ЛЕКЦИЯ № 3

Необходимые условия экстремума и критерии оптимальности

1. Необходимые условия оптимальности Куна–Таккера: линейный случай

2. Теорема Куна-Таккера в локальной форме

3. Теорема Куна-Таккера в нелокальной форме

4. Теория двойственности. Линейное программирование

Необходимые условия оптимальности Куна–Таккера: линейный случай

$Q = \{x | \varphi_i(x) = (a_i, x) - b_i \leq 0, i = \overline{1, m}\}$, условие Слейтора не используется, $f \in C^1$ — выпуклая

Теорема 6 (Необходимые условия оптимальности Куна–Таккера: линейный случай). **Если точка x^* — локальный минимум задачи (1), (2),то найдутся такие множители $\lambda_i \geq 0, i = \overline{1, m}$, что**

$$-f'(x^*) = \sum_{i=1}^{m} \lambda_i a_i, \tag{12}$$

$$\lambda_i \varphi_i(x^*) = 0, i = \overline{1, m}. \tag{13}$$

Доказательство. Найдется $\varepsilon > 0$:

$$f(x^*) < f(y) \; \forall y \in B(x^*, \varepsilon) \cap Q.$$

Необходимые условия оптимальности Куна–Таккера: линейный случай

Пусть $z \in Q, z \neq x^*$. Тогда

1). Для малых $\alpha > 0$

$$\alpha z + (1-\alpha)x^* = x^* + \alpha(z - x^*) \in B(x^*, \varepsilon) \cap Q \Rightarrow$$

$$(f'(x^*), z - x^*) = \lim_{\alpha \to 0} \frac{f(x^* + \alpha(z - x^*)) - f(x^*)}{\alpha} \geq 0. \qquad (*)$$

2). $(*)$ выполняется для любого возможного направления s (т.к. для подходящего $z \in Q: \; s = z - x^*$).

3). Из 2) и леммы 4 $\Rightarrow$

$$\forall s \neq 0: \; (a_i, s) \leq 0, \forall i \in I(x^*) \Rightarrow (*).$$

Необходимые условия оптимальности Куна–Таккера: линейный случай

Т.е. для любого решения $\boldsymbol{s} \neq \boldsymbol{0}$ системы

$$(\boldsymbol{a}_i, \boldsymbol{s}) \leq \boldsymbol{0}, \forall \boldsymbol{i} \in \boldsymbol{I}(\boldsymbol{x}^*)$$

выполняется неравенство

$$(-\boldsymbol{f}'(\boldsymbol{x}^*), \boldsymbol{s}) \leq \boldsymbol{0}.$$

Применяем теорему Фаркаша-Минковского.■

Упражнение Провести соответствующие выкладки самостоятельно.

Вернемся к задаче выпуклого программирования. Выполняется условие Слейтера!

Критерии оптимальности

Теорема 7 (Теорема Куна-Таккера в локальной форме). Точка $x^* \in Q$ — оптимальное решение задачи выпуклого программирования в том и только в том случае, когда существуют такие числа $\lambda_i \geq 0, i = \overline{1, m}$, что

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

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

Доказательство. Необходимость следует из теоремы 5. Достаточность. Множество допустимых решений Q — выпукло $\Longrightarrow$

$$\forall y \in Q \text{ вектор } s = y - x^* - \text{ возможное направление в точке } x^* \Longrightarrow$$

$$(\varphi_i'(x^*), s) \leq 0, \ \forall i \in I(x^*).$$

Но f — выпуклая функция и из леммы 3 имеем

Критерии оптимальности

$$\forall y \in Q: \; f(y) - f(x^*) \geq (f_i'(x^*), y - x^*) = (f_i'(x^*), s) =$$

$$(-\sum_{i=1}^{m} \lambda_i \varphi_i'(x^*), s) = \sum_{i=1}^{m} (-\lambda_i)(\varphi_i'(x^*), s) \geq 0. \blacksquare$$

Если все ограничения линейны, то условие Слейтера в предыдущей теореме можно опустить.

