

Методы оптимизации

Лекция 4: Условия оптимальности.

Введение в теорию двойственности

Александр Катруца

Факультет инноваций и высоких технологий
Физтех-школа прикладной математики и информатики



25 сентября 2019 г.

На прошлой лекции

- ▶ Постановки задач выпуклой оптимизации
- ▶ LP, SOCP, SDP
- ▶ Примеры приложений

Равносильные преобразования задач

Равносильные преобразования задач

Запись через надграфик

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

Равносильные преобразования задач

Запись через надграфик

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

Преобразования ограничений

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

Равносильные преобразования задач

Запись через надграфик

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

Преобразования ограничений

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

Перенос ограничений в целевую функцию

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

Условие оптимальности для задачи без ограничений

$$\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}) \quad (1)$$

Теорема

Если \mathbf{x}^ решение задачи (1), то $f'(\mathbf{x}^*) = 0$.*

Условие оптимальности для задачи без ограничений

$$\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}) \quad (1)$$

Теорема

Если \mathbf{x}^ решение задачи (1), то $f'(\mathbf{x}^*) = 0$.*

Доказательство

Условие оптимальности для задачи без ограничений

$$\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}) \quad (1)$$

Теорема

Если \mathbf{x}^* решение задачи (1), то $f'(\mathbf{x}^*) = 0$.

Доказательство

► $f(\mathbf{y}) = f(\mathbf{x}^*) + \langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x}^* \rangle + r(\mathbf{x}^*, \mathbf{y})$ и

$$\lim_{\mathbf{y} \rightarrow \mathbf{x}^*} \frac{r(\mathbf{x}^*, \mathbf{y})}{\|\mathbf{x}^* - \mathbf{y}\|_2} = 0 \quad (*)$$

Условие оптимальности для задачи без ограничений

$$\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}) \quad (1)$$

Теорема

Если \mathbf{x}^* решение задачи (1), то $f'(\mathbf{x}^*) = 0$.

Доказательство

- ▶ $f(\mathbf{y}) = f(\mathbf{x}^*) + \langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x}^* \rangle + r(\mathbf{x}^*, \mathbf{y})$ и
$$\lim_{\mathbf{y} \rightarrow \mathbf{x}^*} \frac{r(\mathbf{x}^*, \mathbf{y})}{\|\mathbf{x}^* - \mathbf{y}\|_2} = 0 \quad (*)$$
- ▶ Если $f'(\mathbf{x}^*) \neq 0$, тогда рассмотрим $\mathbf{y}(\tau) = \mathbf{x}^* - \tau f'(\mathbf{x}^*)$, $\tau > 0$

Условие оптимальности для задачи без ограничений

$$\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}) \quad (1)$$

Теорема

Если \mathbf{x}^* решение задачи (1), то $f'(\mathbf{x}^*) = 0$.

Доказательство

- ▶ $f(\mathbf{y}) = f(\mathbf{x}^*) + \langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x}^* \rangle + r(\mathbf{x}^*, \mathbf{y})$ и $\lim_{\mathbf{y} \rightarrow \mathbf{x}^*} \frac{r(\mathbf{x}^*, \mathbf{y})}{\|\mathbf{x}^* - \mathbf{y}\|_2} = 0$ (*)
- ▶ Если $f'(\mathbf{x}^*) \neq 0$, тогда рассмотрим $\mathbf{y}(\tau) = \mathbf{x}^* - \tau f'(\mathbf{x}^*)$, $\tau > 0$
- ▶ $f(\mathbf{y}(\tau)) = f(\mathbf{x}^*) + \langle f'(\mathbf{x}^*), \mathbf{y}(\tau) - \mathbf{x}^* \rangle + r(\mathbf{x}^*, \mathbf{y}(\tau)) = f(\mathbf{x}^*) - \tau \|f'(\mathbf{x}^*)\|_2^2 + r(\mathbf{x}^*, \mathbf{y}(\tau))$

Условие оптимальности для задачи без ограничений

$$\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}) \quad (1)$$

Теорема

Если \mathbf{x}^* решение задачи (1), то $f'(\mathbf{x}^*) = 0$.

Доказательство

- ▶ $f(\mathbf{y}) = f(\mathbf{x}^*) + \langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x}^* \rangle + r(\mathbf{x}^*, \mathbf{y})$ и $\lim_{\mathbf{y} \rightarrow \mathbf{x}^*} \frac{r(\mathbf{x}^*, \mathbf{y})}{\|\mathbf{x}^* - \mathbf{y}\|_2} = 0$ (*)
- ▶ Если $f'(\mathbf{x}^*) \neq 0$, тогда рассмотрим $\mathbf{y}(\tau) = \mathbf{x}^* - \tau f'(\mathbf{x}^*)$, $\tau > 0$
- ▶ $f(\mathbf{y}(\tau)) = f(\mathbf{x}^*) + \langle f'(\mathbf{x}^*), \mathbf{y}(\tau) - \mathbf{x}^* \rangle + r(\mathbf{x}^*, \mathbf{y}(\tau)) = f(\mathbf{x}^*) - \tau \|f'(\mathbf{x}^*)\|_2^2 + r(\mathbf{x}^*, \mathbf{y}(\tau))$
- ▶ В силу (*) найдётся $\bar{\tau}$ такое что для всех $\tau \in (0, \bar{\tau})$ выполнено

$$f(\mathbf{y}(\tau)) - f(\mathbf{x}^*) \leq -\frac{1}{2} \|f'(\mathbf{x}^*)\|_2^2 < 0$$

Значит \mathbf{x}^* не минимум, противоречие.

