Гладкая оптимизация с ограничениями: необходимые условия оптимальности

Н.Н. Писарук
pisaruk@yandex.by

Экономический факультет
Белорусский государственный университет

Минск - 2015

План лекции

Постановка задачи

Будем рассматривать задачу

$$f(x) \to \min,$$
$$g_i(x) \le 0, \quad i \in I = \{1, \dots, m\},$$
$$x \in \mathbb{R}^n,$$

где функции f и g_i $(i \in I)$ *непрерывно дифференцируемы.* Обозначим через X множество *допустимых решений* этой задачи, т. е.

$$X = \{x \in \mathbb{R}^n : \ g_i(x) \le 0, \ i \in I\}.$$

Будем предполагать, что множество X непустое; однако допускаем, что оно может иметь пустую внутренность.

Постановка задачи

Будем рассматривать задачу

$$
\begin{aligned}
f(x) &\to \min, \\
g_i(x) &\le 0, \quad i \in I = \{1, \ldots, m\}, \\
x &\in \mathbb{R}^n,
\end{aligned}
$$

где функции f и g_i $(i \in I)$ *непрерывно дифференцируемы.*

Обозначим через X множество *допустимых решений* этой задачи, т. е.

$$X = \{x \in \mathbb{R}^n : \ g_i(x) \le 0, \ i \in I\}.$$

Будем предполагать, что множество X непустое; однако допускаем, что оно может иметь пустую внутренность.

Постановка задачи

Будем рассматривать задачу

$$\begin{aligned} f(x) &\to \min, \\ g_i(x) &\le 0, \quad i \in I = \{1, \ldots, m\}, \\ x &\in \mathbb{R}^n, \end{aligned}$$

где функции f и g_i $(i \in I)$ *непрерывно дифференцируемы.*

Обозначим через X множество *допустимых решений* этой задачи, т. е.

$$X = \{x \in \mathbb{R}^n : \ g_i(x) \le 0, \ i \in I\}.$$

Будем предполагать, что множество X непустое; однако допускаем, что оно может иметь пустую внутренность.

Постановка задачи

Будем рассматривать задачу

$$\begin{aligned} f(x) &\to \min, \\ g_i(x) &\le 0, \quad i \in I = \{1, \ldots, m\}, \\ x &\in \mathbb{R}^n, \end{aligned}$$

где функции f и g_i $(i \in I)$ *непрерывно дифференцируемы*. Обозначим через X множество *допустимых решений* этой задачи, т. е.

$$X = \{x \in \mathbb{R}^n : \; g_i(x) \le 0, \; i \in I\}.$$

Будем предполагать, что множество X непустое; однако допускаем, что оно может иметь пустую внутренность.

Локальные оптимумы

Определение

Точка $x^0 \in X$ есть *локальный минимум* задачи

$$\min\{f(x) : \ g_i(x) \le 0, \ i \in I; \ x \in \mathbb{R}^n\},$$

если для некоторого числа $\epsilon > 0$ выполняется условие

$$f(x^0) \le f(x) \quad \forall\, x \in X, \ \|x - x^0\| \le \epsilon.$$

Локальные оптимумы

Определение

Точка $x^0 \in X$ есть *локальный минимум* задачи

$$\min\{f(x) : \; g_i(x) \leq 0, \; i \in I; \; x \in \mathbb{R}^n\},$$

если для некоторого числа $\epsilon > 0$ выполняется условие

$$f(x^0) \leq f(x) \quad \forall\, x \in X, \; \|x - x^0\| \leq \epsilon.$$

План лекции

1. Задача гладкой оптимизации
2. Условие выделения ограничений
   - Допустимые направления
   - Касательный конус
   - Достаточные условия для выделения ограничений
3. Необходимые условия оптимальности Куна — Таккера
   - Теорема Куна — Таккера
   - Геометрическая интерпретация условий Куна — Таккера

Допустимые направления

- Если $x^0 \in X$ есть локальный оптимум задачи

$$\min\{f(x):\ g_i(x) \le 0,\ i \in I;\ x \in \mathbb{R}^n\},$$

- то функция $f(x)$ не может убывать, когда x описывает дугу кривой (достаточно регулярной), выходящую из x^0 и содержащуюся в множестве решений X.
- Такую *допустимую* дугу кривой можно описать посредством непрерывно дифференцируемой функции $\varphi:\ \mathbb{R}_+ \to \mathbb{R}^n$ параметра $\theta \ge 0$:

$$\varphi(\theta) = [\varphi_1(\theta), \ldots, \varphi_n(\theta)]^T,$$

которая удовлетворяет условиям:

1. $\varphi(0) = x^0$;
2. $\varphi(\theta) \in X$ для достаточно малых $\theta > 0$.

Допустимые направления

- Если $x^0 \in X$ есть локальный оптимум задачи
$$\min\{f(x):\ g_i(x) \le 0,\ i \in I;\ x \in \mathbb{R}^n\},$$
- то функция $f(x)$ не может убывать, когда x описывает дугу кривой (достаточно регулярной), выходящую из x^0 и содержащуюся в множестве решений X.
- Такую *допустимую* дугу кривой можно описать посредством непрерывно дифференцируемой функции $\varphi:\ \mathbb{R}_+ \to \mathbb{R}^n$ параметра $\theta \ge 0$:
$$\varphi(\theta) = [\varphi_1(\theta), \dots, \varphi_n(\theta)]^T,$$
которая удовлетворяет условиям:
  1. $\varphi(0) = x^0$;
  2. $\varphi(\theta) \in X$ для достаточно малых $\theta > 0$.

Допустимые направления

- Если $x^0 \in X$ есть локальный оптимум задачи
$$\min\{f(x) : \ g_i(x) \le 0, \ i \in I; \ x \in \mathbb{R}^n\},$$
- то функция $f(x)$ не может убывать, когда x описывает дугу кривой (достаточно регулярной), выходящую из x^0 и содержащуюся в множестве решений X.
- Такую *допустимую* дугу кривой можно описать посредством непрерывно дифференцируемой функции $\varphi : \ \mathbb{R}_+ \to \mathbb{R}^n$ параметра $\theta \ge 0$:
$$\varphi(\theta) = [\varphi_1(\theta), \ldots, \varphi_n(\theta)]^T,$$
которая удовлетворяет условиям:
    1. $\varphi(0) = x^0$;
    2. $\varphi(\theta) \in X$ для достаточно малых $\theta > 0$.

Допустимые направления

- Если $x^0 \in X$ есть локальный оптимум задачи
$$\min\{f(x) :\ g_i(x) \le 0,\ i \in I;\ x \in \mathbb{R}^n\},$$
- то функция $f(x)$ не может убывать, когда x описывает дугу кривой (достаточно регулярной), выходящую из x^0 и содержащуюся в множестве решений X.
- Такую *допустимую* дугу кривой можно описать посредством непрерывно дифференцируемой функции $\varphi :\ \mathbb{R}_+ \to \mathbb{R}^n$ параметра $\theta \ge 0$:
$$\varphi(\theta) = [\varphi_1(\theta), \dots, \varphi_n(\theta)]^T,$$
которая удовлетворяет условиям:
    1. $\varphi(0) = x^0$;
    2. $\varphi(\theta) \in X$ для достаточно малых $\theta > 0$.

Допустимые направления

- Если $x^0 \in X$ есть локальный оптимум задачи
$$\min\{f(x) : \; g_i(x) \le 0, \; i \in I; \; x \in \mathbb{R}^n\},$$
- то функция $f(x)$ не может убывать, когда x описывает дугу кривой (достаточно регулярной), выходящую из x^0 и содержащуюся в множестве решений X.
- Такую *допустимую* дугу кривой можно описать посредством непрерывно дифференцируемой функции $\varphi : \; \mathbb{R}_+ \to \mathbb{R}^n$ параметра $\theta \ge 0$:
$$\varphi(\theta) = [\varphi_1(\theta), \dots, \varphi_n(\theta)]^T,$$
которая удовлетворяет условиям:
  1. $\varphi(0) = x^0$;
  2. $\varphi(\theta) \in X$ для достаточно малых $\theta > 0$.

Допустимые направления

Определение

Допустимым направлением в точке x^0 назовем вектор

$$y = \frac{d\varphi}{d\theta}(0) = \left[\frac{d\varphi_1}{d\theta}(0), \frac{d\varphi_2}{d\theta}(0), \ldots, \frac{d\varphi_n}{d\theta}(0)\right]^T,$$

касающийся некоторой дуги кривой $\varphi(\theta)$, допустимой в x^0.

- Обозначим через $C_{\mathrm{ad}}(x^0)$ конус, образованный множеством допустимых направлений в точке x^0.
- Как проверить включение $y \in C_{\mathrm{ad}}(x^0)$?

Допустимые направления

Определение

Допустимым направлением в точке x^0 назовем вектор

$$y = \frac{d\varphi}{d\theta}(0) = \left[\frac{d\varphi_1}{d\theta}(0), \frac{d\varphi_2}{d\theta}(0), \dots, \frac{d\varphi_n}{d\theta}(0)\right]^T,$$

касающийся некоторой дуги кривой $\varphi(\theta)$, допустимой в x^0.

