ЛЕКЦИЯ № 2

Необходимые условия экстремума

1. **Теорема Фритца-Джона**

2. **Теоремы Куна-Таккера**

3. **Правило множителей Лагранжа**

Необходимые условия оптимальности Фритца–Джона

Далее будем работать со следующей задачей нелинейного программирования
Найти:

$$\min f(x) \tag{1}$$

при условии, что

$$\varphi_i(x) \leq 0, i = \overline{1, m}. \tag{2}$$

Здесь $f, \varphi_i : R^n \longrightarrow R$ и $f, \varphi_i \in C^1$.

Определение 1. Направление s (произвольный ненулевой вектор) в точке $x \in Q$ называется возможным, если существует такое число $\overline{\beta}$, что $x + \beta s \in Q, \ \forall \beta \in [0, \overline{\beta}]$.

Комментарий

Напомню, что множество $\boldsymbol{K}$ называется конусом, если $\forall\ \boldsymbol{\lambda} \geq \mathbf{0}$ и $\forall\ \boldsymbol{x} \in \boldsymbol{K}$ имеем $\boldsymbol{\lambda x} \in \boldsymbol{K}$.

Очевидно, что множество возможных направлений в точке $\boldsymbol{x}$ образует конус, который обозначим как $\boldsymbol{K_f(x)}$.

Необходимые условия оптимальности Фритца–Джона

Ограничение φ_i называется **активным** в точке x, если $\varphi_i(x) = 0$. Обозначим через $I(x)$ множество номеров ограничений активных в данной точке.

Лемма 2. Если вектор $s \neq 0$ удовлетворяет системе

$$(\varphi_i'(x), s) + \sigma \leq 0, i \in I(x),$$

при некотором $\sigma > 0$, то направление s является возможным в точке x.

Доказательство. Можно считать, что $I(x) \neq \varnothing$, т.к. иначе x – внутренняя точка Q и тогда любое направление $s \neq 0$ является возможным.

Необходимые условия оптимальности Фритца–Джона

Если $\boldsymbol{i \notin I(x)}$, то малое перемещение не нарушает строгое ограничение $\boldsymbol{\varphi_i(x) < 0 \Longrightarrow}$ найдется подходящее $\overline{\boldsymbol{\beta}}$.

Пусть $\boldsymbol{i \in I(x)(\equiv \varphi_i(x) = 0)}$. Допустим $\boldsymbol{\varphi_i(x + \beta s) > 0,\ \forall\beta > 0 \Longrightarrow}$ что при $\boldsymbol{\beta \longrightarrow 0}$

$$\boldsymbol{\varphi_i(x + \beta s)/\beta = (\varphi_i(x + \beta s) - \varphi_i(x))/\beta \longrightarrow (\varphi_i'(x), s) \geq 0.}$$

Противоречие. Т.к. по условиям леммы $\boldsymbol{(\varphi_i'(x), s) < 0}$.■

Комментарий

Пусть

$$\boldsymbol{K_<(x) = \{s | s \neq 0} \text{ и } \boldsymbol{(\varphi_i'(x), s) < 0, \forall\, i \in I(x)\}.}$$

Из этого определения следует, что множество $\boldsymbol{K_<(x)}$ конус. Поэтому лемма 2 утверждает, что конус $\boldsymbol{K_<(x)}$ является подмножеством конуса $\boldsymbol{K_f(x)}$, оправдывая название первого из них, как внутренней аппроксимации второго – конуса возможных направлений.

Конус

$$\boldsymbol{K_d(x) = \{s \neq 0 | (f_i'(x), s) < 0\}}$$

называется конусом направлений убывания функции $\boldsymbol{f}$.

Геометрическая форма необходимых условий оптимальности

Теорема 2. Для того, чтобы точка $x \in Q$ являлась точкой локального минимума функции f на множестве Q необходимо, чтобы для любого решения (s, σ) системы

$$(\varphi_i'(x), s) + \sigma \leq 0, i \in I(x), \tag{4}$$

$$(f'(x), s) + \sigma \leq 0, \tag{5}$$

выполнялось условие

$$\sigma \leq 0. \tag{6}$$

Доказательство. Пусть x – локальный минимум функции f на множестве Q и $\exists \ (s, \sigma) : (4), (5)$ и $\sigma > 0$.