Критерий для выпуклой задачи без ограничений

Теорема

Если в задаче (1) функция f выпукла, то \mathbf{x}^ глобальный минимум тогда и только тогда $f'(\mathbf{x}^*) = 0$*

Критерий для выпуклой задачи без ограничений

Теорема

Если в задаче (1) функция f выпукла, то \mathbf{x}^ глобальный минимум тогда и только тогда $f'(\mathbf{x}^*) = 0$*

Доказательство

Критерий для выпуклой задачи без ограничений

Теорема

Если в задаче (1) функция f выпукла, то \mathbf{x}^ глобальный минимум тогда и только тогда $f'(\mathbf{x}^*) = 0$*

Доказательство

- ▶ Если \mathbf{x}^* глобальный минимум, а f выпукла, то \mathbf{x}^* локальный минимум

Критерий для выпуклой задачи без ограничений

Теорема

Если в задаче (1) функция f выпукла, то \mathbf{x}^ глобальный минимум тогда и только тогда $f'(\mathbf{x}^*) = 0$*

Доказательство

- ▶ Если \mathbf{x}^* глобальный минимум, а f выпукла, то \mathbf{x}^* локальный минимум
- ▶ Значит $f'(\mathbf{x}^*) = 0$ по предыдущей теореме

Критерий для выпуклой задачи без ограничений

Теорема

Если в задаче (1) функция f выпукла, то \mathbf{x}^ глобальный минимум тогда и только тогда $f'(\mathbf{x}^*) = 0$*

Доказательство

- ▶ Если \mathbf{x}^* глобальный минимум, а f выпукла, то \mathbf{x}^* локальный минимум
- ▶ Значит $f'(\mathbf{x}^*) = 0$ по предыдущей теореме
- ▶ Пусть \mathbf{x}^* такая точка, что $f'(\mathbf{x}^*) = 0$ и функция выпукла

Критерий для выпуклой задачи без ограничений

Теорема

Если в задаче (1) функция f выпукла, то \mathbf{x}^* глобальный минимум тогда и только тогда $f'(\mathbf{x}^*) = 0$

Доказательство

- ▶ Если \mathbf{x}^* глобальный минимум, а f выпукла, то \mathbf{x}^* локальный минимум
- ▶ Значит $f'(\mathbf{x}^*) = 0$ по предыдущей теореме
- ▶ Пусть \mathbf{x}^* такая точка, что $f'(\mathbf{x}^*) = 0$ и функция выпукла
- ▶ Тогда по критерию выпуклости

$$f(\mathbf{y}) \geq f(\mathbf{x}^*) + \langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x}^* \rangle = f(\mathbf{x}^*)$$

Критерий для выпуклой задачи без ограничений

Теорема

Если в задаче (1) функция f выпукла, то \mathbf{x}^* глобальный минимум тогда и только тогда $f'(\mathbf{x}^*) = 0$

Доказательство

- ▶ Если \mathbf{x}^* глобальный минимум, а f выпукла, то \mathbf{x}^* локальный минимум
- ▶ Значит $f'(\mathbf{x}^*) = 0$ по предыдущей теореме
- ▶ Пусть \mathbf{x}^* такая точка, что $f'(\mathbf{x}^*) = 0$ и функция выпукла
- ▶ Тогда по критерию выпуклости

$$f(\mathbf{y}) \geq f(\mathbf{x}^*) + \langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x}^* \rangle = f(\mathbf{x}^*)$$

- ▶ Значит \mathbf{x}^* – глобальный минимум.

Дифференциальное условие оптимальности для задачи с ограничениями

$$\min_{\mathbf{x} \in X} f(\mathbf{x}) \quad (2)$$

Теорема

Точка \mathbf{x}^ – решение задачи (2), где f – выпуклая функция, тогда и только тогда, когда $\mathbf{x}^* \in X$ и $\langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x}^* \rangle \geq 0$ для всех $\mathbf{y} \in X$.*

Дифференциальное условие оптимальности для задачи с ограничениями

$$\min_{\mathbf{x} \in X} f(\mathbf{x}) \quad (2)$$

Теорема

Точка \mathbf{x}^ – решение задачи (2), где f – выпуклая функция, тогда и только тогда, когда $\mathbf{x}^* \in X$ и $\langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x}^* \rangle \geq 0$ для всех $\mathbf{y} \in X$.*

Доказательство

Дифференциальное условие оптимальности для задачи с ограничениями

$$\min_{\mathbf{x} \in X} f(\mathbf{x}) \quad (2)$$

Теорема

Точка \mathbf{x}^ – решение задачи (2), где f – выпуклая функция, тогда и только тогда, когда $\mathbf{x}^* \in X$ и $\langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x}^* \rangle \geq 0$ для всех $\mathbf{y} \in X$.*

Доказательство

- Пусть $\mathbf{x}^* \in X$ и выполнено неравенство. Тогда по критерию первого порядка для выпуклой функции $f(\mathbf{y}) \geq f(\mathbf{x}^*)$

Дифференциальное условие оптимальности для задачи с ограничениями

$$\min_{\mathbf{x} \in X} f(\mathbf{x}) \quad (2)$$

Теорема

Точка \mathbf{x}^* – решение задачи (2), где f – выпуклая функция, тогда и только тогда, когда $\mathbf{x}^* \in X$ и $\langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x}^* \rangle \geq 0$ для всех $\mathbf{y} \in X$.

Доказательство

- ▶ Пусть $\mathbf{x}^* \in X$ и выполнено неравенство. Тогда по критерию первого порядка для выпуклой функции $f(\mathbf{y}) \geq f(\mathbf{x}^*)$
- ▶ Пусть \mathbf{x}^* решение задачи (2), но найдётся $\tilde{\mathbf{y}}$ такой что $\langle f'(\mathbf{x}^*), \tilde{\mathbf{y}} - \mathbf{x}^* \rangle < 0$

Дифференциальное условие оптимальности для задачи с ограничениями

$$\min_{\mathbf{x} \in X} f(\mathbf{x}) \quad (2)$$

Теорема

Точка \mathbf{x}^* – решение задачи (2), где f – выпуклая функция, тогда и только тогда, когда $\mathbf{x}^* \in X$ и $\langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x}^* \rangle \geq 0$ для всех $\mathbf{y} \in X$.