- Обозначим через $C_{\mathrm{ad}}(x^0)$ конус, образованный множеством допустимых направлений в точке x^0.
- Как проверить включение $y \in C_{\mathrm{ad}}(x^0)$?

Допустимые направления

Определение

Допустимым направлением в точке x^0 назовем вектор

$$y = \frac{d\varphi}{d\theta}(0) = \left[\frac{d\varphi_1}{d\theta}(0), \frac{d\varphi_2}{d\theta}(0), \dots, \frac{d\varphi_n}{d\theta}(0)\right]^T,$$

касающийся некоторой дуги кривой $\varphi(\theta)$, допустимой в x^0.

- Обозначим через $C_{\text{ad}}(x^0)$ конус, образованный множеством допустимых направлений в точке x^0.
- Как проверить включение $y \in C_{\text{ad}}(x^0)$?

План лекции

Касательный конус

Определение

Касательный конус в точке x^0 определяется следующим образом:

$$T(x^0) \stackrel{\text{def}}{=} \{y \in \mathbb{R}^n : \ \nabla g_i^T(x^0) y \le 0 \quad \forall\, i \in I(x^0)\},$$

где $I(x^0) \stackrel{\text{def}}{=} \{i \in I : \ g_i(x^0) = 0\}$ есть множество индексов *насыщенных* ограничений в точке x^0.

Лемма

Справедливо включение $C_{\text{ad}}(x^0) \subseteq T(x^0)$.

Перейти к доказательству

Касательный конус

Определение

Касательный конус в точке x^0 определяется следующим образом:

$$T(x^0) \stackrel{\text{def}}{=} \{y \in \mathbb{R}^n : \ \nabla g_i^T(x^0) y \le 0 \quad \forall\, i \in I(x^0)\},$$

где $I(x^0) \stackrel{\text{def}}{=} \{i \in I : \ g_i(x^0) = 0\}$ есть множество индексов *насыщенных* ограничений в точке x^0.

Лемма

Справедливо включение $C_{\text{ad}}(x^0) \subseteq T(x^0)$.

Перейти к доказательству

Касательный конус

Определение

Касательный конус в точке x^0 определяется следующим образом:

$$T(x^0) \stackrel{\text{def}}{=} \{y \in \mathbb{R}^n : \ \nabla g_i^T(x^0) y \le 0 \quad \forall\, i \in I(x^0)\},$$

где $I(x^0) \stackrel{\text{def}}{=} \{i \in I : \ g_i(x^0) = 0\}$ есть множество индексов *насыщенных* ограничений в точке x^0.

Лемма

Справедливо включение $C_{\text{ad}}(x^0) \subseteq T(x^0)$.

Перейти к доказательству

Включение $T(x^0) \subseteq C_{\text{ad}}(x^0)$ в общем случае неверно

Рассмотрим в $\mathbb{R}^2$ множество X, определяемое ограничениями

$g_1(x) = -x_1 \le 0,$

$g_2(x) = -x_2 \le 0,$

$g_3(x) = -(1-x_1)^3 + x_2 \le 0.$

- $x^0 = (1,0)^T$, $I(x^0) = \{2,3\}$.
- Поскольку $\nabla g_2(x^0) = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$, $\quad \nabla g_3(x^0) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$,

  то $T(x^0) = \{y \in \mathbb{R}^2 : \ -y_2 \le 0,\ y_2 \le 0\} = \{y \in \mathbb{R}^2 : \ y_2 = 0\}$.
- Вектор $y = (1,0)^T \in T(x^0)$, однако это направление не является допустимым, поскольку $x^0 + \theta y \notin X \ \forall\, \theta > 0$.
- Заметим, что $\nabla g_2(x^0)$ и $\nabla g_3(x^0)$ линейно зависимы!

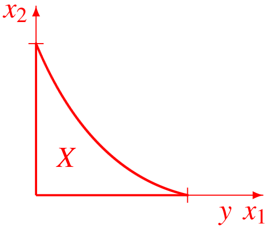
Включение $T(x^0) \subseteq C_{\text{ad}}(x^0)$ в общем случае неверно

Рассмотрим в $\mathbb{R}^2$ множество X, определяемое ограничениями

$g_1(x) = -x_1 \le 0,$

$g_2(x) = -x_2 \le 0,$

$g_3(x) = -(1-x_1)^3 + x_2 \le 0.$



- $x^0 = (1,0)^T$, $I(x^0) = \{2,3\}$.
- Поскольку $\nabla g_2(x^0) = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$, $\quad \nabla g_3(x^0) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$,

  то $T(x^0) = \{y \in \mathbb{R}^2 :\ -y_2 \le 0,\ y_2 \le 0\} = \{y \in \mathbb{R}^2 :\ y_2 = 0\}$.
- Вектор $y = (1,0)^T \in T(x^0)$, однако это направление не является допустимым, поскольку $x^0 + \theta y \notin X\ \forall\ \theta > 0$.
- Заметим, что $\nabla g_2(x^0)$ и $\nabla g_3(x^0)$ линейно зависимы!

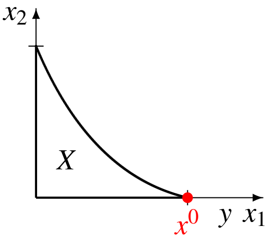
Включение $T(x^0) \subseteq C_{ad}(x^0)$ в общем случае неверно

Рассмотрим в $\mathbb{R}^2$ множество X, определяемое ограничениями

$g_1(x) = -x_1 \leq 0,$

$g_2(x) = -x_2 \leq 0,$

$g_3(x) = -(1 - x_1)^3 + x_2 \leq 0.$



- $x^0 = (1, 0)^T$, $I(x^0) = \{2, 3\}$.
- Поскольку $\nabla g_2(x^0) = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$, $\nabla g_3(x^0) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$,

  то $T(x^0) = \{y \in \mathbb{R}^2 : \ -y_2 \leq 0, \ y_2 \leq 0\} = \{y \in \mathbb{R}^2 : \ y_2 = 0\}$.
- Вектор $y = (1, 0)^T \in T(x^0)$, однако это направление не является допустимым, поскольку $x^0 + \theta y \notin X \ \forall \ \theta > 0$.
- Заметим, что $\nabla g_2(x^0)$ и $\nabla g_3(x^0)$ линейно зависимы!

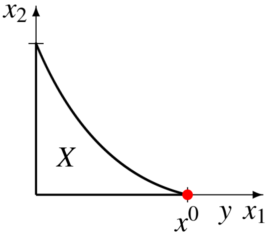
Включение $T(x^0) \subseteq C_{\text{ad}}(x^0)$ в общем случае неверно

Рассмотрим в $\mathbb{R}^2$ множество X, определяемое ограничениями

$g_1(x) = -x_1 \le 0,$

$g_2(x) = -x_2 \le 0,$

$g_3(x) = -(1-x_1)^3 + x_2 \le 0.$



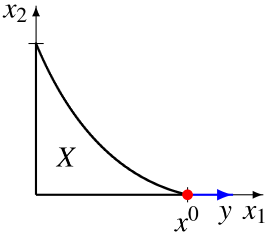
- $x^0 = (1,0)^T$, $I(x^0) = \{2,3\}$.
- Поскольку $\nabla g_2(x^0) = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$, $\quad \nabla g_3(x^0) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$,

  то $T(x^0) = \{y \in \mathbb{R}^2 : \ -y_2 \le 0, \ y_2 \le 0\} = \{y \in \mathbb{R}^2 : \ y_2 = 0\}$.
- Вектор $y = (1,0)^T \in T(x^0)$, однако это направление не является допустимым, поскольку $x^0 + \theta y \notin X \ \forall \ \theta > 0$.
- Заметим, что $\nabla g_2(x^0)$ и $\nabla g_3(x^0)$ линейно зависимы!

Включение $T(x^0) \subseteq C_{ad}(x^0)$ в общем случае неверно

Рассмотрим в $\mathbb{R}^2$ множество X, определяемое ограничениями

$$g_1(x) = -x_1 \le 0,$$
$$g_2(x) = -x_2 \le 0,$$
$$g_3(x) = -(1 - x_1)^3 + x_2 \le 0.$$



- $x^0 = (1, 0)^T$, $I(x^0) = \{2, 3\}$.
- Поскольку $\nabla g_2(x^0) = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$, $\quad \nabla g_3(x^0) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$,

  то $T(x^0) = \{y \in \mathbb{R}^2 : \ -y_2 \le 0, \ y_2 \le 0\} = \{y \in \mathbb{R}^2 : \ y_2 = 0\}$.
- Вектор $y = (1, 0)^T \in T(x^0)$, однако это направление не является допустимым, поскольку $x^0 + \theta y \notin X \ \forall \ \theta > 0$.
- Заметим, что $\nabla g_2(x^0)$ и $\nabla g_3(x^0)$ линейно зависимы!

