Методы оптимизации
Лекция 9: Условия Каруша-Куна-Таккера

Александр Катруца

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

9 ноября 2020 г.

На прошлой лекции

- Преобразования задач и их типы

На прошлой лекции

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

На прошлой лекции

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

На прошлой лекции

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

На прошлой лекции

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

План на эту лекцию

- Теорема о сильной двойственности для выпуклых задач
- Геометрическая интерпретация условий ККТ
- Условия ККТ для задач с обобщёнными неравенствами

Задача оптимизации с функциональными ограничениями

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

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

Задача оптимизации с функциональными ограничениями

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

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

Основная цель

Сформулировать условия оптимальности для таких задач

Задача оптимизации с функциональными ограничениями

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

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

Основная цель

Сформулировать условия оптимальности для таких задач

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

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

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

Основные результаты с прошлой лекции

Теорема

Если $\hat{\mathbf{x}}$ лежит в допустимом множестве и найдутся допустимые $(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ такие что $\boxed{f_0(\hat{\mathbf{x}}) = g(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})}$, то $\hat{\mathbf{x}}$ является решением задачи.

Основные результаты с прошлой лекции

Теорема

Если $\hat{\mathbf{x}}$ лежит в допустимом множестве и найдутся допустимые $(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ такие что $\boxed{f_0(\hat{\mathbf{x}}) = g(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})}$, то $\hat{\mathbf{x}}$ является решением задачи.

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 $\hat{\mu}_j h_j(\hat{\mathbf{x}}) = 0$ для всех $j = 1, \ldots, p$.

Условия Каруша-Куна-Таккера (ККТ)

Следствие

Пусть $\mathbf{x}^*$ решение задачи минимизации такое, что $\mathbf{x}^* \in \mathrm{int}(\mathcal{D})$, 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
- $\hat{\boldsymbol{\mu}} \geq 0$
- $h_j(\mathbf{x}^*) \leq 0$
- $g_i(\mathbf{x}^*) = 0$

Задача выпуклой оптимизации

$$
\begin{aligned}
&\min_{\mathbf{x} \in \mathcal{D}} f_0(\mathbf{x}) \\
\text{s.t. } & f_i(\mathbf{x}) \leq 0, \; i = 1, \dots, p \\
& \mathbf{A}\mathbf{x} = \mathbf{b}
\end{aligned}
$$

Задача выпуклой оптимизации

$$
\begin{aligned}
&\min_{\mathbf{x} \in \mathcal{D}} f_0(\mathbf{x}) \\
\text{s.t. } & f_i(\mathbf{x}) \leq 0, \ i = 1, \dots, p \\
& \mathbf{A}\mathbf{x} = \mathbf{b}
\end{aligned}
$$

- $f_0, f_1, \dots, f_p$ — выпуклые функции

Задача выпуклой оптимизации

$$\begin{aligned} &\min_{\mathbf{x}\in\mathcal{D}} f_0(\mathbf{x}) \\ \text{s.t. } &f_i(\mathbf{x}) \leq 0,\ i = 1, \ldots, p \\ &\mathbf{A}\mathbf{x} = \mathbf{b} \end{aligned}$$

- $f_0, f_1, \ldots, f_p$ — выпуклые функции
- ограничение $\mathbf{A}\mathbf{x} = \mathbf{b}$ задаёт некоторое аффинное множество

Задача выпуклой оптимизации

$$
\begin{aligned}
&\min_{\mathbf{x} \in \mathcal{D}} f_0(\mathbf{x}) \\
\text{s.t. } & f_i(\mathbf{x}) \le 0, \; i = 1, \dots, p \\
& \mathbf{A}\mathbf{x} = \mathbf{b}
\end{aligned}
$$

- $f_0, f_1, \dots, f_p$ — выпуклые функции
- ограничение $\mathbf{A}\mathbf{x} = \mathbf{b}$ задаёт некоторое аффинное множество
- ограничения $f_i(\mathbf{x}) \le 0$, где $i = 1, \dots, p$ задают некоторое выпуклое множество (проверьте!)

Условие регулярности ограничений для выпуклых задач

Условие Слейтера

Говорят, что для выпуклой задачи выполнено условие Слейтера, если существует точка $\bar{\mathbf{x}} \in \operatorname{relint}(\mathcal{D})$ такая что

- $\mathbf{A}\bar{\mathbf{x}} = \mathbf{b}$
- $f_i(\bar{\mathbf{x}}) < 0$ для всех неаффинных ограничений типа неравенств.

Это означает, что найдётся хотя бы одна **внутренняя точка** в допустимом множестве

Основной результат на сегодня

Теорема

Если задача оптимизации выпукла, выполнено условие Слейтера и p^* конечно, тогда выполнена сильная двойственность, то есть найдётся пара допустимых $(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ таких что $g(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}) = p^*$

Основной результат на сегодня

Теорема

Если задача оптимизации выпукла, выполнено условие Слейтера и p^* конечно, тогда выполнена сильная двойственность, то есть найдётся пара допустимых $(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ таких что $g(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}) = p^*$

Замечание

Проверьте, что в примере с прошлой лекции с положительным зазором двойственности условие Слейтера **не выполняется**.

Вспомогательное утверждение

Теорема

Пусть дано выпуклое множество $\mathcal{X}$, на котором определены выпуклые функции f_i, где $i = 1, \ldots, p; p \geq 1$ и аффинные функции g_j, где $j = 1, \ldots, m; m \geq 0$, то есть $g_j(\mathbf{x}) = \mathbf{a}_j^\top \mathbf{x} + b_j$, а также множество $\{\mathbf{x} \in \mathrm{relint}\,(\mathcal{X}) \mid g_j(\mathbf{x}) \leq 0\}$ непусто. Тогда следующие утверждения эквивалентны

1) Система $\begin{cases} f_i(\mathbf{x}) < 0, \ i = 1, \ldots, p \\ g_j(\mathbf{x}) \leq 0, \ j = 1, \ldots, m \end{cases}$ не имеет решения среди $\mathbf{x} \in \mathcal{X}$
2) Существует набор неотрицательных чисел $\mu_1, \ldots \mu_p$, среди которых есть хотя бы один не ноль, и набор неотрицательных чисел $\lambda_1, \ldots, \lambda_m$, такие что $\sum_{i=1}^p \mu_i f_i(\mathbf{x}) + \sum_{j=1}^m \lambda_j g_j(\mathbf{x}) \geq 0$ для всех $\mathbf{x} \in \mathcal{X}$.