Доказательство

- ▶ Пусть $\mathbf{x}^* \in X$ и выполнено неравенство. Тогда по критерию первого порядка для выпуклой функции $f(\mathbf{y}) \geq f(\mathbf{x}^*)$
- ▶ Пусть \mathbf{x}^* решение задачи (2), но найдётся $\tilde{\mathbf{y}}$ такой что $\langle f'(\mathbf{x}^*), \tilde{\mathbf{y}} - \mathbf{x}^* \rangle < 0$
- ▶ Рассмотрим точку $\mathbf{z}(t) = t\tilde{\mathbf{y}} + (1 - t)\mathbf{x}^*$, $t \in [0, 1]$

Дифференциальное условие оптимальности для задачи с ограничениями

$$\min_{\mathbf{x} \in X} f(\mathbf{x}) \quad (2)$$

Теорема

Точка \mathbf{x}^* – решение задачи (2), где f – выпуклая функция, тогда и только тогда, когда $\mathbf{x}^* \in X$ и $\langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x}^* \rangle \geq 0$ для всех $\mathbf{y} \in X$.

Доказательство

- ▶ Пусть $\mathbf{x}^* \in X$ и выполнено неравенство. Тогда по критерию первого порядка для выпуклой функции $f(\mathbf{y}) \geq f(\mathbf{x}^*)$
- ▶ Пусть \mathbf{x}^* решение задачи (2), но найдётся $\tilde{\mathbf{y}}$ такой что $\langle f'(\mathbf{x}^*), \tilde{\mathbf{y}} - \mathbf{x}^* \rangle < 0$
- ▶ Рассмотрим точку $\mathbf{z}(t) = t\tilde{\mathbf{y}} + (1 - t)\mathbf{x}^*$, $t \in [0, 1]$
- ▶ Тогда в силу $\left. \frac{d}{dt} f(\mathbf{z}(t)) \right|_{t=0} = \langle f'(\mathbf{x}^*), \tilde{\mathbf{y}} - \mathbf{x}^* \rangle < 0$

Дифференциальное условие оптимальности для задачи с ограничениями

$$\min_{\mathbf{x} \in X} f(\mathbf{x}) \quad (2)$$

Теорема

Точка \mathbf{x}^* – решение задачи (2), где f – выпуклая функция, тогда и только тогда, когда $\mathbf{x}^* \in X$ и $\langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x}^* \rangle \geq 0$ для всех $\mathbf{y} \in X$.

Доказательство

- ▶ Пусть $\mathbf{x}^* \in X$ и выполнено неравенство. Тогда по критерию первого порядка для выпуклой функции $f(\mathbf{y}) \geq f(\mathbf{x}^*)$
- ▶ Пусть \mathbf{x}^* решение задачи (2), но найдётся $\tilde{\mathbf{y}}$ такой что $\langle f'(\mathbf{x}^*), \tilde{\mathbf{y}} - \mathbf{x}^* \rangle < 0$
- ▶ Рассмотрим точку $\mathbf{z}(t) = t\tilde{\mathbf{y}} + (1 - t)\mathbf{x}^*$, $t \in [0, 1]$
- ▶ Тогда в силу $\left. \frac{d}{dt} f(\mathbf{z}(t)) \right|_{t=0} = \langle f'(\mathbf{x}^*), \tilde{\mathbf{y}} - \mathbf{x}^* \rangle < 0$
- ▶ Значит для малого t выполнено $f(\mathbf{z}(t)) < f(\mathbf{x}^*)$.
Противоречие.

Задача оптимизации с функциональными ограничениями

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

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

Q: как сформулировать условия оптимальности для задачи в таком виде?

Задача оптимизации с функциональными ограничениями

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

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

Q: как сформулировать условия оптимальности для задачи в таком виде?

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

Нижняя оценка оптимального значения функции

Утверждение

Если $\mu \geq 0$, тогда $p^* \geq g(\lambda, \mu)$

Доказательство

Нижняя оценка оптимального значения функции

Утверждение

Если $\mu \geq 0$, тогда $p^* \geq g(\lambda, \mu)$

Доказательство

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

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

Нижняя оценка оптимального значения функции

Утверждение

Если $\mu \geq 0$, тогда $p^* \geq g(\lambda, \mu)$

Доказательство

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

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

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

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

Нижняя оценка оптимального значения функции

Утверждение

Если $\mu \geq 0$, тогда $p^* \geq g(\lambda, \mu)$

Доказательство

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

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

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

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

Q: что теперь надо сделать с двойственной функцией, чтобы получить наилучшее приближение к p^* ?

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

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
- ▶ $d^* = g(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$
- ▶ 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
- ▶ $d^* = g(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$
- ▶ 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 $(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \text{dom } g$

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, подробнее через несколько слайдов

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, подробнее через несколько слайдов
- ▶ Может выполняться и для **невыпуклых** 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, подробнее через несколько слайдов
- ▶ Может выполняться и для **невыпуклых** задач
- ▶ Решение двойственной задачи даёт решение прямой задачи

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, подробнее через несколько слайдов
- ▶ Может выполняться и для **невыпуклых** задач
- ▶ Решение двойственной задачи даёт решение прямой задачи

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

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, подробнее через несколько слайдов
- ▶ Может выполняться и для **невыпуклых** задач
- ▶ Решение двойственной задачи даёт решение прямой задачи

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

- ▶ Оценка точности решения

Слабая и сильная двойственность

Слабая двойственность: $d^* \leq p^*$

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

Сильная двойственность: $d^* = p^*$

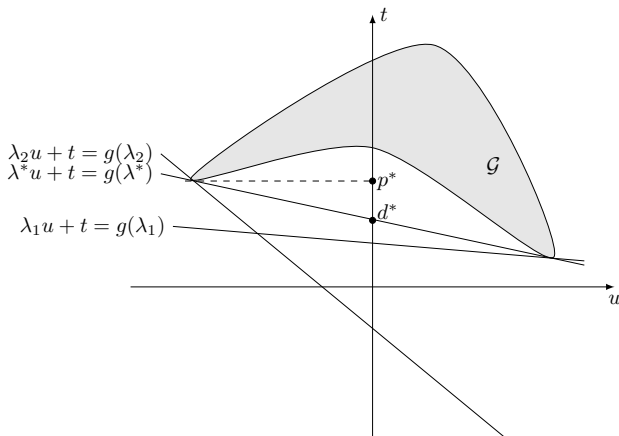
- ▶ В общем случае **НЕ** выполняется
- ▶ *Обычно* выполнена для выпуклых задач
- ▶ Условия регулярности ограничений, подробнее через несколько слайдов
- ▶ Может выполняться и для **невыпуклых** задач
- ▶ Решение двойственной задачи даёт решение прямой задачи