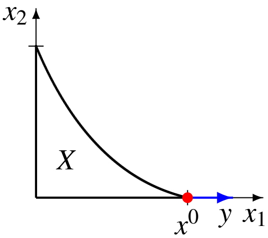
Включение $T(x^0) \subseteq C_{ad}(x^0)$ в общем случае неверно

Рассмотрим в $\mathbb{R}^2$ множество X, определяемое ограничениями

$g_1(x) = -x_1 \le 0,$

$g_2(x) = -x_2 \le 0,$

$g_3(x) = -(1-x_1)^3 + x_2 \le 0.$



- $x^0 = (1,0)^T$, $I(x^0) = \{2,3\}$.
- Поскольку $\nabla g_2(x^0) = \begin{bmatrix} 0 \\ -1 \end{bmatrix}, \quad \nabla g_3(x^0) = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$

  то $T(x^0) = \{y \in \mathbb{R}^2 : \ -y_2 \le 0, \ y_2 \le 0\} = \{y \in \mathbb{R}^2 : \ y_2 = 0\}$.
- Вектор $y = (1,0)^T \in T(x^0)$, однако это направление не является допустимым, поскольку $x^0 + \theta y \notin X \ \forall \ \theta > 0$.
- Заметим, что $\nabla g_2(x^0)$ и $\nabla g_3(x^0)$ линейно зависимы!

Условие выделения ограничений

Определение

Говорят, что в точке $x^0 \in X$ выполняется *условие выделения ограничений*, если касательный конус в этой точке является замыканием конуса допустимых направлений:

$$\mathrm{cl}(C_{\mathrm{ad}}(x^0)) = T(x^0).$$

При выполнение условия выделения ограничений

$$\mathrm{cl}(C_{\mathrm{ad}}(x^0)) = \{y \in \mathbb{R}^n : \nabla g_i^T(x^0) y \le 0 \quad \forall\, i \in I(x^0)\}.$$

Условие выделения ограничений

Определение

Говорят, что в точке $x^0 \in X$ выполняется *условие выделения ограничений*, если касательный конус в этой точке является замыканием конуса допустимых направлений:

$$\mathrm{cl}(C_{\mathrm{ad}}(x^0)) = T(x^0).$$

При выполнение условия выделения ограничений

$$\mathrm{cl}(C_{\mathrm{ad}}(x^0)) = \{y \in \mathbb{R}^n : \nabla g_i^T(x^0) y \le 0 \quad \forall\, i \in I(x^0)\}.$$

Условие выделения ограничений

Определение

Говорят, что в точке $x^0 \in X$ выполняется *условие выделения ограничений*, если касательный конус в этой точке является замыканием конуса допустимых направлений:

$$\mathrm{cl}(C_{\mathrm{ad}}(x^0)) = T(x^0).$$

При выполнение условия выделения ограничений

$$\mathrm{cl}(C_{\mathrm{ad}}(x^0)) = \{y \in \mathbb{R}^n : \nabla g_i^T(x^0) y \le 0 \quad \forall\, i \in I(x^0)\}.$$

План лекции

Достаточные условие для выделения ограничений

Лемма

1. *Условие выделения ограничений выполняется в каждой точке множества X в следующих случаях:*
   - *а) все функции g_i линейны;*
   - *б) (условие Слейтера) все функции g_i выпуклы и множество X имеет непустую внутренность (существует $\bar{x} \in \mathbb{R}^n$, что $g_i(\bar{x}) < 0$ для всех $i \in I$).*
2. *Условие выделения ограничений выполняется в точке $x^0 \in X$, если*
   - *в) градиенты $\nabla g_i(x^0)$ ограничений, которые в точке x^0 выполняются как равенства, линейно независимы.*

Перейти к доказательству

Достаточные условие для выделения ограничений

Лемма

1. *Условие выделения ограничений выполняется в каждой точке множества X в следующих случаях:*
   - *а) все функции g_i линейны;*
   - *б) (условие Слейтера) все функции g_i выпуклы и множество X имеет непустую внутренность (существует $\bar{x} \in \mathbb{R}^n$, что $g_i(\bar{x}) < 0$ для всех $i \in I$).*
2. *Условие выделения ограничений выполняется в точке $x^0 \in X$, если*
   - *в) градиенты $\nabla g_i(x^0)$ ограничений, которые в точке x^0 выполняются как равенства, линейно независимы.*

Перейти к доказательству

Достаточные условие для выделения ограничений

Лемма

1. *Условие выделения ограничений выполняется в каждой точке множества X в следующих случаях:*
   - *а) все функции g_i линейны;*
   - *б) (условие Слейтера) все функции g_i выпуклы и множество X имеет непустую внутренность (существует $\bar{x} \in \mathbb{R}^n$, что $g_i(\bar{x}) < 0$ для всех $i \in I$).*
2. *Условие выделения ограничений выполняется в точке $x^0 \in X$, если*
   - *в) градиенты $\nabla g_i(x^0)$ ограничений, которые в точке x^0 выполняются как равенства, линейно независимы.*

Перейти к доказательству

Достаточные условие для выделения ограничений

Лемма

1. *Условие выделения ограничений выполняется в каждой точке множества X в следующих случаях:*
   - *а) все функции g_i линейны;*
   - *б) (условие Слейтера) все функции g_i выпуклы и множество X имеет непустую внутренность (существует $\bar{x} \in \mathbb{R}^n$, что $g_i(\bar{x}) < 0$ для всех $i \in I$).*
2. *Условие выделения ограничений выполняется в точке $x^0 \in X$, если*
   - *в) градиенты $\nabla g_i(x^0)$ ограничений, которые в точке x^0 выполняются как равенства, линейно независимы.*

Перейти к доказательству

Достаточные условие для выделения ограничений

Лемма

1. *Условие выделения ограничений выполняется в каждой точке множества X в следующих случаях:*
   - *а) все функции g_i линейны;*
   - *б) (условие Слейтера) все функции g_i выпуклы и множество X имеет непустую внутренность (существует $\bar{x} \in \mathbb{R}^n$, что $g_i(\bar{x}) < 0$ для всех $i \in I$).*
2. *Условие выделения ограничений выполняется в точке $x^0 \in X$, если*
   - *в) градиенты $\nabla g_i(x^0)$ ограничений, которые в точке x^0 выполняются как равенства, линейно независимы.*

Перейти к доказательству

Комбинирование достат. условий для выделения ограничений

Линейные и выпуклые функции ограничений

1. Условие выделения ограничений выполняются в любой точке $x \in X$, если
   - часть функций g_i линейные $(i \in I_l)$,
   - а остальные функции выпуклые $(i \in I_c)$
   - и существует точка $\bar{x} \in X$, что $g_i(\bar{x}) < 0$ для всех $i \in I_c$.
2. Условие выделения ограничений также выполняются в точке $x^0 \in X$, если
   - часть функций g_i линейные $(i \in I_l)$,
   - а градиенты $\nabla g_i(x^0)$ линейно независимы для нелинейных ограничений $i \in I_n$
   - и существует точка $\bar{x} \in X$, что $g_i(\bar{x}) < 0$ для всех $i \in I_n$.

Комбинирование достат. условий для выделения ограничений

Линейные и выпуклые функции ограничений

1. Условие выделения ограничений выполняются в любой точке $x \in X$, если
   - часть функций g_i линейные $(i \in I_l)$,
   - а остальные функции выпуклые $(i \in I_c)$
   - и существует точка $\bar{x} \in X$, что $g_i(\bar{x}) < 0$ для всех $i \in I_c$.
2. Условие выделения ограничений также выполняются в точке $x^0 \in X$, если
   - часть функций g_i линейные $(i \in I_l)$,
   - а градиенты $\nabla g_i(x^0)$ линейно независимы для нелинейных ограничений $i \in I_n$
   - и существует точка $\bar{x} \in X$, что $g_i(\bar{x}) < 0$ для всех $i \in I_n$.

Комбинирование достат. условий для выделения ограничений

Линейные и выпуклые функции ограничений

1. Условие выделения ограничений выполняются в любой точке $x \in X$, если
   - часть функций g_i линейные $(i \in I_l)$,
   - а остальные функции выпуклые $(i \in I_c)$
   - и существует точка $\bar{x} \in X$, что $g_i(\bar{x}) < 0$ для всех $i \in I_c$.
2. Условие выделения ограничений также выполняются в точке $x^0 \in X$, если
   - часть функций g_i линейные $(i \in I_l)$,
   - а градиенты $\nabla g_i(x^0)$ линейно независимы для нелинейных ограничений $i \in I_n$
   - и существует точка $\bar{x} \in X$, что $g_i(\bar{x}) < 0$ для всех $i \in I_n$.

Комбинирование достат. условий для выделения ограничений

Линейные и выпуклые функции ограничений

1. Условие выделения ограничений выполняются в любой точке $x \in X$, если
   - часть функций g_i линейные $(i \in I_l)$,
   - а остальные функции выпуклые $(i \in I_c)$
   - и существует точка $\bar{x} \in X$, что $g_i(\bar{x}) < 0$ для всех $i \in I_c$.
2. Условие выделения ограничений также выполняются в точке $x^0 \in X$, если
   - часть функций g_i линейные $(i \in I_l)$,
   - а градиенты $\nabla g_i(x^0)$ линейно независимы для нелинейных ограничений $i \in I_n$
   - и существует точка $\bar{x} \in X$, что $g_i(\bar{x}) < 0$ для всех $i \in I_n$.

