11} + y_{32} + y_{23} = 1 \\ & y_{22} = 0\end{array}$$

Пример выпуклой задачи с положительным зазором двойственности

Рассмотрим задачу

$$\begin{array}{ll}\min & x_2 \\ \text{s.t.} & \begin{bmatrix} x_2 + 1 & 0 & 0 \\ 0 & x_1 & x_2 \\ 0 & x_2 & 0 \end{bmatrix} \succeq 0\end{array}$$

- ▶ Допустимое множество:

$$x_2 \geq -1, x_1 \geq 0, -x_2^2 \geq 0, x_1(x_2 + 1) \geq 0, (x_2 + 1)(-x_2^2) \geq 0$$

- ▶ $p^* = 0$

Двойственная задача имеет вид

$$\begin{array}{ll}\min & -y_{11} \\ \text{s.t.} & \mathbf{Y} \succeq 0 \\ & y_{11} + y_{32} + y_{23} = 1 \\ & y_{22} = 0\end{array}$$

- ▶ Допустимое множество: $y_{11} \geq 0, y_{22}y_{33} - y_{23}y_{32} \geq 0$

Пример выпуклой задачи с положительным зазором двойственности

Рассмотрим задачу

$$\begin{array}{ll}\min & x_2 \\ \text{s.t.} & \begin{bmatrix} x_2 + 1 & 0 & 0 \\ 0 & x_1 & x_2 \\ 0 & x_2 & 0 \end{bmatrix} \succeq 0\end{array}$$

► Допустимое множество:

$$x_2 \geq -1, x_1 \geq 0, -x_2^2 \geq 0, x_1(x_2 + 1) \geq 0, (x_2 + 1)(-x_2^2) \geq 0$$

► $p^* = 0$

Двойственная задача имеет вид

$$\begin{array}{ll}\min & -y_{11} \\ \text{s.t.} & \mathbf{Y} \succeq 0 \\ & y_{11} + y_{32} + y_{23} = 1 \\ & y_{22} = 0\end{array}$$

► Допустимое множество: $y_{11} \geq 0, y_{22}y_{33} - y_{23}y_{32} \geq 0$

► $d^* = -1$

- ▶ Преобразования задач и их типы

Главное

- ▶ Преобразования задач и их типы
- ▶ Двойственная функция и её свойства

Главное

- ▶ Преобразования задач и их типы
- ▶ Двойственная функция и её свойства
- ▶ Двойственная задача и её свойства

- ▶ Преобразования задач и их типы
- ▶ Двойственная функция и её свойства
- ▶ Двойственная задача и её свойства
- ▶ Сильная двойственность и слабая двойственность

- ▶ Преобразования задач и их типы
- ▶ Двойственная функция и её свойства
- ▶ Двойственная задача и её свойства
- ▶ Сильная двойственность и слабая двойственность
- ▶ Обобщённые неравенства