Теорема 8. Точка $x^* \in Q$ — оптимальное решение задачи выпуклого программирования с линейными ограничениями в том и только в том случае, когда существуют такие числа $\lambda_i \geq 0, i = \overline{1, m}$, что

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

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

Доказательство. Необходимость следует из теоремы 6. Достаточность как в теореме 7.

Критерии оптимальности

Определение 2. Функцию

$$\boldsymbol{L(x,\lambda) = f(x) + \sum_{i=1}^{m} \lambda_i \varphi_i(x),}$$

определенную при всех $\boldsymbol{x}$ и $\boldsymbol{\lambda}$, назовем функцией Лагранжа для задачи (1), (2).

Определение 3. Пара $(\boldsymbol{x^*}, \boldsymbol{\lambda^*})$ называется седловой точкой функции Лагранжа, если

$$\boldsymbol{L(x^*,\lambda) \overset{(14)}{\leq} L(x^*,\lambda^*) \overset{(15)}{\leq} L(x,\lambda^*) \ \forall x \in R^n, \forall \lambda \geq 0.}$$

Пусть для задачи выпуклого программирования выполняется условие Слейтера. Тогда верно следующее утверждение.

Критерии оптимальности

Теорема 9 (Теорема Куна-Таккера в нелокальной форме). Допустимое решение $x^* \in Q$ является оптимальным тогда и только тогда, когда существует такой вектор λ^*, что пара (x^*, λ^*) является седловой точкой функции Лагранжа.

Доказательство. Необходимость. Пусть $x^* \in Q$ — оптимальное решение. Тогда из теоремы 5 имеем

$$\exists \lambda^* \geq 0 : f'(x^*) + \sum_{i=1}^{m} \lambda_i^* \varphi_i'(x^*) = 0 \text{ и}$$

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

Фиксируем λ. Функция

$$x \longrightarrow L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i \varphi_i(x)$$

выпукла $\overset{\text{л. 3}}{\Longrightarrow}$

Критерии оптимальности

$$L(x, \lambda^*) \geq L(x^*, \lambda^*) + (\frac{\partial L(x, \lambda)}{\partial x}\Big|_{(x^*, \lambda^*)}, x - x^*) = L(x^*, \lambda^*)$$

$$(\text{так как } \frac{\partial L(x, \lambda)}{\partial x}\Big|_{(x^*, \lambda^*)} = f^{'}(x^*) + \sum_{i=1}^{m} \lambda_i^* \varphi_i^{'}(x^*) = 0).$$

Получили неравенство (15) определения седловой точки.

Фиксируем x. Функция

$$\lambda \longrightarrow L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i \varphi_i(x)$$

линейна $\overset{\text{л. } 3}{\Longrightarrow}$ и равенства $\frac{\partial L(x,\lambda)}{\partial x}\Big|_{(x^*,\lambda^*)} = \varphi(x^*)$ имеем

$$L(x^*, \lambda^*) \geq L(x^*, \lambda) + (\varphi(x^*), \lambda^* - \lambda) =$$

Критерии оптимальности

$$\boldsymbol{L(x^*, \lambda) + (\varphi(x^*), \lambda^*) - (\varphi(x^*), \lambda) \geq L(x^*, \lambda).}$$

Последнее неравенство следует из дополняющей нежесткости $(\boldsymbol{\varphi(x^*), \lambda^*}) = \mathbf{0}$ и неравенств $\boldsymbol{\varphi(x^*) \leq 0, \lambda \geq 0}$.

Таким образом получили неравенство (14) из определения 3 и, следовательно, $(\boldsymbol{x^*, \lambda^*})$ — седловая точка функции Лагранжа.