Геометрическая форма необходимых условий оптимальности

$\Longrightarrow$ и леммы 2 имеем, что $\boldsymbol{s}$ – возможное направление в точке $\boldsymbol{x}$.

Т.к. $\boldsymbol{f}'(\boldsymbol{x})$–непрерывная функция и $(\boldsymbol{f}'(\boldsymbol{x}), \boldsymbol{s}) \leq -\boldsymbol{\sigma} < \mathbf{0}$, то $\exists$ достаточно малое $\boldsymbol{\beta} > \mathbf{0}$:

$$(\boldsymbol{f}'(\boldsymbol{x} + \boldsymbol{\beta s}), \boldsymbol{s}) < \mathbf{0}, \boldsymbol{x} + \boldsymbol{\beta s} \in \boldsymbol{Q}.$$

Тогда по теореме о среднем

$$\boldsymbol{f}(\boldsymbol{x} + \boldsymbol{\beta s}) - \boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{\beta}(\boldsymbol{f}'(\boldsymbol{x} + \boldsymbol{\theta \beta s}), \boldsymbol{s}) < \mathbf{0}, \text{ где } \mathbf{0} \leq \boldsymbol{\theta} \leq \mathbf{1}.$$

Противоречие с тем, что $\boldsymbol{x}$ – локальный минимум функции $\boldsymbol{f}$.$\blacksquare$

Комментарий

Теорема 2 эквивалентна утверждению, что для локальной оптимальности точки $\boldsymbol{x}$ необходимо выполнение условия

$$\boldsymbol{K_d(x) \cap K_<(x) = \emptyset.}$$

Название теоремы связано с тем, что в ней необходимые условия формулируются в виде требования пустоты пересечения двух геометрических объектов – двух конусов. Эту теорему можно сформулировать в более сильном варианте:

$$\boldsymbol{K_d(x) \cap \overline{K}_f(x) = \emptyset,}$$

который, однако, неконструктивен, т.к. отсутствует аналитическое описание замыкания конуса возможных направлений. Впоследствии мы увидим, что наличие дополнительной информации о задаче приводит к нужному описанию этого конуса. Эта информация будет формулироваться в виде так называемых условий регулярности.

Необходимые условия оптимальности Фритца–Джона

Теорема 3 (Необходимые условия оптимальности Фритца–Джона). Пусть x^* — локальный экстремум задачи (1), (2), функции f, $\varphi_i, i = \overline{1,m}$, непрерывны и непрерывно дифференцируемы. Тогда найдутся такие не все равные 0 множители $\lambda_i \geq 0, i = \overline{0,m}$, что

$$\lambda_0 f'(x^*) + \sum_{i=1}^{m} \lambda_i \varphi_i'(x^*) = 0^n, \tag{7}$$

$$\lambda_i \varphi_i(x^*) = 0, i = \overline{1,m}. \tag{8}$$

Доказательство. Перепишем неравенство (6) ($\sigma \leq 0$) теоремы 2 в следующем виде: $(0, s) + 1 \cdot \sigma \leq 0$. Эта теорема утверждает, что

Необходимые условия оптимальности Фритца–Джона

если вектор $(\boldsymbol{s}, \boldsymbol{\sigma})$ удовлетворяет системе уравнений (4), (5):

$$(s_1, \ldots, s_n, \sigma)\begin{pmatrix} f'(x^*) & \ldots & \varphi_i'(x^*) & \ldots \\ 1 & \ldots & 1 & \ldots \end{pmatrix} \leq 0,$$

то выполняется неравенство (6)

$$(s_1, \ldots, s_n, \sigma)(0, \ldots, 0, 1)^{\mathsf{T}} \leq 0.$$

Тогда по теореме Фаркаша-Минковского $\exists$ неотрицательное решение системы уравнений:

$$\begin{pmatrix} f'(x^*) & \ldots & \varphi_i'(x^*) & \ldots \\ 1 & \ldots & 1 & \ldots \end{pmatrix}(\lambda_0, \ldots, \lambda_i, \ldots)^{\mathsf{T}} = (0, \ldots, 0, 1)^{\mathsf{T}}.$$