Комбинирование достат. условий для выделения ограничений

Линейные и выпуклые функции ограничений

1. Условие выделения ограничений выполняются в любой точке $x \in X$, если
   - часть функций g_i линейные $(i \in I_l)$,
   - а остальные функции выпуклые $(i \in I_c)$
   - и существует точка $\bar{x} \in X$, что $g_i(\bar{x}) < 0$ для всех $i \in I_c$.
2. Условие выделения ограничений также выполняются в точке $x^0 \in X$, если
   - часть функций g_i линейные $(i \in I_l)$,
   - а градиенты $\nabla g_i(x^0)$ линейно независимы для нелинейных ограничений $i \in I_n$
   - и существует точка $\bar{x} \in X$, что $g_i(\bar{x}) < 0$ для всех $i \in I_n$.

Комбинирование достат. условий для выделения ограничений

Линейные и выпуклые функции ограничений

1. Условие выделения ограничений выполняются в любой точке $x \in X$, если
   - часть функций g_i линейные $(i \in I_l)$,
   - а остальные функции выпуклые $(i \in I_c)$
   - и существует точка $\bar{x} \in X$, что $g_i(\bar{x}) < 0$ для всех $i \in I_c$.
2. Условие выделения ограничений также выполняются в точке $x^0 \in X$, если
   - часть функций g_i линейные $(i \in I_l)$,
   - а градиенты $\nabla g_i(x^0)$ линейно независимы для нелинейных ограничений $i \in I_n$
   - и существует точка $\bar{x} \in X$, что $g_i(\bar{x}) < 0$ для всех $i \in I_n$.

Комбинирование достат. условий для выделения ограничений

Линейные и выпуклые функции ограничений

1. Условие выделения ограничений выполняются в любой точке $x \in X$, если
   - часть функций g_i линейные $(i \in I_l)$,
   - а остальные функции выпуклые $(i \in I_c)$
   - и существует точка $\bar{x} \in X$, что $g_i(\bar{x}) < 0$ для всех $i \in I_c$.
2. Условие выделения ограничений также выполняются в точке $x^0 \in X$, если
   - часть функций g_i линейные $(i \in I_l)$,
   - а градиенты $\nabla g_i(x^0)$ линейно независимы для нелинейных ограничений $i \in I_n$
   - и существует точка $\bar{x} \in X$, что $g_i(\bar{x}) < 0$ для всех $i \in I_n$.

Комбинирование достат. условий для выделения ограничений

Линейные и выпуклые функции ограничений

1. Условие выделения ограничений выполняются в любой точке $x \in X$, если
   - часть функций g_i линейные $(i \in I_l)$,
   - а остальные функции выпуклые $(i \in I_c)$
   - и существует точка $\bar{x} \in X$, что $g_i(\bar{x}) < 0$ для всех $i \in I_c$.
2. Условие выделения ограничений также выполняются в точке $x^0 \in X$, если
   - часть функций g_i линейные $(i \in I_l)$,
   - а градиенты $\nabla g_i(x^0)$ линейно независимы для нелинейных ограничений $i \in I_n$
   - и существует точка $\bar{x} \in X$, что $g_i(\bar{x}) < 0$ для всех $i \in I_n$.

План лекции

Условия Куна - Таккера

Теорема (Куна — Таккера)

Предположим, что все функции f и g_i $(i = 1, \dots, m)$ непрерывно дифференцируемы и в точке $x^0 \in X$ выполняется условие выделения ограничений. *Если x^0 есть точка локального минимума, то существуют такие числа (множители Куна-Таккера) $\lambda_i \geq 0$ $(i = 1, \dots, m)$, что*

$$\nabla f(x^0) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^0) = 0,$$

$$\lambda_i g_i(x^0) = 0, \quad i = 1, \dots, m. \quad (\text{усл. доп. нежесткости})$$

Перейти к доказательству

Условия Куна - Таккера

Теорема (Куна — Таккера)

Предположим, что все функции f и g_i $(i = 1, \dots, m)$ непрерывно дифференцируемы и в точке $x^0 \in X$ выполняется условие выделения ограничений. Если x^0 есть точка локального минимума, то существуют такие числа (множители Куна-Таккера) $\lambda_i \geq 0$ $(i = 1, \dots, m)$, что

$$\nabla f(x^0) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^0) = 0,$$

$$\lambda_i g_i(x^0) = 0, \quad i = 1, \dots, m. \quad \text{(усл. доп. нежесткости)}$$

Перейти к доказательству

Условия Куна - Таккера

Теорема (Куна — Таккера)

Предположим, что все функции f и g_i $(i = 1, \ldots, m)$ непрерывно дифференцируемы и в точке $x^0 \in X$ выполняется условие выделения ограничений. Если x^0 есть точка локального минимума, то существуют такие числа (множители Куна-Таккера) $\lambda_i \geq 0$ $(i = 1, \ldots, m)$, что

$$\nabla f(x^0) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^0) = 0,$$

$$\lambda_i g_i(x^0) = 0, \quad i = 1, \ldots, m. \quad (\text{усл. доп. нежесткости})$$

▸ Перейти к доказательству

Условия Куна - Таккера

Теорема (Куна — Таккера)

Предположим, что все функции f *и* g_i $(i = 1, \dots, m)$ *непрерывно дифференцируемы и в точке* $x^0 \in X$ *выполняется условие выделения ограничений. Если* x^0 *есть точка локального минимума, то существуют такие числа (множители Куна-Таккера)* $\lambda_i \geq 0$ $(i = 1, \dots, m)$, *что*

$$\nabla f(x^0) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^0) = 0,$$

$$\lambda_i g_i(x^0) = 0, \quad i = 1, \dots, m. \quad \textit{(усл. доп. нежесткости)}$$

Перейти к доказательству

Ограничения разных знаков

Утверждение

Если в задаче оптимизации присутствует ограничение

1. $g_i(x) = 0$, *то множитель* λ_i *может быть любого знака, т. е.* $\lambda_i \in \mathbb{R}$;
2. $g_i(x) \geq 0$, *то множитель* λ_i *должен быть неположительным, т. е.* $\lambda_i \leq 0$.

Перейти к доказательству

Ограничения разных знаков

Утверждение

Если в задаче оптимизации присутствует ограничение

1. $g_i(x) = 0$, *то множитель* λ_i *может быть любого знака, т. е.* $\lambda_i \in \mathbb{R}$;
2. $g_i(x) \geq 0$, *то множитель* λ_i *должен быть неположительным, т. е.* $\lambda_i \leq 0$.

Перейти к доказательству

Ограничения разных знаков

Утверждение

Если в задаче оптимизации присутствует ограничение

1. $g_i(x) = 0$, *то множитель* λ_i *может быть любого знака, т. е.* $\lambda_i \in \mathbb{R}$;
2. $g_i(x) \geq 0$, *то множитель* λ_i *должен быть неположительным, т. е.* $\lambda_i \leq 0$.

Перейти к доказательству

Ограничения разных знаков

Утверждение

Если в задаче оптимизации присутствует ограничение

1. $g_i(x) = 0$, *то множитель* λ_i *может быть любого знака, т. е.* $\lambda_i \in \mathbb{R}$;
2. $g_i(x) \geq 0$, *то множитель* λ_i *должен быть неположительным, т. е.* $\lambda_i \leq 0$.

▸ Перейти к доказательству

Стационарные точки

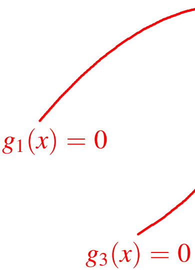
Определение

Точка $x^0 \in X$, которая удовлетворяет условиям Куна — Таккера называется *стационарной точкой*.

План лекции

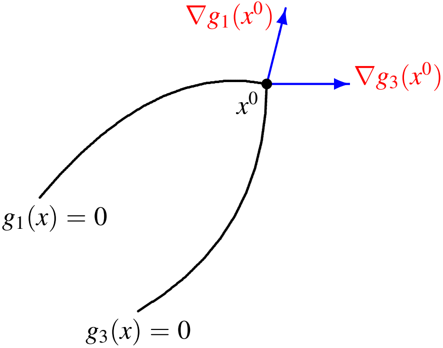
Геометр. интерпретация условий Куна - Таккера

- В точке x^0 локального минимума обращаются в равенство два нер-ва $g_1(x) \le 0$ и $g_3(x) \le 0$; поэтому $I(x^0) = \{1, 3\}$.
- Вектор $-\nabla f(x^0)$ составляет тупой угол с любым вектором $y \in C_{\text{ad}}(x^0) = T(x^0)$.
- Поэтому $-\nabla f(x^0) = \lambda_1 \nabla g_1(x^0) + \lambda_3 \nabla g_3(x^0)$ для некоторых $\lambda_1 > 0$ и $\lambda_3 > 0$.



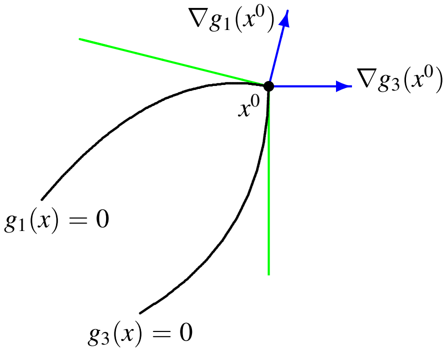
Геометр. интерпретация условий Куна - Таккера

- В точке x^0 локального минимума обращаются в равенство два нер-ва $g_1(x) \le 0$ и $g_3(x) \le 0$; поэтому $I(x^0) = \{1, 3\}$.
- Вектор $-\nabla f(x^0)$ составляет тупой угол с любым вектором $y \in C_{\text{ad}}(x^0) = T(x^0)$.
- Поэтому $-\nabla f(x^0) = \lambda_1 \nabla g_1(x^0) + \lambda_3 \nabla g_3(x^0)$ для некоторых $\lambda_1 > 0$ и $\lambda_3 > 0$.