Доказательство

- Если система из 1) имеет решение $\hat{\mathbf{x}}$, тогда сумма в утверждении 2) будет отрицательной в этой точке, поскольку коэффициенты неотрицательные.

Доказательство

- Если система из 1) имеет решение $\hat{\mathbf{x}}$, тогда сумма в утверждении 2) будет отрицательной в этой точке, поскольку коэффициенты неотрицательные.
- Таким образом, 2) $\to$ 1)

Доказательство

- Если система из 1) имеет решение $\hat{\mathbf{x}}$, тогда сумма в утверждении 2) будет отрицательной в этой точке, поскольку коэффициенты неотрицательные.
- Таким образом, 2) $\to$ 1)
- Пусть система из 1) не имеет решения. Зададим множество $\mathcal{M} \subset \mathbb{R}^{m+p}$ такое что если $\mathbf{y} \in \mathcal{M}$, то система $\begin{cases} f_i(\mathbf{x}) < y_i, \ i = 1, \dots, p \\ g_j(\mathbf{x}) = y_{p+j}, \ j = 1, \dots, m \end{cases}$ имеет решение на $\mathcal{X}$

Доказательство

- Если система из 1) имеет решение $\hat{\mathbf{x}}$, тогда сумма в утверждении 2) будет отрицательной в этой точке, поскольку коэффициенты неотрицательные.
- Таким образом, 2) $\to$ 1)
- Пусть система из 1) не имеет решения. Зададим множество $\mathcal{M} \subset \mathbb{R}^{m+p}$ такое что если $\mathbf{y} \in \mathcal{M}$, то система $\begin{cases} f_i(\mathbf{x}) < y_i, \ i = 1, \ldots, p \\ g_j(\mathbf{x}) = y_{p+j}, \ j = 1, \ldots, m \end{cases}$ имеет решение на $\mathcal{X}$
- Множество $\mathcal{M}$ является выпуклым (проверьте!)

Доказательство

- Если система из 1) имеет решение $\hat{\mathbf{x}}$, тогда сумма в утверждении 2) будет отрицательной в этой точке, поскольку коэффициенты неотрицательные.
- Таким образом, 2) $\to$ 1)
- Пусть система из 1) не имеет решения. Зададим множество $\mathcal{M} \subset \mathbb{R}^{m+p}$ такое что если $\mathbf{y} \in \mathcal{M}$, то система
$$\begin{cases} f_i(\mathbf{x}) < y_i, \; i = 1, \dots, p \\ g_j(\mathbf{x}) = y_{p+j}, \; j = 1, \dots, m \end{cases} \quad \text{имеет решение на } \mathcal{X}$$
- Множество $\mathcal{M}$ является выпуклым (проверьте!)
- Наше предположение об отсутствии решения у системы из 1) означает, что $\mathcal{M} \cap \mathbb{R}^{m+p}_{-} = \varnothing$, где $\mathbb{R}^{m+p}_{-} = \{\mathbf{z} \mid z_i \leq 0\}$

Доказательство

- Если система из 1) имеет решение $\hat{\mathbf{x}}$, тогда сумма в утверждении 2) будет отрицательной в этой точке, поскольку коэффициенты неотрицательные.
- Таким образом, 2) $\to$ 1)
- Пусть система из 1) не имеет решения. Зададим множество $\mathcal{M} \subset \mathbb{R}^{m+p}$ такое что если $\mathbf{y} \in \mathcal{M}$, то система $\begin{cases} f_i(\mathbf{x}) < y_i, \ i = 1, \ldots, p \\ g_j(\mathbf{x}) = y_{p+j}, \ j = 1, \ldots, m \end{cases}$ имеет решение на $\mathcal{X}$
- Множество $\mathcal{M}$ является выпуклым (проверьте!)
- Наше предположение об отсутствии решения у системы из 1) означает, что $\mathcal{M} \cap \mathbb{R}^{m+p}_{-} = \varnothing$, где $\mathbb{R}^{m+p}_{-} = \{\mathbf{z} \mid z_i \leq 0\}$
- Значит найдётся разделяющая гиперплоскость, такая что $\langle \mathbf{q}, \mathbf{y} \rangle \geq \alpha$ для $\mathbf{y} \in \mathcal{M}$ и $\langle \mathbf{q}, \mathbf{z} \rangle \leq \alpha$ для $\mathbf{z} \in \mathbb{R}^{m+p}_{-}$

- Так как $0 \in \mathbb{R}_{-}^{m+p}$, то $\langle \mathbf{q}, 0 \rangle = 0 \leq \alpha$

- Так как $0 \in \mathbb{R}^{m+p}_{-}$, то $\langle \mathbf{q}, 0 \rangle = 0 \leq \alpha$
- Если взять для каждого $i = 1, \ldots, m+p$ и $\mathbf{z} = t\mathbf{e}_i$ и перейти к пределу $t \to -\infty$, то окажется что $\mathbf{q} \geq 0$

- Так как $0 \in \mathbb{R}^{m+p}_{-}$, то $\langle \mathbf{q}, 0 \rangle = 0 \le \alpha$
- Если взять для каждого $i = 1, \ldots, m+p$ и $\mathbf{z} = t\mathbf{e}_i$ и перейти к пределу $t \to -\infty$, то окажется что $\mathbf{q} \ge 0$
- Так как для всех $\mathbf{x} \in \mathcal{X}$ и произвольно малого $\varepsilon > 0$ вектор $\bar{\mathbf{y}} = (f_1(\mathbf{x}) + \varepsilon, \ldots, f_p(\mathbf{x}) + \varepsilon, g_1(\mathbf{x}), \ldots, g_m(\mathbf{x})) \in \mathcal{M}$, то $\langle \mathbf{q}, \bar{\mathbf{y}} \rangle \ge \alpha \ge 0$