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

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

Геометрическая интерпретация

$$\begin{array}{ll} \min_{\mathbf{x} \in \mathcal{G}} f_0(\mathbf{x}) & g(\lambda) = \inf_{(u,t) \in \mathcal{G}} (t + \lambda u) \\ \text{s.t. } f_1(\mathbf{x}) \leq 0 & \mathcal{G} = \{(f_1(\mathbf{x}), f_0(\mathbf{x})) \mid \mathbf{x} \in \mathcal{G}\} \end{array}$$



Условие Слейтера и сильная двойственность

Условие Слейтера

Говорят, что выполнено условие Слейтера, если

$$\exists \bar{\mathbf{x}} \in \text{int } \mathcal{D} : f_i(\bar{\mathbf{x}}) < 0, \mathbf{A}\bar{\mathbf{x}} = \mathbf{b}$$

Теорема

Сильная двойственность выполняется для выпуклой задачи

$$\begin{aligned} \min f_0(\mathbf{x}) \\ \text{s.t. } f_i(\mathbf{x}) \leq 0, \quad i = 1, \dots, m \\ \mathbf{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{A} \in \mathbb{R}^{p \times n} \end{aligned}$$

если выполнено условие Слейтера.

Доказательство

Предположения

- ▶ $\text{int } \mathcal{D} \neq \emptyset$
- ▶ $\text{rank } \mathbf{A} = p$

Этапы доказательства

1. Введём два множества

$$\mathcal{A} = \{(\mathbf{u}, \mathbf{v}, t) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R} \mid \exists x \in \mathcal{D} : f_i(\mathbf{x}) \leq u_i, \mathbf{Ax} - \mathbf{b} = \mathbf{v}, f_0(\mathbf{x}) \leq t\} \text{ и}$$
$$\mathcal{B} = \{(0, 0, s) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R} \mid s < p^*\}$$

Доказательство

Предположения

- ▶ $\text{int } \mathcal{D} \neq \emptyset$
- ▶ $\text{rank } \mathbf{A} = p$

Этапы доказательства

1. Введём два множества

$$\mathcal{A} = \{(\mathbf{u}, \mathbf{v}, t) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R} \mid \exists x \in \mathcal{D} : f_i(\mathbf{x}) \leq u_i, \mathbf{Ax} - \mathbf{b} = \mathbf{v}, f_0(\mathbf{x}) \leq t\} \text{ и}$$
$$\mathcal{B} = \{(0, 0, s) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R} \mid s < p^*\}$$

2. Наблюдение: $p^* = \inf\{t \mid (0, 0, t) \in \mathcal{A}\}$

Доказательство

Предположения

- ▶ $\text{int } \mathcal{D} \neq \emptyset$
- ▶ $\text{rank } \mathbf{A} = p$

Этапы доказательства

1. Введём два множества

$$\mathcal{A} = \{(\mathbf{u}, \mathbf{v}, t) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R} \mid \exists x \in \mathcal{D} : f_i(\mathbf{x}) \leq u_i, \mathbf{Ax} - \mathbf{b} = \mathbf{v}, f_0(\mathbf{x}) \leq t\} \text{ и}$$

$$\mathcal{B} = \{(0, 0, s) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R} \mid s < p^*\}$$

2. Наблюдение: $p^* = \inf\{t \mid (0, 0, t) \in \mathcal{A}\}$
3. Множества \mathcal{A} и \mathcal{B} выпуклы, как декартово произведение выпуклых множеств

Доказательство

Предположения

- ▶ $\text{int } \mathcal{D} \neq \emptyset$
- ▶ $\text{rank } \mathbf{A} = p$

Этапы доказательства

1. Введём два множества

$$\mathcal{A} = \{(\mathbf{u}, \mathbf{v}, t) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R} \mid \exists x \in \mathcal{D} : f_i(\mathbf{x}) \leq u_i, \mathbf{Ax} - \mathbf{b} = \mathbf{v}, f_0(\mathbf{x}) \leq t\} \text{ и}$$

$$\mathcal{B} = \{(0, 0, s) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R} \mid s < p^*\}$$

2. Наблюдение: $p^* = \inf\{t \mid (0, 0, t) \in \mathcal{A}\}$
3. Множества \mathcal{A} и \mathcal{B} выпуклы, как декартово произведение выпуклых множеств
4. $\mathcal{A} \cap \mathcal{B} = \emptyset$ иначе противоречие с тем, что p^* – минимальное значение f_0

Доказательство

Предположения

- ▶ $\text{int } \mathcal{D} \neq \emptyset$
- ▶ $\text{rank } \mathbf{A} = p$

Этапы доказательства

1. Введём два множества

$$\mathcal{A} = \{(\mathbf{u}, \mathbf{v}, t) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R} \mid \exists x \in \mathcal{D} : f_i(\mathbf{x}) \leq u_i, \mathbf{Ax} - \mathbf{b} = \mathbf{v}, f_0(\mathbf{x}) \leq t\} \text{ и}$$
$$\mathcal{B} = \{(0, 0, s) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R} \mid s < p^*\}$$

2. Наблюдение: $p^* = \inf\{t \mid (0, 0, t) \in \mathcal{A}\}$
3. Множества \mathcal{A} и \mathcal{B} выпуклы, как декартово произведение выпуклых множеств
4. $\mathcal{A} \cap \mathcal{B} = \emptyset$ иначе противоречие с тем, что p^* – минимальное значение f_0
5. По теореме об отделимости существует разделяющая гиперплоскость

6. $(\mathbf{u}, \mathbf{v}, t) \in \mathcal{A} \rightarrow \tilde{\boldsymbol{\lambda}}^\top \mathbf{u} + \tilde{\boldsymbol{\nu}}^\top \mathbf{v} + \mu t \geq \alpha$ (i)
 $(\mathbf{u}, \mathbf{v}, t) \in \mathcal{B} \rightarrow \mu t \leq \alpha$ (ii)