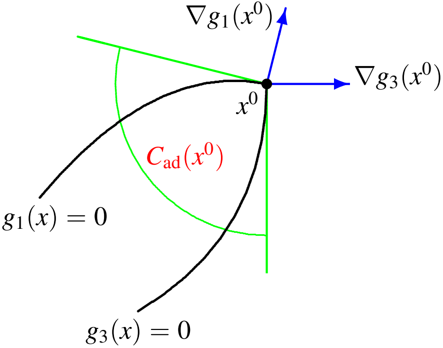
Геометр. интерпретация условий Куна - Таккера

- В точке x^0 локального минимума обращаются в равенство два нер-ва $g_1(x) \le 0$ и $g_3(x) \le 0$; поэтому $I(x^0) = \{1, 3\}$.
- Вектор $-\nabla f(x^0)$ составляет тупой угол с любым вектором $y \in C_{\text{ad}}(x^0) = T(x^0)$.
- Поэтому $-\nabla f(x^0) = \lambda_1 \nabla g_1(x^0) + \lambda_3 \nabla g_3(x^0)$ для некоторых $\lambda_1 > 0$ и $\lambda_3 > 0$.



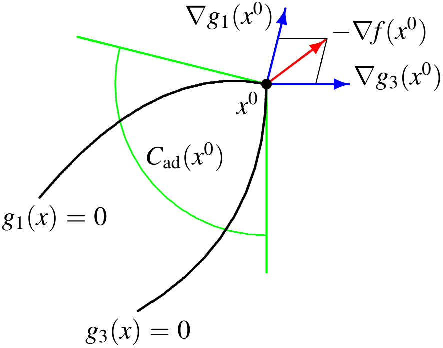
Геометр. интерпретация условий Куна - Таккера

- В точке x^0 локального минимума обращаются в равенство два нер-ва $g_1(x) \leq 0$ и $g_3(x) \leq 0$; поэтому $I(x^0) = \{1, 3\}$.
- Вектор $-\nabla f(x^0)$ составляет тупой угол с любым вектором $y \in C_{\text{ad}}(x^0) = T(x^0)$.
- Поэтому $-\nabla f(x^0) = \lambda_1 \nabla g_1(x^0) + \lambda_3 \nabla g_3(x^0)$ для некоторых $\lambda_1 > 0$ и $\lambda_3 > 0$.



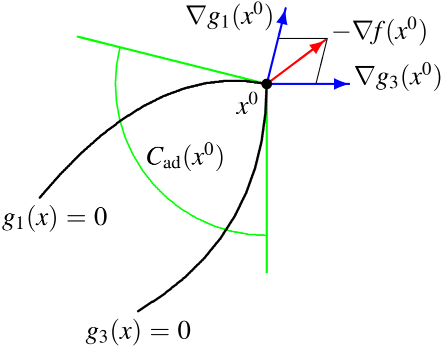
Геометр. интерпретация условий Куна - Таккера

- В точке x^0 локального минимума обращаются в равенство два нер-ва $g_1(x) \le 0$ и $g_3(x) \le 0$; поэтому $I(x^0) = \{1, 3\}$.
- Вектор $-\nabla f(x^0)$ составляет тупой угол с любым вектором $y \in C_{\text{ad}}(x^0) = T(x^0)$.
- Поэтому $-\nabla f(x^0) = \lambda_1 \nabla g_1(x^0) + \lambda_3 \nabla g_3(x^0)$ для некоторых $\lambda_1 > 0$ и $\lambda_3 > 0$.



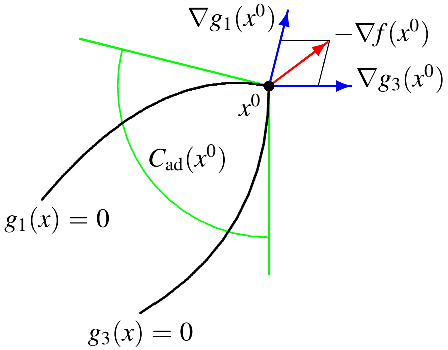
Геометр. интерпретация условий Куна - Таккера

- В точке x^0 локального минимума обращаются в равенство два нер-ва $g_1(x) \leq 0$ и $g_3(x) \leq 0$; поэтому $I(x^0) = \{1, 3\}$.
- Вектор $-\nabla f(x^0)$ составляет тупой угол с любым вектором $y \in C_{\mathrm{ad}}(x^0) = T(x^0)$.
- Поэтому $-\nabla f(x^0) = \lambda_1 \nabla g_1(x^0) + \lambda_3 \nabla g_3(x^0)$ для некоторых $\lambda_1 > 0$ и $\lambda_3 > 0$.



Геометр. интерпретация условий Куна - Таккера

- В точке x^0 локального минимума обращаются в равенство два нер-ва $g_1(x) \leq 0$ и $g_3(x) \leq 0$; поэтому $I(x^0) = \{1, 3\}$.
- Вектор $-\nabla f(x^0)$ составляет тупой угол с любым вектором $y \in C_{\mathrm{ad}}(x^0) = T(x^0)$.
- Поэтому $-\nabla f(x^0) = \lambda_1 \nabla g_1(x^0) + \lambda_3 \nabla g_3(x^0)$ для некоторых $\lambda_1 > 0$ и $\lambda_3 > 0$.



Доказательство включения $C_{\text{ad}}(x^0) \subseteq T(x^0)$

- Пусть $\varphi(\theta)$ есть дуга допустимой кривой в x^0, а $y = \frac{d\varphi}{d\theta}(0)$ — допустимое направление в x^0.
- Для $i \in I(x^0)$ при достаточно малом $\theta > 0$ должно выполняться неравенство $g_i(\varphi(\theta)) \leq 0$.
- Разлагая функцию $g_i(\varphi(\theta))$ в ряд Тейлора в окрестности точки $\theta = 0$, получим неравенство

$$g_i(x^0) + \theta \nabla g_i^T(x^0) \frac{d\varphi}{d\theta}(0) + \theta o(\theta) \leq 0, \quad \text{где} \ \lim_{\theta \to 0} o(\theta) = 0.$$

- В силу равенства $g_i(x^0) = 0$ необходимо, чтобы направление $y = \frac{d\varphi}{d\theta}(0)$ удовлетворяло условию $\nabla g_i^T(x^0) y \leq 0 \quad \forall\, i \in I(x^0)$.
- Из $y \in C_{\text{ad}}(x^0)$ следует, что $y \in T(x^0)$, и $C_{\text{ad}}(x^0) \subseteq T(x^0)$. $\square$

Вернуться к формулировке

Доказательство включения $C_{\text{ad}}(x^0) \subseteq T(x^0)$

- Пусть $\varphi(\theta)$ есть дуга допустимой кривой в x^0, а $y = \frac{d\varphi}{d\theta}(0)$ — допустимое направление в x^0.
- Для $i \in I(x^0)$ при достаточно малом $\theta > 0$ должно выполняться неравенство $g_i(\varphi(\theta)) \le 0$.
- Разлагая функцию $g_i(\varphi(\theta))$ в ряд Тейлора в окрестности точки $\theta = 0$, получим неравенство

$$g_i(x^0) + \theta \nabla g_i^T(x^0) \frac{d\varphi}{d\theta}(0) + \theta o(\theta) \le 0, \quad \text{где} \lim_{\theta \to 0} o(\theta) = 0.$$

- В силу равенства $g_i(x^0) = 0$ необходимо, чтобы направление $y = \frac{d\varphi}{d\theta}(0)$ удовлетворяло условию $\nabla g_i^T(x^0) y \le 0 \quad \forall\, i \in I(x^0)$.
- Из $y \in C_{\text{ad}}(x^0)$ следует, что $y \in T(x^0)$, и $C_{\text{ad}}(x^0) \subseteq T(x^0)$. $\square$

Вернуться к формулировке

Доказательство включения $C_{\text{ad}}(x^0) \subseteq T(x^0)$

- Пусть $\varphi(\theta)$ есть дуга допустимой кривой в x^0, а $y = \frac{d\varphi}{d\theta}(0)$ — допустимое направление в x^0.
- Для $i \in I(x^0)$ при достаточно малом $\theta > 0$ должно выполняться неравенство $g_i(\varphi(\theta)) \le 0$.
- Разлагая функцию $g_i(\varphi(\theta))$ в ряд Тейлора в окрестности точки $\theta = 0$, получим неравенство

$$g_i(x^0) + \theta \nabla g_i^T(x^0)\frac{d\varphi}{d\theta}(0) + \theta o(\theta) \le 0, \quad \text{где } \lim_{\theta\to 0} o(\theta) = 0.$$