- Так как $0 \in \mathbb{R}_-^{m+p}$, то $\langle \mathbf{q}, 0 \rangle = 0 \le \alpha$
- Если взять для каждого $i = 1, \ldots, m + p$ и $\mathbf{z} = t\mathbf{e}_i$ и перейти к пределу $t \to -\infty$, то окажется что $\mathbf{q} \ge 0$
- Так как для всех $\mathbf{x} \in \mathcal{X}$ и произвольно малого $\varepsilon > 0$ вектор $\bar{\mathbf{y}} = (f_1(\mathbf{x}) + \varepsilon, \ldots, f_p(\mathbf{x}) + \varepsilon, g_1(\mathbf{x}), \ldots, g_m(\mathbf{x})) \in \mathcal{M}$, то $\langle \mathbf{q}, \bar{\mathbf{y}} \rangle \ge \alpha \ge 0$
- Переходя к пределу для $\varepsilon \to 0$, получим для всех $\mathbf{x} \in \mathcal{X}$

$$\sum_{i=1}^{p} q_i f_i(\mathbf{x}) + \sum_{j=1}^{m} q_{p+j} g_j(\mathbf{x}) \ge 0$$

- Так как $0 \in \mathbb{R}_-^{m+p}$, то $\langle \mathbf{q}, 0 \rangle = 0 \leq \alpha$
- Если взять для каждого $i = 1, \dots, m+p$ и $\mathbf{z} = t\mathbf{e}_i$ и перейти к пределу $t \to -\infty$, то окажется что $\mathbf{q} \geq 0$
- Так как для всех $\mathbf{x} \in \mathcal{X}$ и произвольно малого $\varepsilon > 0$ вектор $\bar{\mathbf{y}} = (f_1(\mathbf{x}) + \varepsilon, \dots, f_p(\mathbf{x}) + \varepsilon, g_1(\mathbf{x}), \dots, g_m(\mathbf{x})) \in \mathcal{M}$, то $\langle \mathbf{q}, \bar{\mathbf{y}} \rangle \geq \alpha \geq 0$
- Переходя к пределу для $\varepsilon \to 0$, получим для всех $\mathbf{x} \in \mathcal{X}$

$$\sum_{i=1}^{p} q_i f_i(\mathbf{x}) + \sum_{j=1}^{m} q_{p+j} g_j(\mathbf{x}) \geq 0$$

- Если $m = 0$, теорема доказана, так как все компоненты вектора $\mathbf{q}$ неотрицательны.

- Так как $0 \in \mathbb{R}_{-}^{m+p}$, то $\langle \mathbf{q}, 0 \rangle = 0 \leq \alpha$
- Если взять для каждого $i = 1, \ldots, m+p$ и $\mathbf{z} = t\mathbf{e}_i$ и перейти к пределу $t \to -\infty$, то окажется что $\mathbf{q} \geq 0$
- Так как для всех $\mathbf{x} \in \mathcal{X}$ и произвольно малого $\varepsilon > 0$ вектор $\bar{\mathbf{y}} = (f_1(\mathbf{x}) + \varepsilon, \ldots, f_p(\mathbf{x}) + \varepsilon, g_1(\mathbf{x}), \ldots, g_m(\mathbf{x})) \in \mathcal{M}$, то $\langle \mathbf{q}, \bar{\mathbf{y}} \rangle \geq \alpha \geq 0$
- Переходя к пределу для $\varepsilon \to 0$, получим для всех $\mathbf{x} \in \mathcal{X}$

$$\sum_{i=1}^{p} q_i f_i(\mathbf{x}) + \sum_{j=1}^{m} q_{p+j} g_j(\mathbf{x}) \geq 0$$

- Если $m = 0$, теорема доказана, так как все компоненты вектора $\mathbf{q}$ неотрицательны.
- Если $m > 0$, покажем, что среди коэффициентов $q_1, \ldots, q_p$ найдётся ненулевой

- Так как $0 \in \mathbb{R}^{m+p}_{-}$, то $\langle \mathbf{q}, 0 \rangle = 0 \leq \alpha$
- Если взять для каждого $i = 1, \ldots, m+p$ и $\mathbf{z} = t\mathbf{e}_i$ и перейти к пределу $t \to -\infty$, то окажется что $\mathbf{q} \geq 0$
- Так как для всех $\mathbf{x} \in \mathcal{X}$ и произвольно малого $\varepsilon > 0$ вектор $\bar{\mathbf{y}} = (f_1(\mathbf{x}) + \varepsilon, \ldots, f_p(\mathbf{x}) + \varepsilon, g_1(\mathbf{x}), \ldots, g_m(\mathbf{x})) \in \mathcal{M}$, то $\langle \mathbf{q}, \bar{\mathbf{y}} \rangle \geq \alpha \geq 0$
- Переходя к пределу для $\varepsilon \to 0$, получим для всех $\mathbf{x} \in \mathcal{X}$

$$\sum_{i=1}^{p} q_i f_i(\mathbf{x}) + \sum_{j=1}^{m} q_{p+j} g_j(\mathbf{x}) \geq 0$$

- Если $m = 0$, теорема доказана, так как все компоненты вектора $\mathbf{q}$ неотрицательны.
- Если $m > 0$, покажем, что среди коэффициентов $q_1, \ldots, q_p$ найдётся ненулевой
- Пусть это не так, то есть $q_1 = \ldots = q_p = 0$, тогда рассмотрим функцию $h(\mathbf{x}) = \sum_{j=1}^{m} q_{p+j} g_j(\mathbf{x})$