- 6. $(\mathbf{u}, \mathbf{v}, t) \in \mathcal{A} \rightarrow \tilde{\boldsymbol{\lambda}}^\top \mathbf{u} + \tilde{\boldsymbol{\nu}}^\top \mathbf{v} + \mu t \geq \alpha$ (i)
 $(\mathbf{u}, \mathbf{v}, t) \in \mathcal{B} \rightarrow \mu t \leq \alpha$ (ii)
- 7. Так как (ii) выполнено для $t < p^*$, то $\mu p^* \leq \alpha$

- 6. $(\mathbf{u}, \mathbf{v}, t) \in \mathcal{A} \rightarrow \tilde{\boldsymbol{\lambda}}^\top \mathbf{u} + \tilde{\boldsymbol{\nu}}^\top \mathbf{v} + \mu t \geq \alpha$ (i)
 $(\mathbf{u}, \mathbf{v}, t) \in \mathcal{B} \rightarrow \mu t \leq \alpha$ (ii)
- 7. Так как (ii) выполнено для $t < p^*$, то $\mu p^* \leq \alpha$
- 8. Из (i) следует, что $\tilde{\boldsymbol{\lambda}} \geq 0$ и $\mu \geq 0$, иначе выражение слева будет неограничено снизу

6. $(\mathbf{u}, \mathbf{v}, t) \in \mathcal{A} \rightarrow \tilde{\boldsymbol{\lambda}}^\top \mathbf{u} + \tilde{\boldsymbol{\nu}}^\top \mathbf{v} + \mu t \geq \alpha$ (i)
 $(\mathbf{u}, \mathbf{v}, t) \in \mathcal{B} \rightarrow \mu t \leq \alpha$ (ii)
7. Так как (ii) выполнено для $t < p^*$, то $\mu p^* \leq \alpha$
8. Из (i) следует, что $\tilde{\boldsymbol{\lambda}} \geq 0$ и $\mu \geq 0$, иначе выражение слева будет неограничено снизу
9. Перейдём в (i) от записи через \mathcal{A} к записи через \mathbf{x} и ограничения

$$\sum_{i=1}^m \tilde{\lambda}_i f_i(\mathbf{x}) + \tilde{\boldsymbol{\nu}}(\mathbf{A}\mathbf{x} - \mathbf{b}) + \mu f_0(\mathbf{x}) \geq \alpha \geq \mu p^*$$

6. $(\mathbf{u}, \mathbf{v}, t) \in \mathcal{A} \rightarrow \tilde{\boldsymbol{\lambda}}^\top \mathbf{u} + \tilde{\boldsymbol{\nu}}^\top \mathbf{v} + \mu t \geq \alpha$ (i)
 $(\mathbf{u}, \mathbf{v}, t) \in \mathcal{B} \rightarrow \mu t \leq \alpha$ (ii)
7. Так как (ii) выполнено для $t < p^*$, то $\mu p^* \leq \alpha$
8. Из (i) следует, что $\tilde{\boldsymbol{\lambda}} \geq 0$ и $\mu \geq 0$, иначе выражение слева будет неограничено снизу
9. Перейдём в (i) от записи через \mathcal{A} к записи через \mathbf{x} и ограничения

$$\sum_{i=1}^m \tilde{\lambda}_i f_i(\mathbf{x}) + \tilde{\boldsymbol{\nu}}(\mathbf{A}\mathbf{x} - \mathbf{b}) + \mu f_0(\mathbf{x}) \geq \alpha \geq \mu p^*$$

10. Пусть $\mu > 0$ тогда

$$L(\mathbf{x}, \tilde{\boldsymbol{\lambda}}/\mu, \tilde{\boldsymbol{\nu}}/\mu) \geq p^*, \quad \mathbf{x} \in \mathcal{D} \rightarrow \min_{\mathbf{x} \in \mathcal{D}} L = g(\boldsymbol{\lambda}, \boldsymbol{\mu}) \geq p^*$$

6. $(\mathbf{u}, \mathbf{v}, t) \in \mathcal{A} \rightarrow \tilde{\boldsymbol{\lambda}}^\top \mathbf{u} + \tilde{\boldsymbol{\nu}}^\top \mathbf{v} + \mu t \geq \alpha$ (i)
 $(\mathbf{u}, \mathbf{v}, t) \in \mathcal{B} \rightarrow \mu t \leq \alpha$ (ii)
7. Так как (ii) выполнено для $t < p^*$, то $\mu p^* \leq \alpha$
8. Из (i) следует, что $\tilde{\boldsymbol{\lambda}} \geq 0$ и $\mu \geq 0$, иначе выражение слева будет неограничено снизу
9. Перейдём в (i) от записи через \mathcal{A} к записи через \mathbf{x} и ограничения

$$\sum_{i=1}^m \tilde{\lambda}_i f_i(\mathbf{x}) + \tilde{\boldsymbol{\nu}}(\mathbf{A}\mathbf{x} - \mathbf{b}) + \mu f_0(\mathbf{x}) \geq \alpha \geq \mu p^*$$

10. Пусть $\mu > 0$ тогда

$$L(\mathbf{x}, \tilde{\boldsymbol{\lambda}}/\mu, \tilde{\boldsymbol{\nu}}/\mu) \geq p^*, \quad \mathbf{x} \in \mathcal{D} \rightarrow \min_{\mathbf{x} \in \mathcal{D}} L = g(\boldsymbol{\lambda}, \boldsymbol{\mu}) \geq p^*$$

11. Но в силу слабой двойственности $g(\boldsymbol{\lambda}, \boldsymbol{\mu}) \leq p^*$, значит $g(\boldsymbol{\lambda}, \boldsymbol{\mu}) = p^*$ и выполнена сильная двойственность.