- В силу равенства $g_i(x^0) = 0$ необходимо, чтобы направление $y = \frac{d\varphi}{d\theta}(0)$ удовлетворяло условию $\nabla g_i^T(x^0)y \le 0 \quad \forall\, i \in I(x^0)$.
- Из $y \in C_{\text{ad}}(x^0)$ следует, что $y \in T(x^0)$, и $C_{\text{ad}}(x^0) \subseteq T(x^0)$. $\square$

Вернуться к формулировке

Доказательство включения $C_{\text{ad}}(x^0) \subseteq T(x^0)$

- Пусть $\varphi(\theta)$ есть дуга допустимой кривой в x^0, а $y = \frac{d\varphi}{d\theta}(0)$ — допустимое направление в x^0.
- Для $i \in I(x^0)$ при достаточно малом $\theta > 0$ должно выполняться неравенство $g_i(\varphi(\theta)) \le 0$.
- Разлагая функцию $g_i(\varphi(\theta))$ в ряд Тейлора в окрестности точки $\theta = 0$, получим неравенство

$$g_i(x^0) + \theta \nabla g_i^T(x^0) \frac{d\varphi}{d\theta}(0) + \theta o(\theta) \le 0, \quad \text{где } \lim_{\theta \to 0} o(\theta) = 0.$$

- В силу равенства $g_i(x^0) = 0$ необходимо, чтобы направление $y = \frac{d\varphi}{d\theta}(0)$ удовлетворяло условию $\nabla g_i^T(x^0) y \le 0 \quad \forall\, i \in I(x^0)$.
- Из $y \in C_{\text{ad}}(x^0)$ следует, что $y \in T(x^0)$, и $C_{\text{ad}}(x^0) \subseteq T(x^0)$. $\square$

Вернуться к формулировке

Доказательство включения $C_{\mathrm{ad}}(x^0) \subseteq T(x^0)$

- Пусть $\varphi(\theta)$ есть дуга допустимой кривой в x^0, а $y = \frac{d\varphi}{d\theta}(0)$ — допустимое направление в x^0.
- Для $i \in I(x^0)$ при достаточно малом $\theta > 0$ должно выполняться неравенство $g_i(\varphi(\theta)) \le 0$.
- Разлагая функцию $g_i(\varphi(\theta))$ в ряд Тейлора в окрестности точки $\theta = 0$, получим неравенство

$$g_i(x^0) + \theta \nabla g_i^T(x^0) \frac{d\varphi}{d\theta}(0) + \theta o(\theta) \le 0, \quad \text{где} \ \lim_{\theta \to 0} o(\theta) = 0.$$

- В силу равенства $g_i(x^0) = 0$ необходимо, чтобы направление $y = \frac{d\varphi}{d\theta}(0)$ удовлетворяло условию $\nabla g_i^T(x^0) y \le 0 \quad \forall\, i \in I(x^0)$.
- Из $y \in C_{\mathrm{ad}}(x^0)$ следует, что $y \in T(x^0)$, и $C_{\mathrm{ad}}(x^0) \subseteq T(x^0)$. $\square$

Вернуться к формулировке

Док-во достат. условий для выделения ограничений

Обоснование условия а) элементарно. Поэтому мы ограничимся обоснованием условий б) и в).

- Рассмотрим конус *возможных направлений* в точке $x^0 \in X$: $C_{\mathrm{pd}}(x^0) \stackrel{\mathrm{def}}{=} \{y \in \mathbb{R}^n : \left(\nabla g_i(x^0)\right)^T y < 0,\ i \in I(x^0)\}$.
- Понятно, что $C_{\mathrm{pd}}(x^0) \subset C_{\mathrm{ad}}(x^0)$ и $\mathrm{cl}(C_{\mathrm{pd}}(x^0)) \subset \mathrm{cl}(C_{\mathrm{ad}}(x^0))$. Так как $C_{\mathrm{ad}}(x^0) \subseteq T(x^0)$, для завершения доказательства достаточно показать, что при выполнении условий б) и в) справедливо равенство $\mathrm{cl}(C_{\mathrm{pd}}) = T(x^0)$.
- Сначала докажем, что нужное равенство справедливо, если $C_{\mathrm{pd}}(x^0) \neq \emptyset$. Действительно, для $\bar{y} \in C_{\mathrm{pd}}(x^0)$ выполняются неравенства: $\left(\nabla g_i(x^0)\right)^T \bar{y} < 0, \quad i \in I(x^0)$.
- Для любого $y \in T(x^0)$ выполняются неравенства:

$$\left(\nabla g_i(x^0)\right)^T y \leq 0, \quad i \in I(x^0).$$

Поэтому $y(\lambda) \stackrel{\mathrm{def}}{=} (1-\lambda)\bar{y} + \lambda y \in C_{\mathrm{pd}}(x^0) \quad \forall\, \lambda \in [0,1)$.
- Следовательно, $\lim_{k\to\infty} y(\lambda_k) = y$, если все $\lambda_k \in [0,1)$ и

Док-во достат. условий для выделения ограничений

- Рассмотрим конус *возможных направлений* в точке $x^0 \in X$: $C_{\rm pd}(x^0) \stackrel{\rm def}{=} \{y \in \mathbb{R}^n : \left(\nabla g_i(x^0)\right)^T y < 0,\ i \in I(x^0)\}$.
- Понятно, что $C_{\rm pd}(x^0) \subset C_{\rm ad}(x^0)$ и $\mathrm{cl}(C_{\rm pd}(x^0)) \subset \mathrm{cl}(C_{\rm ad}(x^0))$. Так как $C_{\rm ad}(x^0) \subseteq T(x^0)$, для завершения доказательства достаточно показать, что при выполнении условий б) и в) справедливо равенство $\mathrm{cl}(C_{\rm pd}) = T(x^0)$.
- Сначала докажем, что нужное равенство справедливо, если $C_{\rm pd}(x^0) \neq \emptyset$. Действительно, для $\bar y \in C_{\rm pd}(x^0)$ выполняются неравенства: $\left(\nabla g_i(x^0)\right)^T \bar y < 0, \quad i \in I(x^0)$.
- Для любого $y \in T(x^0)$ выполняются неравенства:
$$\left(\nabla g_i(x^0)\right)^T y \le 0, \quad i \in I(x^0).$$
Поэтому $y(\lambda) \stackrel{\rm def}{=} (1-\lambda)\bar y + \lambda y \in C_{\rm pd}(x^0) \quad \forall\, \lambda \in [0,1)$.
- Следовательно, $\lim_{k\to\infty} y(\lambda_k) = y$, если все $\lambda_k \in [0,1)$ и $\lim_{k\to\infty} \lambda_k = 1$. Это означает, что $\mathrm{cl}(C_{\rm pd}) = T(x^0)$.

Вернуться к формулировке

Док-во достат. условий для выделения ограничений

- Рассмотрим конус *возможных направлений* в точке $x^0 \in X$: $C_{\text{pd}}(x^0) \stackrel{\text{def}}{=} \{y \in \mathbb{R}^n : \left(\nabla g_i(x^0)\right)^T y < 0,\ i \in I(x^0)\}$.
- Понятно, что $C_{\text{pd}}(x^0) \subset C_{\text{ad}}(x^0)$ и $\text{cl}(C_{\text{pd}}(x^0)) \subset \text{cl}(C_{\text{ad}}(x^0))$. Так как $C_{\text{ad}}(x^0) \subseteq T(x^0)$, для завершения доказательства достаточно показать, что при выполнении условий б) и в) справедливо равенство $\text{cl}(C_{\text{pd}}) = T(x^0)$.
- Сначала докажем, что нужное равенство справедливо, если $C_{\text{pd}}(x^0) \neq \emptyset$. Действительно, для $\bar{y} \in C_{\text{pd}}(x^0)$ выполняются неравенства: $\left(\nabla g_i(x^0)\right)^T \bar{y} < 0, \quad i \in I(x^0)$.
- Для любого $y \in T(x^0)$ выполняются неравенства:
$$\left(\nabla g_i(x^0)\right)^T y \le 0, \quad i \in I(x^0).$$
Поэтому $y(\lambda) \stackrel{\text{def}}{=} (1-\lambda)\bar{y} + \lambda y \in C_{\text{pd}}(x^0) \quad \forall\, \lambda \in [0,1)$.
- Следовательно, $\lim_{k\to\infty} y(\lambda_k) = y$, если все $\lambda_k \in [0,1)$ и $\lim_{k\to\infty} \lambda_k = 1$. Это означает, что $\text{cl}(C_{\text{pd}}) = T(x^0)$.