- Так как $0 \in \mathbb{R}^{m+p}_{-}$, то $\langle \mathbf{q}, 0 \rangle = 0 \leq \alpha$
- Если взять для каждого $i = 1, \ldots, m+p$ и $\mathbf{z} = t\mathbf{e}_i$ и перейти к пределу $t \to -\infty$, то окажется что $\mathbf{q} \geq 0$
- Так как для всех $\mathbf{x} \in \mathcal{X}$ и произвольно малого $\varepsilon > 0$ вектор $\bar{\mathbf{y}} = (f_1(\mathbf{x}) + \varepsilon, \ldots, f_p(\mathbf{x}) + \varepsilon, g_1(\mathbf{x}), \ldots, g_m(\mathbf{x})) \in \mathcal{M}$, то $\langle \mathbf{q}, \bar{\mathbf{y}} \rangle \geq \alpha \geq 0$
- Переходя к пределу для $\varepsilon \to 0$, получим для всех $\mathbf{x} \in \mathcal{X}$

$$\sum_{i=1}^{p} q_i f_i(\mathbf{x}) + \sum_{j=1}^{m} q_{p+j} g_j(\mathbf{x}) \geq 0$$

- Если $m = 0$, теорема доказана, так как все компоненты вектора $\mathbf{q}$ неотрицательны.
- Если $m > 0$, покажем, что среди коэффициентов $q_1, \ldots, q_p$ найдётся ненулевой
- Пусть это не так, то есть $q_1 = \ldots = q_p = 0$, тогда рассмотрим функцию $h(\mathbf{x}) = \sum_{j=1}^{m} q_{p+j} g_j(\mathbf{x})$
- Она является аффинной и $h(\mathbf{x}) \geq 0$ для $\mathbf{x} \in \mathcal{X}$

- Но по условию теоремы существует точка $\mathbf{x}_0$ из $\operatorname{relint}(\mathcal{X})$ такая что $g_j(\mathbf{x}_0) \leq 0$ для $j = 1, \dots, m$. Значит $h(\mathbf{x}_0) \leq 0$, тогда $h(\mathbf{x}_0) = 0$

- Но по условию теоремы существует точка $\mathbf{x}_0$ из $\operatorname{relint}(\mathcal{X})$ такая что $g_j(\mathbf{x}_0) \leq 0$ для $j = 1, \ldots, m$. Значит $h(\mathbf{x}_0) \leq 0$, тогда $h(\mathbf{x}_0) = 0$
- Итак, $h(\mathbf{x})$ вогнутая функция (поскольку аффинная), $h(\mathbf{x}) \geq 0$ и $h(\mathbf{x}_0) = 0$, то есть вогнутая функция достигает своего минимума в точке $\mathbf{x}_0$ и он равен 0

- Но по условию теоремы существует точка $\mathbf{x}_0$ из $\mathrm{relint}\,(\mathcal{X})$ такая что $g_j(\mathbf{x}_0) \leq 0$ для $j = 1, \ldots, m$. Значит $h(\mathbf{x}_0) \leq 0$, тогда $h(\mathbf{x}_0) = 0$
- Итак, $h(\mathbf{x})$ вогнутая функция (поскольку аффинная), $h(\mathbf{x}) \geq 0$ и $h(\mathbf{x}_0) = 0$, то есть вогнутая функция достигает своего минимума в точке $\mathbf{x}_0$ и он равен 0
- Это означает, что $h(\mathbf{x}) = 0$ для всех $\mathbf{x} \in \mathcal{X}$ (докажите!)

- Но по условию теоремы существует точка $\mathbf{x}_0$ из $\operatorname{relint}(\mathcal{X})$ такая что $g_j(\mathbf{x}_0) \leq 0$ для $j = 1, \ldots, m$. Значит $h(\mathbf{x}_0) \leq 0$, тогда $h(\mathbf{x}_0) = 0$
- Итак, $h(\mathbf{x})$ вогнутая функция (поскольку аффинная), $h(\mathbf{x}) \geq 0$ и $h(\mathbf{x}_0) = 0$, то есть вогнутая функция достигает своего минимума в точке $\mathbf{x}_0$ и он равен 0
- Это означает, что $h(\mathbf{x}) = 0$ для всех $\mathbf{x} \in \mathcal{X}$ (докажите!)
- Но для каждого $\mathbf{y} \in \mathcal{M}$ найдётся $\mathbf{x} \in \mathcal{X}$ для которого $y_{p+j} = g_j(\mathbf{x})$.

- Но по условию теоремы существует точка $\mathbf{x}_0$ из $\operatorname{relint}(\mathcal{X})$ такая что $g_j(\mathbf{x}_0) \le 0$ для $j = 1, \ldots, m$. Значит $h(\mathbf{x}_0) \le 0$, тогда $h(\mathbf{x}_0) = 0$
- Итак, $h(\mathbf{x})$ вогнутая функция (поскольку аффинная), $h(\mathbf{x}) \ge 0$ и $h(\mathbf{x}_0) = 0$, то есть вогнутая функция достигает своего минимума в точке $\mathbf{x}_0$ и он равен 0
- Это означает, что $h(\mathbf{x}) = 0$ для всех $\mathbf{x} \in \mathcal{X}$ (докажите!)
- Но для каждого $\mathbf{y} \in \mathcal{M}$ найдётся $\mathbf{x} \in \mathcal{X}$ для которого $y_{p+j} = g_j(\mathbf{x})$.
- Это означает, что $\langle \mathbf{q}, \mathbf{y} \rangle = \sum_{j=1}^{m} q_{p+j} g_j(\mathbf{x}) = h(\mathbf{x}) = 0$

- Но по условию теоремы существует точка $\mathbf{x}_0$ из $\mathrm{relint}\,(\mathcal{X})$ такая что $g_j(\mathbf{x}_0) \leq 0$ для $j = 1, \dots, m$. Значит $h(\mathbf{x}_0) \leq 0$, тогда $h(\mathbf{x}_0) = 0$
- Итак, $h(\mathbf{x})$ вогнутая функция (поскольку аффинная), $h(\mathbf{x}) \geq 0$ и $h(\mathbf{x}_0) = 0$, то есть вогнутая функция достигает своего минимума в точке $\mathbf{x}_0$ и он равен 0
- Это означает, что $h(\mathbf{x}) = 0$ для всех $\mathbf{x} \in \mathcal{X}$ (докажите!)
- Но для каждого $\mathbf{y} \in \mathcal{M}$ найдётся $\mathbf{x} \in \mathcal{X}$ для которого $y_{p+j} = g_j(\mathbf{x})$.
- Это означает, что $\langle \mathbf{q}, \mathbf{y} \rangle = \sum_{j=1}^{m} q_{p+j} g_j(\mathbf{x}) = h(\mathbf{x}) = 0$
- То есть $\alpha = 0$ и $\mathcal{M}$ является подмножеством разделяющей гиперплоскости, получили противоречие.