Необходимые условия оптимальности Фритца–Джона

Следовательно

$$\boldsymbol{\lambda_0 f'(x^*) + \sum_{i \in I(x^*)} \lambda_i \varphi_i'(x^*) = 0^n,}$$

$$\boldsymbol{\lambda_0 + \sum_{i \in I(x^*)} \lambda_i = 1.} \tag{9}$$

Положим $\boldsymbol{\lambda_i = 0, i \notin I(x^*)}$ и получим

$$\boldsymbol{\lambda_0 f'(x^*) + \sum_{i=1}^{m} \lambda_i \varphi_i'(x^*) = 0^n,}$$

$$\boldsymbol{\lambda_i \varphi_i(x^*) = 0, i = \overline{1, m}.} \ \blacksquare$$

Комментарий

Необходимые условия Фритца–Джона дают уже чисто алгебраическую характеризацию локальных экстремумов как решений системы уравнений. Компоненты решения $\boldsymbol{\lambda_i \geq 0, i = \overline{0, m}}$, системы (7)-(8) называют множителями Лагранжа локального оптимума $\boldsymbol{x}^*$. Соотношения (8) называют соотношениями дополняющей нежёсткости, в следствии которых множители $\boldsymbol{\lambda_i}$ соответствующие неактивным ограничениям $\boldsymbol{i \notin I(x^*)}$ равны 0.

Множитель $\boldsymbol{\lambda_i}$ характеризует чувствительность данного локального оптимума $\boldsymbol{x}^*$ относительно малых изменений величины $\boldsymbol{\varphi_i(x^*)}$.

Перейдём к выводу классических условий оптимальности Куна-Таккера, которые совпадают с необходимыми условиями Фритца-Джона затем лишь исключением, что в них $\boldsymbol{\lambda_0 \neq 0.}$

Необходимые условия оптимальности Куна–Таккера

Теорема 4 (Необходимые условия оптимальности Куна–Таккера). Пусть x^* — локальный экстремум задачи (1), (2), функции f, $\varphi_i, i = \overline{1,m}$, непрерывны и непрерывно дифференцируемы и вектора $\varphi_i'(x^*), i \in I(x^*)$, линейно независимы. Тогда найдутся такие множители $\lambda_i \geq 0, i = \overline{1,m}$, что

$$-f'(x^*) = \sum_{i=1}^{m} \lambda_i \varphi_i'(x^*), \tag{10}$$

$$\lambda_i \varphi_i(x^*) = 0, i = \overline{1,m}. \tag{11}$$

Доказательство. Из теоремы 3 $\Longrightarrow$ найдутся такие не все равные 0 множители $\lambda_i \geq 0, i = \overline{0,m}$, что

Необходимые условия оптимальности Куна–Таккера

$$\boldsymbol{\lambda_0 f'(x^*) + \sum_{i=1}^{m} \lambda_i \varphi_i'(x^*) = 0^n.} \tag{7}$$

$$\boldsymbol{\lambda_i \varphi_i(x^*) = 0, i = \overline{1, m}.} \tag{8}$$

Также получили равенство

$$\boldsymbol{\lambda_0 + \sum_{i \in I(x^*)} \lambda_i = 1.} \tag{9}$$

Покажем, что $\boldsymbol{\lambda_0 \neq 0}$ (от противного). Пусть $\boldsymbol{\lambda_0 = 0}$. Из (9) $\Longrightarrow \exists\ \boldsymbol{\lambda_i \neq 0}$ $\Longrightarrow \boldsymbol{\{\varphi_i'(x^*)\}_{i \in I(x^*)}}$ линейно зависимые вектора. Противоречие. $\Longrightarrow \boldsymbol{\lambda_0 > 0}$

Необходимые условия оптимальности Куна–Таккера: выпуклый случай

Задача (1), (2) называется задачей выпуклого программирования, если функции $\boldsymbol{f}$, $\boldsymbol{\varphi_i, i = \overline{1, m}}$, — выпуклы.