12. Пусть $\mu = 0$ тогда

$$\sum_{i=1}^m \tilde{\lambda}_i f_i(\mathbf{x}) + \tilde{\boldsymbol{\nu}}^\top (\mathbf{Ax} - \mathbf{b}) \geq 0 \quad (3)$$

для всех $\mathbf{x} \in \mathcal{D}$

12. Пусть $\mu = 0$ тогда

$$\sum_{i=1}^m \tilde{\lambda}_i f_i(\mathbf{x}) + \tilde{\boldsymbol{\nu}}^\top (\mathbf{Ax} - \mathbf{b}) \geq 0 \quad (3)$$

для всех $\mathbf{x} \in \mathcal{D}$

13. Возьмём $\tilde{\mathbf{x}} \in \text{int } \mathcal{D}$, для которого выполнено условие
Слейтера, тогда $\sum_{i=1}^m \tilde{\lambda}_i f_i(\tilde{\mathbf{x}}) \geq 0$, но $f_i(\tilde{\mathbf{x}}) < 0$ и
 $\tilde{\lambda}_i \geq 0 \rightarrow \tilde{\boldsymbol{\lambda}} = \mathbf{0}$

12. Пусть $\mu = 0$ тогда

$$\sum_{i=1}^m \tilde{\lambda}_i f_i(\mathbf{x}) + \tilde{\boldsymbol{\nu}}^\top (\mathbf{Ax} - \mathbf{b}) \geq 0 \quad (3)$$

для всех $\mathbf{x} \in \mathcal{D}$

13. Возьмём $\tilde{\mathbf{x}} \in \text{int } \mathcal{D}$, для которого выполнено условие Слейтера, тогда $\sum_{i=1}^m \tilde{\lambda}_i f_i(\tilde{\mathbf{x}}) \geq 0$, но $f_i(\tilde{\mathbf{x}}) < 0$ и $\tilde{\lambda}_i \geq 0 \rightarrow \tilde{\boldsymbol{\lambda}} = 0$

14. Из (3) следует, что $\tilde{\boldsymbol{\nu}}^\top (\mathbf{Ax} - \mathbf{b}) \geq 0$ для всех $\mathbf{x} \in \mathcal{D}$, но $\tilde{\boldsymbol{\nu}}^\top (\mathbf{A}\tilde{\mathbf{x}} - \mathbf{b}) = (\mathbf{A}^\top \tilde{\boldsymbol{\nu}})^\top \tilde{\mathbf{x}} - \tilde{\boldsymbol{\nu}}^\top \mathbf{b} = 0$

12. Пусть $\mu = 0$ тогда

$$\sum_{i=1}^m \tilde{\lambda}_i f_i(\mathbf{x}) + \tilde{\boldsymbol{\nu}}^\top (\mathbf{Ax} - \mathbf{b}) \geq 0 \quad (3)$$

для всех $\mathbf{x} \in \mathcal{D}$

13. Возьмём $\tilde{\mathbf{x}} \in \text{int } \mathcal{D}$, для которого выполнено условие

Слейтера, тогда $\sum_{i=1}^m \tilde{\lambda}_i f_i(\tilde{\mathbf{x}}) \geq 0$, но $f_i(\tilde{\mathbf{x}}) < 0$ и

$$\tilde{\lambda}_i \geq 0 \rightarrow \tilde{\boldsymbol{\lambda}} = \mathbf{0}$$

14. Из (3) следует, что $\tilde{\boldsymbol{\nu}}^\top (\mathbf{Ax} - \mathbf{b}) \geq 0$ для всех $\mathbf{x} \in \mathcal{D}$, но $\tilde{\boldsymbol{\nu}}^\top (\mathbf{A}\tilde{\mathbf{x}} - \mathbf{b}) = (\mathbf{A}^\top \tilde{\boldsymbol{\nu}})^\top \tilde{\mathbf{x}} - \tilde{\boldsymbol{\nu}}^\top \mathbf{b} = 0$

15. Значит если $\mathbf{A}^\top \tilde{\boldsymbol{\nu}} \neq \mathbf{0}$, то найдётся точка $\bar{\mathbf{x}}$, в которой $\tilde{\boldsymbol{\nu}}^\top (\mathbf{A}\bar{\mathbf{x}} - \mathbf{b}) < 0$, противоречие

12. Пусть $\mu = 0$ тогда

$$\sum_{i=1}^m \tilde{\lambda}_i f_i(\mathbf{x}) + \tilde{\boldsymbol{\nu}}^\top (\mathbf{Ax} - \mathbf{b}) \geq 0 \quad (3)$$

для всех $\mathbf{x} \in \mathcal{D}$

13. Возьмём $\tilde{\mathbf{x}} \in \text{int } \mathcal{D}$, для которого выполнено условие

Слейтера, тогда $\sum_{i=1}^m \tilde{\lambda}_i f_i(\tilde{\mathbf{x}}) \geq 0$, но $f_i(\tilde{\mathbf{x}}) < 0$ и

$$\tilde{\lambda}_i \geq 0 \rightarrow \tilde{\boldsymbol{\lambda}} = 0$$

14. Из (3) следует, что $\tilde{\boldsymbol{\nu}}^\top (\mathbf{Ax} - \mathbf{b}) \geq 0$ для всех $\mathbf{x} \in \mathcal{D}$, но $\tilde{\boldsymbol{\nu}}^\top (\mathbf{A}\tilde{\mathbf{x}} - \mathbf{b}) = (\mathbf{A}^\top \tilde{\boldsymbol{\nu}})^\top \tilde{\mathbf{x}} - \tilde{\boldsymbol{\nu}}^\top \mathbf{b} = 0$

15. Значит если $\mathbf{A}^\top \tilde{\boldsymbol{\nu}} \neq 0$, то найдётся точка $\bar{\mathbf{x}}$, в которой $\tilde{\boldsymbol{\nu}}^\top (\mathbf{A}\bar{\mathbf{x}} - \mathbf{b}) < 0$, противоречие

16. Если $\mathbf{A}^\top \tilde{\boldsymbol{\nu}} = 0$ и $\tilde{\boldsymbol{\nu}} \neq 0$, то $\text{rank } \mathbf{A} < p$, противоречие.
Значит $\mu \neq 0$.