- Но по условию теоремы существует точка $\mathbf{x}_0$ из $\mathrm{relint}\,(\mathcal{X})$ такая что $g_j(\mathbf{x}_0) \leq 0$ для $j = 1, \ldots, m$. Значит $h(\mathbf{x}_0) \leq 0$, тогда $h(\mathbf{x}_0) = 0$
- Итак, $h(\mathbf{x})$ вогнутая функция (поскольку аффинная), $h(\mathbf{x}) \geq 0$ и $h(\mathbf{x}_0) = 0$, то есть вогнутая функция достигает своего минимума в точке $\mathbf{x}_0$ и он равен 0
- Это означает, что $h(\mathbf{x}) = 0$ для всех $\mathbf{x} \in \mathcal{X}$ (докажите!)
- Но для каждого $\mathbf{y} \in \mathcal{M}$ найдётся $\mathbf{x} \in \mathcal{X}$ для которого $y_{p+j} = g_j(\mathbf{x})$.
- Это означает, что $\langle \mathbf{q}, \mathbf{y} \rangle = \sum_{j=1}^{m} q_{p+j} g_j(\mathbf{x}) = h(\mathbf{x}) = 0$
- То есть $\alpha = 0$ и $\mathcal{M}$ является подмножеством разделяющей гиперплоскости, получили противоречие.
- Таким образом, хотя бы один из коэффициентов $q_1, \ldots, q_p$ не равен нулю.

Возвращаемся к основному факту

$$
\begin{aligned}
&\min_{\mathbf{x} \in \mathcal{D}} f_0(\mathbf{x}) \\
\text{s.t. } & f_i(\mathbf{x}) \leq 0, \; i = 1, \ldots, p \\
& \mathbf{A}\mathbf{x} = \mathbf{b}
\end{aligned}
$$

Теорема

Если задача оптимизации выпукла, выполнено условие Слейтера и p^* конечно, тогда выполнена сильная двойственность, то есть найдётся пара допустимых $(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ таких что $g(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}) = p^*$

Возвращаемся к основному факту

$$
\begin{aligned}
&\min_{\mathbf{x} \in \mathcal{D}} f_0(\mathbf{x}) \\
\text{s.t. } & f_i(\mathbf{x}) \leq 0, \; i = 1, \ldots, p \\
& \mathbf{A}\mathbf{x} = \mathbf{b}
\end{aligned}
$$

Теорема

Если задача оптимизации выпукла, выполнено условие Слейтера и p^* конечно, тогда выполнена сильная двойственность, то есть найдётся пара допустимых $(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ таких что $g(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}) = p^*$

Доказательство

Возвращаемся к основному факту

$$
\begin{aligned}
&\min_{\mathbf{x} \in \mathcal{D}} f_0(\mathbf{x}) \\
\text{s.t. } & f_i(\mathbf{x}) \leq 0, \; i = 1, \ldots, p \\
& \mathbf{A}\mathbf{x} = \mathbf{b}
\end{aligned}
$$

Теорема

Если задача оптимизации выпукла, выполнено условие Слейтера и p^* конечно, тогда выполнена сильная двойственность, то есть найдётся пара допустимых $(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ таких что $g(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}) = p^*$

Доказательство

- Пусть в задаче выпуклой оптимизации есть только ограничения типа неравенств

Возвращаемся к основному факту

$$
\begin{aligned}
&\min_{\mathbf{x} \in \mathcal{D}} f_0(\mathbf{x}) \\
\text{s.t. } & f_i(\mathbf{x}) \leq 0, \; i = 1, \dots, p \\
& \mathbf{A}\mathbf{x} = \mathbf{b}
\end{aligned}
$$

Теорема

Если задача оптимизации выпукла, выполнено условие Слейтера и p^* конечно, тогда выполнена сильная двойственность, то есть найдётся пара допустимых $(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ таких что $g(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}) = p^*$

Доказательство

- Пусть в задаче выпуклой оптимизации есть только ограничения типа неравенств
- Также пусть первые k ограничений выпуклые и неаффинные, а оставшиеся $p - k + 1$ аффинные

- В силу условия Слейтера у системы
$$\begin{cases} f_i(\mathbf{x}) < 0, \; i = 1, \dots, k \\ f_i(\mathbf{x}) \le 0, \; i = k+1, \dots, p \end{cases}$$ есть решение для $\mathbf{x} \in \mathcal{D}$

- В силу условия Слейтера у системы
  $\begin{cases} f_i(\mathbf{x}) < 0, \ i = 1, \dots, k \\ f_i(\mathbf{x}) \le 0, \ i = k+1, \dots, p \end{cases}$ есть решение для $\mathbf{x} \in \mathcal{D}$
- Вместе с тем у системы $\begin{cases} f_0(\mathbf{x}) - p^* < 0 \\ f_i(\mathbf{x}) < 0, \ i = 1, \dots, k \\ f_i(\mathbf{x}) \le 0, \ i = k+1, \dots, p \end{cases}$
  решения нет в силу определения p^*