Вернуться к формулировке

Док-во достат. условий для выделения ограничений

- Рассмотрим конус *возможных направлений* в точке $x^0 \in X$: $C_{\mathrm{pd}}(x^0) \stackrel{\mathrm{def}}{=} \{y \in \mathbb{R}^n : \left(\nabla g_i(x^0)\right)^T y < 0,\ i \in I(x^0)\}$.
- Понятно, что $C_{\mathrm{pd}}(x^0) \subset C_{\mathrm{ad}}(x^0)$ и $\mathrm{cl}(C_{\mathrm{pd}}(x^0)) \subset \mathrm{cl}(C_{\mathrm{ad}}(x^0))$. Так как $C_{\mathrm{ad}}(x^0) \subseteq T(x^0)$, для завершения доказательства достаточно показать, что при выполнении условий б) и в) справедливо равенство $\mathrm{cl}(C_{\mathrm{pd}}) = T(x^0)$.
- Сначала докажем, что нужное равенство справедливо, если $C_{\mathrm{pd}}(x^0) \neq \emptyset$. Действительно, для $\bar{y} \in C_{\mathrm{pd}}(x^0)$ выполняются неравенства: $\left(\nabla g_i(x^0)\right)^T \bar{y} < 0, \quad i \in I(x^0)$.
- Для любого $y \in T(x^0)$ выполняются неравенства:
$$\left(\nabla g_i(x^0)\right)^T y \le 0, \quad i \in I(x^0).$$
Поэтому $y(\lambda) \stackrel{\mathrm{def}}{=} (1-\lambda)\bar{y} + \lambda y \in C_{\mathrm{pd}}(x^0) \quad \forall\, \lambda \in [0,1)$.
- Следовательно, $\lim_{k\to\infty} y(\lambda_k) = y$, если все $\lambda_k \in [0,1)$ и $\lim_{k\to\infty} \lambda_k = 1$. Это означает, что $\mathrm{cl}(C_{\mathrm{pd}}) = T(x^0)$.

Вернуться к формулировке

Док-во достат. условий для выделения ограничений

- Рассмотрим конус *возможных направлений* в точке $x^0 \in X$: $C_{\rm pd}(x^0) \stackrel{\rm def}{=} \{y \in \mathbb{R}^n : \left(\nabla g_i(x^0)\right)^T y < 0,\ i \in I(x^0)\}$.
- Понятно, что $C_{\rm pd}(x^0) \subset C_{\rm ad}(x^0)$ и $\mathrm{cl}(C_{\rm pd}(x^0)) \subset \mathrm{cl}(C_{\rm ad}(x^0))$. Так как $C_{\rm ad}(x^0) \subseteq T(x^0)$, для завершения доказательства достаточно показать, что при выполнении условий б) и в) справедливо равенство $\mathrm{cl}(C_{\rm pd}) = T(x^0)$.
- Сначала докажем, что нужное равенство справедливо, если $C_{\rm pd}(x^0) \neq \emptyset$. Действительно, для $\bar{y} \in C_{\rm pd}(x^0)$ выполняются неравенства: $\left(\nabla g_i(x^0)\right)^T \bar{y} < 0, \quad i \in I(x^0)$.
- Для любого $y \in T(x^0)$ выполняются неравенства:
$$\left(\nabla g_i(x^0)\right)^T y \le 0, \quad i \in I(x^0).$$
Поэтому $y(\lambda) \stackrel{\rm def}{=} (1-\lambda)\bar{y} + \lambda y \in C_{\rm pd}(x^0) \quad \forall\, \lambda \in [0, 1)$.
- Следовательно, $\lim_{k\to\infty} y(\lambda_k) = y$, если все $\lambda_k \in [0,1)$ и $\lim_{k\to\infty} \lambda_k = 1$. Это означает, что $\mathrm{cl}(C_{\rm pd}) = T(x^0)$.

Вернуться к формулировке

Док-во достат. условий для выделения ограничений

- Рассмотрим конус *возможных направлений* в точке $x^0 \in X$: $C_{\text{pd}}(x^0) \stackrel{\text{def}}{=} \{y \in \mathbb{R}^n : \left(\nabla g_i(x^0)\right)^T y < 0,\ i \in I(x^0)\}$.
- Понятно, что $C_{\text{pd}}(x^0) \subset C_{\text{ad}}(x^0)$ и $\text{cl}(C_{\text{pd}}(x^0)) \subset \text{cl}(C_{\text{ad}}(x^0))$. Так как $C_{\text{ad}}(x^0) \subseteq T(x^0)$, для завершения доказательства достаточно показать, что при выполнении условий б) и в) справедливо равенство $\text{cl}(C_{\text{pd}}) = T(x^0)$.
- Сначала докажем, что нужное равенство справедливо, если $C_{\text{pd}}(x^0) \neq \emptyset$. Действительно, для $\bar{y} \in C_{\text{pd}}(x^0)$ выполняются неравенства: $\left(\nabla g_i(x^0)\right)^T \bar{y} < 0, \quad i \in I(x^0)$.
- Для любого $y \in T(x^0)$ выполняются неравенства:
$$\left(\nabla g_i(x^0)\right)^T y \le 0, \quad i \in I(x^0).$$
Поэтому $y(\lambda) \stackrel{\text{def}}{=} (1-\lambda)\bar{y} + \lambda y \in C_{\text{pd}}(x^0) \quad \forall\, \lambda \in [0,1)$.
- Следовательно, $\lim_{k\to\infty} y(\lambda_k) = y$, если все $\lambda_k \in [0,1)$ и $\lim_{k\to\infty} \lambda_k = 1$. Это означает, что $\text{cl}(C_{\text{pd}}) = T(x^0)$.

Вернуться к формулировке

Док-во достат. условий для выделения ограничений

- Предположим теперь, что выполняется условие б). Тогда существует $\bar{x} \in X$, что
$$g_i(\bar{x}) < 0, \quad i \in I.$$
- В силу выпуклости функций g_i справедливы нер-ва: $\left(\nabla g_i(x^0)\right)^T (\bar{x} - x^0) \le g_i(\bar{x}) - g_i(x^0) = g_i(\bar{x}) < 0, \quad i \in I(x^0),$ т. е. $\bar{y} = \bar{x} - x^0 \in C_{\mathrm{fs}}(x^0)$ и, следовательно, $C_{\mathrm{fs}}(x^0) \ne \emptyset$.
- Теперь предположим, что в точке x^0 выполняется условие в). Тогда векторное равенство
$$\sum_{i \in I(x^0)} \lambda_i \nabla g_i(x^0) = \mathbf{0}$$
несовместно относительно неизвестных λ_i.
- По лемме Фаркаша существует такой вектор $y \in \mathbb{R}^n$, что
$$\left(\nabla g_i(x^0)\right)^T y < 0, \quad i \in I(x^0),$$
и, значит, $C_{\mathrm{fs}}(x^0) \ne \emptyset$. □

Вернуться к формулировке

Док-во достат. условий для выделения ограничений

- Предположим теперь, что выполняется условие б). Тогда существует $\bar{x} \in X$, что
$$g_i(\bar{x}) < 0, \quad i \in I.$$
- В силу выпуклости функций g_i справедливы нер-ва: $\left(\nabla g_i(x^0)\right)^T (\bar{x} - x^0) \le g_i(\bar{x}) - g_i(x^0) = g_i(\bar{x}) < 0, \quad i \in I(x^0),$ т. е. $\bar{y} = \bar{x} - x^0 \in C_{\mathrm{fs}}(x^0)$ и, следовательно, $C_{\mathrm{fs}}(x^0) \ne \emptyset$.
- Теперь предположим, что в точке x^0 выполняется условие в). Тогда векторное равенство
$$\sum_{i \in I(x^0)} \lambda_i \nabla g_i(x^0) = \mathbf{0}$$
несовместно относительно неизвестных λ_i.
- По лемме Фаркаша существует такой вектор $y \in \mathbb{R}^n$, что
$$\left(\nabla g_i(x^0)\right)^T y < 0, \quad i \in I(x^0),$$
и, значит, $C_{\mathrm{fs}}(x^0) \ne \emptyset$. □

Вернуться к формулировке

Док-во достат. условий для выделения ограничений

- Предположим теперь, что выполняется условие б). Тогда существует $\bar{x} \in X$, что
$$g_i(\bar{x}) < 0, \quad i \in I.$$
- В силу выпуклости функций g_i справедливы нер-ва: $\left(\nabla g_i(x^0)\right)^T (\bar{x} - x^0) \le g_i(\bar{x}) - g_i(x^0) = g_i(\bar{x}) < 0, \quad i \in I(x^0),$ т. е. $\bar{y} = \bar{x} - x^0 \in C_{\text{fs}}(x^0)$ и, следовательно, $C_{\text{fs}}(x^0) \neq \emptyset$.
- Теперь предположим, что в точке x^0 выполняется условие в). Тогда векторное равенство
$$\sum_{i \in I(x^0)} \lambda_i \nabla g_i(x^0) = \mathbf{0}$$
несовместно относительно неизвестных λ_i.
- По лемме Фаркаша существует такой вектор $y \in \mathbb{R}^n$, что
$$\left(\nabla g_i(x^0)\right)^T y < 0, \quad i \in I(x^0),$$
и, значит, $C_{\text{fs}}(x^0) \neq \emptyset$. $\square$

Док-во достат. условий для выделения ограничений

- Предположим теперь, что выполняется условие б). Тогда существует $\bar{x} \in X$, что
$$g_i(\bar{x}) < 0, \quad i \in I.$$
- В силу выпуклости функций g_i справедливы нер-ва: $\left(\nabla g_i(x^0)\right)^T (\bar{x} - x^0) \le g_i(\bar{x}) - g_i(x^0) = g_i(\bar{x}) < 0, \quad i \in I(x^0),$ т. е. $\bar{y} = \bar{x} - x^0 \in C_{\text{fs}}(x^0)$ и, следовательно, $C_{\text{fs}}(x^0) \ne \emptyset$.
- Теперь предположим, что в точке x^0 выполняется условие в). Тогда векторное равенство
$$\sum_{i \in I(x^0)} \lambda_i \nabla g_i(x^0) = \mathbf{0}$$
несовместно относительно неизвестных λ_i.
- По лемме Фаркаша существует такой вектор $y \in \mathbb{R}^n$, что
$$\left(\nabla g_i(x^0)\right)^T y < 0, \quad i \in I(x^0),$$
и, значит, $C_{\text{fs}}(x^0) \ne \emptyset$. $\square$