Дополняющая нежёсткость

Пусть выполнена сильная двойственность,

x^* – решение прямой задачи,

(λ^*, μ^*) – решение двойственной задачи, тогда

Дополняющая нежёсткость

Пусть выполнена сильная двойственность,

\mathbf{x}^* – решение прямой задачи,

$(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ – решение двойственной задачи, тогда

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

Дополняющая нежёсткость

Пусть выполнена сильная двойственность,

\mathbf{x}^* – решение прямой задачи,

$(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ – решение двойственной задачи, тогда

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

► \mathbf{x}^* минимизирует $L(\mathbf{x}, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$

Дополняющая нежёсткость

Пусть выполнена сильная двойственность,

\mathbf{x}^* – решение прямой задачи,

$(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ – решение двойственной задачи, тогда

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

- ▶ \mathbf{x}^* минимизирует $L(\mathbf{x}, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$
- ▶ Условие дополняющей нежёсткости

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

$$\mu_j^* > 0 \Rightarrow h_j(\mathbf{x}^*) = 0, \quad h_j(\mathbf{x}^*) < 0 \Rightarrow \mu_j^* = 0$$

Условия Каруша-Куна-Таккера (ККТ): невыпуклые задачи

Пусть $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ решения прямой и двойственной задачи и выполнена сильная двойственность, тогда

Условия Каруша-Куна-Таккера (ККТ): невыпуклые задачи

Пусть $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ решения прямой и двойственной задачи и выполнена сильная двойственность, тогда

1. $h_j(\mathbf{x}^*) \leq 0$

Условия Каруша-Куна-Таккера (ККТ): невыпуклые задачи

Пусть $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ решения прямой и двойственной задачи и выполнена сильная двойственность, тогда

1. $h_j(\mathbf{x}^*) \leq 0$
2. $g_i(\mathbf{x}^*) = 0$

Условия Каруша-Куна-Таккера (ККТ): невыпуклые задачи

Пусть $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ решения прямой и двойственной задачи и выполнена сильная двойственность, тогда

1. $h_j(\mathbf{x}^*) \leq 0$
2. $g_i(\mathbf{x}^*) = 0$
3. $\boldsymbol{\mu}^* \geq 0$

Условия Каруша-Куна-Таккера (ККТ): невыпуклые задачи

Пусть $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ решения прямой и двойственной задачи и выполнена сильная двойственность, тогда

1. $h_j(\mathbf{x}^*) \leq 0$
2. $g_i(\mathbf{x}^*) = 0$
3. $\boldsymbol{\mu}^* \geq 0$
4. $\mu_j^* h_j(\mathbf{x}^*) = 0$

Условия Каруша-Куна-Таккера (ККТ): невыпуклые задачи

Пусть $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ решения прямой и двойственной задачи и выполнена сильная двойственность, тогда

1. $h_j(\mathbf{x}^*) \leq 0$
2. $g_i(\mathbf{x}^*) = 0$
3. $\boldsymbol{\mu}^* \geq 0$
4. $\mu_j^* h_j(\mathbf{x}^*) = 0$
5. $f'_0(\mathbf{x}^*) + \sum_{i=1}^m \lambda_i^* g'_i(\mathbf{x}^*) + \sum_{j=1}^p \mu_j^* h'_j(\mathbf{x}^*) = 0$

Условия Каруша-Куна-Таккера (ККТ): невыпуклые задачи

Пусть $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ решения прямой и двойственной задачи и выполнена сильная двойственность, тогда

1. $h_j(\mathbf{x}^*) \leq 0$
2. $g_i(\mathbf{x}^*) = 0$
3. $\boldsymbol{\mu}^* \geq 0$
4. $\mu_j^* h_j(\mathbf{x}^*) = 0$
5. $f'_0(\mathbf{x}^*) + \sum_{i=1}^m \lambda_i^* g'_i(\mathbf{x}^*) + \sum_{j=1}^p \mu_j^* h'_j(\mathbf{x}^*) = 0$

Последнее равенство выполнено в силу

$$\min_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) = L(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$$

и необходимого условия минимума.

Условия Каруша-Куна-Таккера (ККТ): невыпуклые задачи

Пусть $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ решения прямой и двойственной задачи и выполнена сильная двойственность, тогда

1. $h_j(\mathbf{x}^*) \leq 0$
2. $g_i(\mathbf{x}^*) = 0$
3. $\boldsymbol{\mu}^* \geq 0$
4. $\mu_j^* h_j(\mathbf{x}^*) = 0$
5. $f'_0(\mathbf{x}^*) + \sum_{i=1}^m \lambda_i^* g'_i(\mathbf{x}^*) + \sum_{j=1}^p \mu_j^* h'_j(\mathbf{x}^*) = 0$

Последнее равенство выполнено в силу

$$\min_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) = L(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$$

и необходимого условия минимума.

Замечание

Сначала эти условия были известны как условия Куна-Таккера (работа 1951 г.). Потом обнаружили, что Вильям Каруш вывел их в своей дипломной работе 1939 г.

ККТ для выпуклых задач

Утверждение 1

Пусть прямая задача выпукла (f_0, h_j – выпуклы, g_i – аффинны) и для $(\hat{x}, \hat{\lambda}, \hat{\mu})$ выполнены условия ККТ, тогда

- ▶ выполнена сильная двойственность
- ▶ $(\hat{x}, \hat{\lambda}, \hat{\mu})$ – решения прямой и двойственной задач

ККТ для выпуклых задач

Утверждение 1

Пусть прямая задача выпукла (f_0, h_j – выпуклы, g_i – аффинны) и для $(\hat{x}, \hat{\lambda}, \hat{\mu})$ выполнены условия ККТ, тогда

- ▶ выполнена сильная двойственность
- ▶ $(\hat{x}, \hat{\lambda}, \hat{\mu})$ – решения прямой и двойственной задач

Доказательство

ККТ для выпуклых задач

Утверждение 1

Пусть прямая задача выпукла (f_0, h_j – выпуклы, g_i – аффинны) и для $(\hat{x}, \hat{\lambda}, \hat{\mu})$ выполнены условия ККТ, тогда

- ▶ выполнена сильная двойственность
- ▶ $(\hat{x}, \hat{\lambda}, \hat{\mu})$ – решения прямой и двойственной задач

Доказательство

- ▶ Первые два условия из ККТ $\rightarrow \hat{x}$ лежит в допустимом множестве

ККТ для выпуклых задач

Утверждение 1

Пусть прямая задача выпукла (f_0, h_j – выпуклы, g_i – аффинны) и для $(\hat{x}, \hat{\lambda}, \hat{\mu})$ выполнены условия ККТ, тогда

- ▶ выполнена сильная двойственность
- ▶ $(\hat{x}, \hat{\lambda}, \hat{\mu})$ – решения прямой и двойственной задач