- В силу условия Слейтера у системы $\begin{cases} f_i(\mathbf{x}) < 0, \ i = 1, \dots, k \\ f_i(\mathbf{x}) \le 0, \ i = k+1, \dots, p \end{cases}$ есть решение для $\mathbf{x} \in \mathcal{D}$
- Вместе с тем у системы $\begin{cases} f_0(\mathbf{x}) - p^* < 0 \\ f_i(\mathbf{x}) < 0, \ i = 1, \dots, k \\ f_i(\mathbf{x}) \le 0, \ i = k+1, \dots, p \end{cases}$ решения нет в силу определения p^*
- Поэтому в силу ранее доказанного факта найдутся неотрицательные коэффициенты $\mu_0, \mu_1, \dots, \mu_p$ такие что хотя бы одно из чисел $\mu_0, \mu_1, \dots, \mu_k$ больше нуля и для всех $\mathbf{x} \in \mathcal{D}$ выполнено

$$\mu_0(f_0(\mathbf{x}) - p^*) + \sum_{i=1}^{p} \mu_i f_i(\mathbf{x}) \ge 0$$

- В силу условия Слейтера у системы $\begin{cases} f_i(\mathbf{x}) < 0, \; i = 1, \ldots, k \\ f_i(\mathbf{x}) \leq 0, \; i = k+1, \ldots, p \end{cases}$ есть решение для $\mathbf{x} \in \mathcal{D}$
- Вместе с тем у системы $\begin{cases} f_0(\mathbf{x}) - p^* < 0 \\ f_i(\mathbf{x}) < 0, \; i = 1, \ldots, k \\ f_i(\mathbf{x}) \leq 0, \; i = k+1, \ldots, p \end{cases}$ решения нет в силу определения p^*
- Поэтому в силу ранее доказанного факта найдутся неотрицательные коэффициенты $\mu_0, \mu_1, \ldots, \mu_p$ такие что хотя бы одно из чисел $\mu_0, \mu_1, \ldots, \mu_k$ больше нуля и для всех $\mathbf{x} \in \mathcal{D}$ выполнено
$$\mu_0(f_0(\mathbf{x}) - p^*) + \sum_{i=1}^{p} \mu_i f_i(\mathbf{x}) \geq 0$$
- $\mu_0 > 0$ так как иначе мы получили бы противоречие с условием Слейтера

- В силу условия Слейтера у системы $\begin{cases} f_i(\mathbf{x}) < 0, \ i = 1, \ldots, k \\ f_i(\mathbf{x}) \le 0, \ i = k+1, \ldots, p \end{cases}$ есть решение для $\mathbf{x} \in \mathcal{D}$
- Вместе с тем у системы $\begin{cases} f_0(\mathbf{x}) - p^* < 0 \\ f_i(\mathbf{x}) < 0, \ i = 1, \ldots, k \\ f_i(\mathbf{x}) \le 0, \ i = k+1, \ldots, p \end{cases}$ решения нет в силу определения p^*
- Поэтому в силу ранее доказанного факта найдутся неотрицательные коэффициенты $\mu_0, \mu_1, \ldots, \mu_p$ такие что хотя бы одно из чисел $\mu_0, \mu_1, \ldots, \mu_k$ больше нуля и для всех $\mathbf{x} \in \mathcal{D}$ выполнено
$$\mu_0(f_0(\mathbf{x}) - p^*) + \sum_{i=1}^{p} \mu_i f_i(\mathbf{x}) \ge 0$$
- $\mu_0 > 0$ так как иначе мы получили бы противоречие с условием Слейтера
- Поделив на μ_0, получим, что для всех $\mathbf{x} \in \mathcal{D}$ выполнено
$$L(\mathbf{x}, \boldsymbol{\mu}) \ge p^*$$

- Значит $g(\boldsymbol{\mu}) = \inf\limits_{\mathbf{x} \in \mathcal{D}} L(\mathbf{x}, \boldsymbol{\mu}) \geq p^*$

- Значит $g(\boldsymbol{\mu}) = \inf\limits_{\mathbf{x} \in \mathcal{D}} L(\mathbf{x}, \boldsymbol{\mu}) \geq p^*$
- Но мы знаем, что $g(\boldsymbol{\mu}) \leq p^*$ из прошлой лекции, а значит $p^* = g(\boldsymbol{\mu})$ и выполнена сильная двойственность

- Значит $g(\boldsymbol{\mu}) = \inf\limits_{\mathbf{x} \in \mathcal{D}} L(\mathbf{x}, \boldsymbol{\mu}) \geq p^*$
- Но мы знаем, что $g(\boldsymbol{\mu}) \leq p^*$ из прошлой лекции, а значит $p^* = g(\boldsymbol{\mu})$ и выполнена сильная двойственность
- Пусть теперь есть m аффинных ограничений типа равенств. Обозначим их $g_i(\mathbf{x}) = 0$

- Значит $g(\boldsymbol{\mu}) = \inf\limits_{\mathbf{x} \in \mathcal{D}} L(\mathbf{x}, \boldsymbol{\mu}) \geq p^*$
- Но мы знаем, что $g(\boldsymbol{\mu}) \leq p^*$ из прошлой лекции, а значит $p^* = g(\boldsymbol{\mu})$ и выполнена сильная двойственность
- Пусть теперь есть m аффинных ограничений типа равенств. Обозначим их $g_i(\mathbf{x}) = 0$
- Перепишем их в виде $2m$ аффинных ограничений типа неравенств вида $\mathbf{a}_i^\top \mathbf{x} + b_i \leq 0$ и $-\mathbf{a}_i^\top \mathbf{x} - b_i \leq 0$

- Значит $g(\boldsymbol{\mu}) = \inf\limits_{\mathbf{x} \in \mathcal{D}} L(\mathbf{x}, \boldsymbol{\mu}) \geq p^*$
- Но мы знаем, что $g(\boldsymbol{\mu}) \leq p^*$ из прошлой лекции, а значит $p^* = g(\boldsymbol{\mu})$ и выполнена сильная двойственность
- Пусть теперь есть m аффинных ограничений типа равенств. Обозначим их $g_i(\mathbf{x}) = 0$
- Перепишем их в виде $2m$ аффинных ограничений типа неравенств вида $\mathbf{a}_i^\top \mathbf{x} + b_i \leq 0$ и $-\mathbf{a}_i^\top \mathbf{x} - b_i \leq 0$
- Тогда в силу доказанного выше найдутся такие неотрицательные множители Лагранжа $\mu_1, \ldots, \mu_p$, $\beta_1, \ldots, \beta_m$ и $\gamma_1, \ldots, \gamma_m$, что для всех $\mathbf{x} \in \mathcal{D}$

$$f(\mathbf{x}) + \sum_{i=1}^{p} \mu_i f_i(\mathbf{x}) + \sum_{j=1}^{m} (\beta_j - \gamma_j) g_i(\mathbf{x}) \geq p^*$$