Множество $\boldsymbol{Q}$ выпукло. По прежнему считаем, что $\boldsymbol{f}$, $\boldsymbol{\varphi_i} \in \boldsymbol{C^1}$.

Условие регулярности для выпуклого случая:

$$\forall \boldsymbol{i, i = \overline{1, m},}\ \exists \boldsymbol{x^i} \in \boldsymbol{Q} : \varphi_i(\boldsymbol{x^i}) < \boldsymbol{0.}$$

Эквивалентно условию регулярности Слейтера

$$\exists \widetilde{\boldsymbol{x}} \in \boldsymbol{Q} : \varphi_i(\widetilde{\boldsymbol{x}}) < \boldsymbol{0, i = \overline{1, m}.}$$

Необходимые условия оптимальности Куна–Таккера: выпуклый случай

Лемма 3. Функция f дифференцируемая на выпуклом множестве Q, выпукла в том и только в том случае, когда для любых $x, y \in Q$: $(f'(x), y - x) \leq f(y) - f(x)$.

Доказательство. $\forall x \neq y \in Q, \forall \alpha \; 0 < \alpha \leq 1$

$$f(x + \alpha(y - x)) \leq f(x) + \alpha(f(y) - f(x)).$$

Перепишем неравенство:

$$\|y - x\| \frac{f(x + \beta s) - f(x)}{\beta} \leq f(y) - f(x),$$

где $s = \frac{y-x}{\|y-x\|}$, $\beta = \alpha\|y - x\|$. Устремим β к 0. В пределе получим

Необходимые условия оптимальности Куна–Таккера: выпуклый случай

$$\frac{\partial f}{\partial s}\,\|y-x\| \le f(y)-f(x).$$

Но

$$\frac{\partial f}{\partial s}\,\|y-x\| = (f'(x), s)\,\|y-x\| = (f'(x), y-x).$$

Итак, доказали неравенство

$$(f'(x), y-x) \le f(y)-f(x).$$

В обратную сторону. Пусть $\forall x, y \in Q : (f'(x), y-x) \le f(y)-f(x)$.

Необходимые условия оптимальности Куна–Таккера: выпуклый случай

По условию $\forall\, \alpha \in [0,1] : z = \alpha x + (1-\alpha) y \in Q$.

Умножим неравенство $(f'(z), x - z) \le f(x) - f(z)$ на α, а

неравенство $(f'(z), y - z) \le f(y) - f(z)$ на $(1-\alpha)$ и сложим их

$$0 = (f'(z), \alpha(x - z) + (1-\alpha)(y - z)) \le \alpha f(x) + (1-\alpha) f(y) - f(z)$$

$$\Longrightarrow f(\alpha x + (1-\alpha) y) \le \alpha f(x) + (1-\alpha) f(y).$$

Необходимые условия оптимальности Куна–Таккера: выпуклый случай

Лемма 4. Если

$$Q = \{x | \varphi_i(x) = (a_i, x) - b_i \le 0, i = \overline{1, m}\},$$

то условия

$$(a_i, s) \le 0, i \in I(x^*)$$

необходимы и достаточны для того, чтобы направление s было возможным в точке $x^* \in Q$.

Доказательство. Пусть $\beta > 0$. Рассмотрим

$$\varphi_i(x^* + \beta s) = (a_i, x^* + \beta s) - b_i = (a_i, x^*) - b_i + \beta(a_i, s).$$

$$\forall i \notin I(x^*)\ (a_i, x^*) - b_i < 0 \Rightarrow \forall i \notin I(x^*)\ \varphi_i(x^* + \beta s) \le 0,$$

для достаточно малых β.