Вернуться к формулировке

Доказательство теоремы Куна - Таккера

- Точка x^0 не является локальным минимумом задачи
$$\min\{f(x) : \; g_i(x) \le 0, \; i \in I; \; x \in \mathbb{R}^n\},$$
если существует такое допустимое направление, вдоль которого целевая функция убывает.
- Алгебраически это условие эквивалентно тому, что имеет решение система неравенств
$$\left(\nabla f(x^0)\right)^T y \le -1; \quad \left(\nabla g_i(x^0)\right)^T y \le 0, \quad i \in I(x^0).$$
- По лемме Фаркаша, эта система несовместна тогда и только тогда, когда существуют неотрицательные числа λ_0 и λ_i $(i \in I(x^0))$, такие, что
$$\lambda_0 \nabla f(x^0) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^0) = 0 \quad \text{и} \quad -1 \cdot \lambda_0 + \sum_{i=1}^{m} 0 \cdot \lambda_i = -1.$$
Из последнего равенства следует, что $\lambda_0 = 1$.
- Если положить $\lambda_i = 0$ для всех $i \in I \setminus I(x^0)$, то получим:
$$\lambda_i g_i(x^0) = 0, \quad i = 1, \dots, m. \qquad \square$$

Доказательство теоремы Куна - Таккера

- Точка x^0 не является локальным минимумом задачи
$$\min\{f(x) : \ g_i(x) \le 0, \ i \in I; \ x \in \mathbb{R}^n\},$$
если существует такое допустимое направление, вдоль которого целевая функция убывает.
- Алгебраически это условие эквивалентно тому, что имеет решение система неравенств
$$\left(\nabla f(x^0)\right)^T y \le -1; \quad \left(\nabla g_i(x^0)\right)^T y \le 0, \quad i \in I(x^0).$$
- По лемме Фаркаша, эта система несовместна тогда и только тогда, когда существуют неотрицательные числа λ_0 и λ_i $(i \in I(x^0))$, такие, что
$$\lambda_0 \nabla f(x^0) + \sum_{i=1}^m \lambda_i \nabla g_i(x^0) = 0 \quad \text{и} \quad -1 \cdot \lambda_0 + \sum_{i=1}^m 0 \cdot \lambda_i = -1.$$
Из последнего равенства следует, что $\lambda_0 = 1$.
- Если положить $\lambda_i = 0$ для всех $i \in I \setminus I(x^0)$, то получим:
$$\lambda_i g_i(x^0) = 0, \quad i = 1, \dots, m.$$
□

Доказательство теоремы Куна - Таккера

- Точка x^0 не является локальным минимумом задачи
$$\min\{f(x) : \ g_i(x) \le 0, \ i \in I; \ x \in \mathbb{R}^n\},$$
если существует такое допустимое направление, вдоль которого целевая функция убывает.
- Алгебраически это условие эквивалентно тому, что имеет решение система неравенств
$$\left(\nabla f(x^0)\right)^T y \le -1; \quad \left(\nabla g_i(x^0)\right)^T y \le 0, \quad i \in I(x^0).$$
- По лемме Фаркаша, эта система несовместна тогда и только тогда, когда существуют неотрицательные числа λ_0 и λ_i ($i \in I(x^0)$), такие, что
$$\lambda_0 \nabla f(x^0) + \sum_{i=1}^m \lambda_i \nabla g_i(x^0) = 0 \quad \text{и} \quad -1 \cdot \lambda_0 + \sum_{i=1}^m 0 \cdot \lambda_i = -1.$$
Из последнего равенства следует, что $\lambda_0 = 1$.
- Если положить $\lambda_i = 0$ для всех $i \in I \setminus I(x^0)$, то получим:
$$\lambda_i g_i(x^0) = 0, \quad i = 1, \ldots, m. \qquad \square$$

Доказательство теоремы Куна - Таккера

- Точка x^0 не является локальным минимумом задачи
$$\min\{f(x) : \; g_i(x) \le 0, \; i \in I; \; x \in \mathbb{R}^n\},$$
если существует такое допустимое направление, вдоль которого целевая функция убывает.
- Алгебраически это условие эквивалентно тому, что имеет решение система неравенств
$$\left(\nabla f(x^0)\right)^T y \le -1; \quad \left(\nabla g_i(x^0)\right)^T y \le 0, \quad i \in I(x^0).$$
- По лемме Фаркаша, эта система несовместна тогда и только тогда, когда существуют неотрицательные числа λ_0 и λ_i $(i \in I(x^0))$, такие, что
$$\lambda_0 \nabla f(x^0) + \sum_{i=1}^m \lambda_i \nabla g_i(x^0) = 0 \quad \text{и} \quad -1 \cdot \lambda_0 + \sum_{i=1}^m 0 \cdot \lambda_i = -1.$$
Из последнего равенства следует, что $\lambda_0 = 1$.
- Если положить $\lambda_i = 0$ для всех $i \in I \setminus I(x^0)$, то получим:
$$\lambda_i g_i(x^0) = 0, \quad i = 1, \ldots, m.$$
$\square$

Множители для ограничений в форме равенств

- Уравнение $g_i(x) = 0$ можно заменить двумя неравенствами $g_i(x) \le 0$ и $-g_i(x) \le 0$, которым ставятся в соответствие два неотрицательных множителя λ_i^+ и λ_i^-.
- Тогда в векторном равенстве

$$f(x^0) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^0) = 0, \quad (*)$$

- будут присутствовать два слагаемых $\lambda_i^+ \nabla g_i(x^0)$ и $-\lambda_i^- \nabla g_i(x^0)$, которые можно заменить их суммой $(\lambda_i^+ - \lambda_i^-) \nabla g_i(x^0)$.
- Вводя новый множитель $\lambda_i = \lambda_i^+ - \lambda_i^-$, мы вернемся к представлению $(*)$, где ограничение $g_i(x) = 0$ представлено одним слагаемым $\lambda_i \nabla g_i(x^0)$, но со множителем λ_i произвольного знака.

Вернуться к формулировке

Множители для ограничений в форме равенств

- Уравнение $g_i(x) = 0$ можно заменить двумя неравенствами $g_i(x) \le 0$ и $-g_i(x) \le 0$, которым ставятся в соответствие два неотрицательных множителя λ_i^+ и λ_i^-.
- Тогда в векторном равенстве

$$f(x^0) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^0) = 0, \quad (*)$$

- будут присутствовать два слагаемых $\lambda_i^+ \nabla g_i(x^0)$ и $-\lambda_i^- \nabla g_i(x^0)$, которые можно заменить их суммой $(\lambda_i^+ - \lambda_i^-) \nabla g_i(x^0)$.
- Вводя новый множитель $\lambda_i = \lambda_i^+ - \lambda_i^-$, мы вернемся к представлению $(*)$, где ограничение $g_i(x) = 0$ представлено одним слагаемым $\lambda_i \nabla g_i(x^0)$, но со множителем λ_i произвольного знака.

Вернуться к формулировке

Множители для ограничений в форме равенств

- Уравнение $g_i(x) = 0$ можно заменить двумя неравенствами $g_i(x) \le 0$ и $-g_i(x) \le 0$, которым ставятся в соответствие два неотрицательных множителя λ_i^+ и λ_i^-.
- Тогда в векторном равенстве

$$f(x^0) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^0) = 0, \quad (*)$$

- будут присутствовать два слагаемых $\lambda_i^+ \nabla g_i(x^0)$ и $-\lambda_i^- \nabla g_i(x^0)$, которые можно заменить их суммой $(\lambda_i^+ - \lambda_i^-)\nabla g_i(x^0)$.
- Вводя новый множитель $\lambda_i = \lambda_i^+ - \lambda_i^-$, мы вернемся к представлению $(*)$, где ограничение $g_i(x) = 0$ представлено одним слагаемым $\lambda_i \nabla g_i(x^0)$, но со множителем λ_i произвольного знака.

Вернуться к формулировке

Множители для ограничений в форме равенств

- Уравнение $g_i(x) = 0$ можно заменить двумя неравенствами $g_i(x) \le 0$ и $-g_i(x) \le 0$, которым ставятся в соответствие два неотрицательных множителя λ_i^+ и λ_i^-.
- Тогда в векторном равенстве

$$f(x^0) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^0) = 0, \quad (*)$$

- будут присутствовать два слагаемых $\lambda_i^+ \nabla g_i(x^0)$ и $-\lambda_i^- \nabla g_i(x^0)$, которые можно заменить их суммой $(\lambda_i^+ - \lambda_i^-) \nabla g_i(x^0)$.
- Вводя новый множитель $\lambda_i = \lambda_i^+ - \lambda_i^-$, мы вернемся к представлению $(*)$, где ограничение $g_i(x) = 0$ представлено одним слагаемым $\lambda_i \nabla g_i(x^0)$, но со множителем λ_i произвольного знака.

Вернуться к формулировке