- Значит $g(\boldsymbol{\mu}) = \inf\limits_{\mathbf{x} \in \mathcal{D}} L(\mathbf{x}, \boldsymbol{\mu}) \geq p^*$
- Но мы знаем, что $g(\boldsymbol{\mu}) \leq p^*$ из прошлой лекции, а значит $p^* = g(\boldsymbol{\mu})$ и выполнена сильная двойственность
- Пусть теперь есть m аффинных ограничений типа равенств. Обозначим их $g_i(\mathbf{x}) = 0$
- Перепишем их в виде $2m$ аффинных ограничений типа неравенств вида $\mathbf{a}_i^\top \mathbf{x} + b_i \leq 0$ и $-\mathbf{a}_i^\top \mathbf{x} - b_i \leq 0$
- Тогда в силу доказанного выше найдутся такие неотрицательные множители Лагранжа $\mu_1, \ldots, \mu_p$, $\beta_1, \ldots, \beta_m$ и $\gamma_1, \ldots, \gamma_m$, что для всех $\mathbf{x} \in \mathcal{D}$

$$f(\mathbf{x}) + \sum_{i=1}^{p} \mu_i f_i(\mathbf{x}) + \sum_{j=1}^{m} (\beta_j - \gamma_j) g_i(\mathbf{x}) \geq p^*$$

- Обозначим $\lambda_j = \beta_j - \gamma_j$ и поскольку $\gamma_j \geq 0, \beta_j \geq 0$, то ограничений на знак λ_j нет

- Значит $g(\boldsymbol{\mu}) = \inf\limits_{\mathbf{x} \in \mathcal{D}} L(\mathbf{x}, \boldsymbol{\mu}) \geq p^*$
- Но мы знаем, что $g(\boldsymbol{\mu}) \leq p^*$ из прошлой лекции, а значит $p^* = g(\boldsymbol{\mu})$ и выполнена сильная двойственность
- Пусть теперь есть m аффинных ограничений типа равенств. Обозначим их $g_i(\mathbf{x}) = 0$
- Перепишем их в виде $2m$ аффинных ограничений типа неравенств вида $\mathbf{a}_i^\top \mathbf{x} + b_i \leq 0$ и $-\mathbf{a}_i^\top \mathbf{x} - b_i \leq 0$
- Тогда в силу доказанного выше найдутся такие неотрицательные множители Лагранжа $\mu_1, \ldots, \mu_p$, $\beta_1, \ldots, \beta_m$ и $\gamma_1, \ldots, \gamma_m$, что для всех $\mathbf{x} \in \mathcal{D}$

$$f(\mathbf{x}) + \sum_{i=1}^{p} \mu_i f_i(\mathbf{x}) + \sum_{j=1}^{m} (\beta_j - \gamma_j) g_i(\mathbf{x}) \geq p^*$$

- Обозначим $\lambda_j = \beta_j - \gamma_j$ и поскольку $\gamma_j \geq 0, \beta_j \geq 0$, то ограничений на знак λ_j нет
- Значит нашли допустимую точку $(\boldsymbol{\lambda}, \boldsymbol{\mu})$, такую что выполнена сильная двойственность

Теорема Каруша-Куна-Таккера

Формулировка

Пусть дана выпуклая задача и функции $f_0, \ldots, f_p$ дифференцируемы в допустимой точке $\hat{\mathbf{x}}$. Тогда

- Если $(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ допустимы и для тройки $(\hat{\mathbf{x}}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ выполнены условия ККТ, то выполнена сильная двойственность, $\hat{\mathbf{x}}$ — решение прямой задачи, $(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ — решение двойственной задачи
- Если выполнено условие Слейтера и $\hat{\mathbf{x}}$ решение задачи, тогда существуют множители Лагранжа $(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$, такие что для тройки $(\hat{\mathbf{x}}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ выполнены условия ККТ

Теорема Каруша-Куна-Таккера

Формулировка

Пусть дана выпуклая задача и функции $f_0, \ldots, f_p$ дифференцируемы в допустимой точке $\hat{\mathbf{x}}$. Тогда

- Если $(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ допустимы и для тройки $(\hat{\mathbf{x}}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ выполнены условия ККТ, то выполнена сильная двойственность, $\hat{\mathbf{x}}$ — решение прямой задачи, $(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ — решение двойственной задачи
- Если выполнено условие Слейтера и $\hat{\mathbf{x}}$ решение задачи, тогда существуют множители Лагранжа $(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$, такие что для тройки $(\hat{\mathbf{x}}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ выполнены условия ККТ

Доказательство

Теорема Каруша-Куна-Таккера

Формулировка

Пусть дана выпуклая задача и функции $f_0, \ldots, f_p$ дифференцируемы в допустимой точке $\hat{\mathbf{x}}$. Тогда

- Если $(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ допустимы и для тройки $(\hat{\mathbf{x}}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ выполнены условия ККТ, то выполнена сильная двойственность, $\hat{\mathbf{x}}$ — решение прямой задачи, $(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ — решение двойственной задачи
- Если выполнено условие Слейтера и $\hat{\mathbf{x}}$ решение задачи, тогда существуют множители Лагранжа $(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$, такие что для тройки $(\hat{\mathbf{x}}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ выполнены условия ККТ

Доказательство

- Если задача выпукла и точка $(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ допустима, то лагранжиан есть выпуклая функция по $\mathbf{x}$ (проверьте это!)