Доказательство

- ▶ Первые два условия из ККТ $\rightarrow \hat{x}$ лежит в допустимом множестве
- ▶ $\hat{\mu} \geq 0 \rightarrow L(x, \hat{\lambda}, \hat{\mu})$ выпуклый по x

ККТ для выпуклых задач

Утверждение 1

Пусть прямая задача выпукла (f_0, h_j – выпуклы, g_i – аффинны) и для $(\hat{\mathbf{x}}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ выполнены условия ККТ, тогда

- ▶ выполнена сильная двойственность
- ▶ $(\hat{\mathbf{x}}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ – решения прямой и двойственной задач

Доказательство

- ▶ Первые два условия из ККТ $\rightarrow \hat{\mathbf{x}}$ лежит в допустимом множестве
- ▶ $\hat{\boldsymbol{\mu}} \geq 0 \rightarrow L(\mathbf{x}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ выпуклый по x
- ▶ Последнее условие $\rightarrow \hat{\mathbf{x}}$ минимизирует L

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

ККТ для выпуклых задач

Утверждение 1

Пусть прямая задача выпукла (f_0, h_j – выпуклы, g_i – аффинны) и для $(\hat{x}, \hat{\lambda}, \hat{\mu})$ выполнены условия ККТ, тогда

- ▶ выполнена сильная двойственность
- ▶ $(\hat{x}, \hat{\lambda}, \hat{\mu})$ – решения прямой и двойственной задач

Доказательство

- ▶ Первые два условия из ККТ $\rightarrow \hat{x}$ лежит в допустимом множестве
- ▶ $\hat{\mu} \geq 0 \rightarrow L(x, \hat{\lambda}, \hat{\mu})$ выпуклый по x
- ▶ Последнее условие $\rightarrow \hat{x}$ минимизирует L

$$g(\hat{\lambda}, \hat{\mu}) = L(\hat{x}, \hat{\lambda}, \hat{\mu}) = f_0(\hat{x}) + \sum_{i=1}^m \hat{\lambda}_i g_i(\hat{x}) + \sum_{j=1}^p \hat{\mu}_j h_j(\hat{x}) = f_0(\hat{x})$$

- ▶ Выполнена сильная двойственность

ККТ для выпуклых задач

Утверждение 1

Пусть прямая задача выпукла (f_0, h_j – выпуклы, g_i – аффинны) и для $(\hat{x}, \hat{\lambda}, \hat{\mu})$ выполнены условия ККТ, тогда

- ▶ выполнена сильная двойственность
- ▶ $(\hat{x}, \hat{\lambda}, \hat{\mu})$ – решения прямой и двойственной задач

Доказательство

- ▶ Первые два условия из ККТ $\rightarrow \hat{x}$ лежит в допустимом множестве
- ▶ $\hat{\mu} \geq 0 \rightarrow L(x, \hat{\lambda}, \hat{\mu})$ выпуклый по x
- ▶ Последнее условие $\rightarrow \hat{x}$ минимизирует L

$$g(\hat{\lambda}, \hat{\mu}) = L(\hat{x}, \hat{\lambda}, \hat{\mu}) = f_0(\hat{x}) + \sum_{i=1}^m \hat{\lambda}_i g_i(\hat{x}) + \sum_{j=1}^p \hat{\mu}_j h_j(\hat{x}) = f_0(\hat{x})$$

- ▶ Выполнена сильная двойственность
- ▶ $(\hat{x}, \hat{\lambda}, \hat{\mu})$ – решения прямой и двойственной задач

ККТ для выпуклых задач

Утверждение 2

Пусть для выпуклой задачи выполнено условие Слейтера.

Тогда x^* решение прямой задачи тогда и только тогда, когда существуют (λ, μ) такие, что для них выполнены условия ККТ

ККТ для выпуклых задач

Утверждение 2

Пусть для выпуклой задачи выполнено условие Слейтера.

Тогда x^* решение прямой задачи тогда и только тогда, когда существуют (λ, μ) такие, что для них выполнены условия ККТ

Доказательство

ККТ для выпуклых задач

Утверждение 2

Пусть для выпуклой задачи выполнено условие Слейтера.
Тогда x решение прямой задачи тогда и только тогда, когда существуют (λ, μ) такие, что для них выполнены условия ККТ

Доказательство

- Из выпуклости и условий Слейтера следует выполнение сильной двойственности и достижимость минимума p^*

ККТ для выпуклых задач

Утверждение 2

Пусть для выпуклой задачи выполнено условие Слейтера.
Тогда x^* решение прямой задачи тогда и только тогда, когда существуют (λ, μ) такие, что для них выполнены условия ККТ

Доказательство

- ▶ Из выпуклости и условий Слейтера следует выполнение сильной двойственности и достижимость минимума p^*
- ▶ Необходимость выполнения ККТ следует из утверждения для общего случая

ККТ для выпуклых задач

Утверждение 2

Пусть для выпуклой задачи выполнено условие Слейтера.
Тогда x решение прямой задачи тогда и только тогда, когда существуют (λ, μ) такие, что для них выполнены условия ККТ

Доказательство

- ▶ Из выпуклости и условий Слейтера следует выполнение сильной двойственности и достижимость минимума p^*
- ▶ Необходимость выполнения ККТ следует из утверждения для общего случая
- ▶ Достаточность следует из утверждения 1

- ▶ Условие оптимальности для задачи безусловной оптимизации

- ▶ Условие оптимальности для задачи безусловной оптимизации
- ▶ Дифференциальный условий оптимальности для задачи условной оптимизации

- ▶ Условие оптимальности для задачи безусловной оптимизации
- ▶ Дифференциальный условий оптимальности для задачи условной оптимизации
- ▶ Двойственная функция и двойственная задача

- ▶ Условие оптимальности для задачи безусловной оптимизации
- ▶ Дифференциальный условий оптимальности для задачи условной оптимизации
- ▶ Двойственная функция и двойственная задача
- ▶ Сильная двойственность и условие Слейтера

- ▶ Условие оптимальности для задачи безусловной оптимизации
- ▶ Дифференциальный условий оптимальности для задачи условной оптимизации
- ▶ Двойственная функция и двойственная задача
- ▶ Сильная двойственность и условие Слейтера
- ▶ Условия Каруша-Куна-Таккера