Необходимые условия оптимальности Куна–Таккера: выпуклый случай

$$\forall \boldsymbol{i} \in \boldsymbol{I}(\boldsymbol{x}^*)\ \varphi_i(\boldsymbol{x}^* + \boldsymbol{\beta s}) = (\boldsymbol{a}_i, \boldsymbol{x}^* + \boldsymbol{\beta s}) - \boldsymbol{b}_i = \boldsymbol{\beta}(\boldsymbol{a}_i, \boldsymbol{s}) \Rightarrow$$

$$\Rightarrow \boldsymbol{x}^* + \boldsymbol{\beta s} \in \boldsymbol{Q}\ \forall \boldsymbol{\beta} > \mathbf{0} \Leftrightarrow (\boldsymbol{a}_i, \boldsymbol{s}) \leq \mathbf{0}\ \forall \boldsymbol{i} \in \boldsymbol{I}(\boldsymbol{x}^*). \blacksquare$$

Эта лемма позволяет элиминировать условие Слейтера в задаче выпуклого программирования в случае линейных ограничений.

Помним: функции $\boldsymbol{f}, \varphi_i$ — выпуклые, непрерывно-дифференцируемые. Множество допустимых решений $\boldsymbol{Q}$ удовлетворяет условию Слейтера.

Необходимые условия оптимальности Куна–Таккера: выпуклый случай

Теорема 5 (Необходимые условия оптимальности Куна–Таккера: выпуклый случай). Если точка x^* — локальный минимум задачи (1), (2),то найдутся такие множители $\lambda_i \geq 0, i = \overline{1, m}$, что

$$-f'(x^*) = \sum_{i=1}^{m} \lambda_i \varphi_i'(x^*), \tag{10}$$

$$\lambda_i \varphi_i(x^*) = 0, i = \overline{1, m}. \tag{11}$$

Доказательство. Из теоремы 3 $\Longrightarrow$ найдутся такие не все равные 0 множители $\lambda_i \geq 0, i = \overline{0, m}$, что

Необходимые условия оптимальности Куна–Таккера: выпуклый случай

$$\lambda_0 f'(x^*) + \sum_{i=1}^{m} \lambda_i \varphi_i'(x^*) = 0^n. \tag{7}$$

$$\lambda_i \varphi_i(x^*) = 0, i = \overline{1, m}. \tag{8}$$

Также получили равенство

$$\lambda_0 + \sum_{i \in I(x^*)} \lambda_i = 1. \tag{9}$$

Покажем, что $\lambda_0 > 0$ (от противного). Пусть $\lambda_0 = 0$.

Из (9) $\Longrightarrow \sum\limits_{i \in I(x^*)} \lambda_i = 1$.

Из условия Слейтера $\Longrightarrow \exists\, z \in Q : \varphi_i(z) < 0, i = \overline{1, m}$.

Необходимые условия оптимальности Куна–Таккера: выпуклый случай

Из леммы 3 $\Longrightarrow$

$$(\varphi_i'(x^*), z - x^*) \le \varphi_i(z) - \varphi_i(x^*) = \varphi_i(z) < 0, \forall i \in I(x^*).$$

Умножим (7) скалярно на вектор $s = z - x^* \Longrightarrow$

$$\sum_{i \in I(x^*)} \lambda_i (\varphi_i'(x^*), s) = 0. \qquad (*)$$

Но $\sum\limits_{i \in I(x^*)} \lambda_i = 1, \lambda_i \ge 0, i \in I(x^*)$, $(\varphi_i'(x^*), s) < 0, i \in I(x^*)$
$\Longrightarrow \sum\limits_{i \in I(x^*)} \lambda_i (\varphi_i'(x^*), s) < 0$. Противоречие с $(*) \Longrightarrow \lambda_0 > 0$. ■

Далее рассмотрим задачу с линейными ограничениями, т.е. $Q = \{x | \varphi_i(x) = (a_i, x) - b_i \le 0, i = \overline{1, m}\}$, условие Слейтера не используется, $f \in C^1$ — выпуклая

Правило множителей Лагранжа

Теорема 7 (правило множителей Лагранжа). Пусть x^* — локальный экстремум задачи (1), (2), функции f, $\varphi_i, i = \overline{1,m}$, непрерывны и непрерывно дифференцируемы и вектора $\varphi_i'(x^*), i = \overline{1,m}$, линейно независимы. Тогда найдутся такие множители $\lambda_i, i = \overline{1,m}$, что

$$f'(x^*) = \sum_{i=1}^{m} \lambda_i \varphi_i'.$$