Теорема Каруша-Куна-Таккера

Формулировка

Пусть дана выпуклая задача и функции $f_0, \ldots, f_p$ дифференцируемы в допустимой точке $\hat{\mathbf{x}}$. Тогда

- Если $(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ допустимы и для тройки $(\hat{\mathbf{x}}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ выполнены условия ККТ, то выполнена сильная двойственность, $\hat{\mathbf{x}}$ — решение прямой задачи, $(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ — решение двойственной задачи
- Если выполнено условие Слейтера и $\hat{\mathbf{x}}$ решение задачи, тогда существуют множители Лагранжа $(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$, такие что для тройки $(\hat{\mathbf{x}}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ выполнены условия ККТ

Доказательство

- Если задача выпукла и точка $(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ допустима, то лагранжиан есть выпуклая функция по $\mathbf{x}$ (проверьте это!)
- Так как градиент лагранжиана (выпуклой функции) по $\mathbf{x}$ равен 0, то $\hat{\mathbf{x}}$ есть точка минимума лагранжиана

Доказательство

- Выполнено условие 3) из теоремы о седловой точке функции Лагранжа и выполнена сильная двойственность

Доказательство

- Выполнено условие 3) из теоремы о седловой точке функции Лагранжа и выполнена сильная двойственность
- Обратно, если выполнено условие Слейтера и $\hat{\mathbf{x}}$ решение, то выполнена сильная двойственность.

Доказательство

- Выполнено условие 3) из теоремы о седловой точке функции Лагранжа и выполнена сильная двойственность
- Обратно, если выполнено условие Слейтера и $\hat{\mathbf{x}}$ решение, то выполнена сильная двойственность.
- Значит найдётся допустимая пара $(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ такая что

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

Доказательство

- Выполнено условие 3) из теоремы о седловой точке функции Лагранжа и выполнена сильная двойственность
- Обратно, если выполнено условие Слейтера и $\hat{\mathbf{x}}$ решение, то выполнена сильная двойственность.
- Значит найдётся допустимая пара $(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}})$ такая что

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

- Отсюда следует условие дополняющей нежёсткости и стационарность лагранжиана

Геометрическая интерпретация условий ККТ

- Условие стационарности лагранжиана в точке минимума

$$f_0'(\mathbf{x}^*) + \sum_{i=1}^{p} \mu_i^* f_i'(\mathbf{x}^*) + \mathbf{A}^\top \boldsymbol{\lambda}^* = 0$$

- Если рассмотреть задачу только с ограничениями типа неравенств, то $-f_0'(\mathbf{x}^*)$ должен лежать в конусе, натянутом на градиенты активных ограничений в $\mathbf{x}^*$

Условия ККТ для задачи с обобщёнными неравенствами

1. $h_j(\mathbf{x}^*) \preceq_{\mathcal{K}} 0$

Условия ККТ для задачи с обобщёнными неравенствами

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

Условия ККТ для задачи с обобщёнными неравенствами

1. $h_j(\mathbf{x}^*) \preceq_{\mathcal{K}} 0$
2. $g_i(\mathbf{x}^*) = 0$
3. $\boldsymbol{\mu}^* \succeq_{\mathcal{K}^*} 0$

Условия ККТ для задачи с обобщёнными неравенствами

1. $h_j(\mathbf{x}^*) \preceq_{\mathcal{K}} 0$
2. $g_i(\mathbf{x}^*) = 0$
3. $\boldsymbol{\mu}^* \succeq_{\mathcal{K}^*} 0$
4. $\mu_j^* h_j(\mathbf{x}^*) = 0$

Условия ККТ для задачи с обобщёнными неравенствами

1. $h_j(\mathbf{x}^*) \preceq_{\mathcal{K}} 0$
2. $g_i(\mathbf{x}^*) = 0$
3. $\boldsymbol{\mu}^* \succeq_{\mathcal{K}^*} 0$
4. $\mu_j^* h_j(\mathbf{x}^*) = 0$
5. $f_0'(\mathbf{x}^*) + \sum\limits_{i=1}^m \lambda_i^* g_i'(\mathbf{x}^*) + \sum\limits_{j=1}^p \mu_j^* h_j'(\mathbf{x}^*) = 0$

Условия ККТ для задачи с обобщёнными неравенствами

1. $h_j(\mathbf{x}^*) \preceq_{\mathcal{K}} 0$
2. $g_i(\mathbf{x}^*) = 0$
3. $\boldsymbol{\mu}^* \succeq_{\mathcal{K}^*} 0$
4. $\mu_j^* h_j(\mathbf{x}^*) = 0$
5. $f_0'(\mathbf{x}^*) + \sum\limits_{i=1}^{m} \lambda_i^* g_i'(\mathbf{x}^*) + \sum\limits_{j=1}^{p} \mu_j^* h_j'(\mathbf{x}^*) = 0$

Условия оптимальности

Все утверждения для обычных скалярных неравенств (то есть для конуса $\mathbb{R}^n_+$) переносятся на случай произвольного конуса $\mathcal{K}$ с точностью до отмеченных отличий.

Условия ККТ для задачи с обобщёнными неравенствами

1. $h_j(\mathbf{x}^*) \preceq_{\mathcal{K}} 0$
2. $g_i(\mathbf{x}^*) = 0$
3. $\boldsymbol{\mu}^* \succeq_{\mathcal{K}^*} 0$
4. $\mu_j^* h_j(\mathbf{x}^*) = 0$
5. $f_0'(\mathbf{x}^*) + \sum\limits_{i=1}^{m} \lambda_i^* g_i'(\mathbf{x}^*) + \sum\limits_{j=1}^{p} \mu_j^* h_j'(\mathbf{x}^*) = 0$

Условия оптимальности

Все утверждения для обычных скалярных неравенств (то есть для конуса $\mathbb{R}^n_+$) переносятся на случай произвольного конуса $\mathcal{K}$ с точностью до отмеченных отличий.

Условие Слейтера для выпуклой задачи

Говорят, что выполнено условие Слейтера, если найдётся $\hat{\mathbf{x}} \in \mathcal{D}$ такой что $\mathbf{A}\hat{\mathbf{x}} = \mathbf{b}$ и $f_i(\hat{\mathbf{x}}) <_{\mathcal{K}} 0$

Главное

- Условие Слейтера

Главное

- Условие Слейтера
- Выпуклость + условие Слейтера = сильная двойственность

Главное

- Условие Слейтера
- Выпуклость + условие Слейтера = сильная двойственность
- Теорема Каруша-Куна-Таккера

Главное

- Условие Слейтера
- Выпуклость + условие Слейтера = сильная двойственность
- Теорема Каруша-Куна-Таккера
- Геометрическая интерпретация