Достаточность. Пусть $(\boldsymbol{x^*, \lambda^*})$ — седловая точка функции $\boldsymbol{L}$. Из неравенства

$$\boldsymbol{L(x^*, \lambda^*) \overset{(15)}{\leq} L(x, \lambda^*) \ \forall x \in R^n \Rightarrow}$$

$$\boldsymbol{L(x^*, \lambda^*) = \min_{x \in R^n} L(x, \lambda^*).} \qquad (*)$$

Из неравенства

$$\boldsymbol{L(x^*, \lambda) \overset{(14)}{\leq} L(x^*, \lambda^*) \ \forall \lambda \geq 0 \Rightarrow}$$

Критерии оптимальности

$$\forall \boldsymbol{\lambda} \geq \mathbf{0} \ \sum_{i=1}^{m} \boldsymbol{\lambda}_i \varphi_i(\boldsymbol{x}^*) \leq \sum_{i=1}^{m} \boldsymbol{\lambda}_i^* \varphi_i(\boldsymbol{x}^*) < \infty. \qquad (**)$$

Из $(**) \Rightarrow \boldsymbol{x}^*$ — допустимое решение $\Rightarrow \sum\limits_{i=1}^{m} \boldsymbol{\lambda}_i^* \varphi_i(\boldsymbol{x}^*) \leq \mathbf{0}$.

$\Big($От противного. Пусть $\exists \boldsymbol{i_0} : \varphi_{i_0}(\boldsymbol{x}^*) > \mathbf{0}$. Тогда $\sum\limits_{i=1}^{m} \boldsymbol{\lambda}_i \varphi_i(\boldsymbol{x}^*) \longrightarrow +\infty$ при $\boldsymbol{\lambda}_{i_0} \longrightarrow +\infty$. Противоречие с $(**)$. $\Big)$

Но $\sum\limits_{i=1}^{m} \boldsymbol{\lambda}_i^* \varphi_i(\boldsymbol{x}^*) \geq \mathbf{0}$ (положить в $(**)$ $\boldsymbol{\lambda}_i = \mathbf{0} \ \forall \boldsymbol{i}$).

Т.к. равенство $\sum\limits_{i=1}^{m} \boldsymbol{\lambda}_i^* \varphi_i(\boldsymbol{x}^*) = \mathbf{0}$ эквивалентно условиям дополняющей нежесткости $\boldsymbol{\lambda}_i^* \varphi_i(\boldsymbol{x}^*) = \mathbf{0} \ \forall \boldsymbol{i}$, то отсюда и $(*)$ получим

$$\forall \boldsymbol{x} \in \boldsymbol{Q} : \ \boldsymbol{f}(\boldsymbol{x}^*) \leq \boldsymbol{f}(\boldsymbol{x}) + \sum_{i=1}^{m} \boldsymbol{\lambda}_i^* \varphi_i(\boldsymbol{x}) \leq \boldsymbol{f}(\boldsymbol{x}) \Rightarrow \boldsymbol{x}^* \text{ — опт. реш.} \blacksquare$$

Теория двойственности

Рассмотрим задачу $\boldsymbol{P}$:

$$f(x) \longrightarrow \min$$

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

Функции f и φ_i произвольны. Пусть

$$g(x) = \sup_{\lambda \ge 0} L(x, \lambda), \text{ где}$$

$$L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i \varphi_i(x) - \text{функция Лагранжа.}$$

Тогда

$$g(x) = \begin{cases} f(x), \ x \in Q, \\ +\infty, \ x \notin Q. \end{cases} \Longrightarrow$$

Теория двойственности

Задача $\boldsymbol{P}$ эквивалентна следующей:

$$\boldsymbol{g(x) \longrightarrow \min}$$

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

Пусть

$$\boldsymbol{h(\lambda) = \inf_{x \in R^n} L(x, \lambda).}$$

Рассмотрим задачу $(\boldsymbol{D})$:

$$\boldsymbol{h(\lambda) \longrightarrow \max_{\lambda \geq 0}.}$$

Назовем ее двойственной задачей к прямой (или исходной) задаче $\boldsymbol{P}$. Переменные $\boldsymbol{\lambda_1, \ldots, \lambda_m}$ назовем двойственными, а переменные $\boldsymbol{x_1, \ldots, x_n}$ — прямыми.

Если $\boldsymbol{x \in Q, \lambda \geq 0}$, то $\boldsymbol{x}$ — допустимое решение прямой задачи, а $\boldsymbol{\lambda}$ — допустимое решение двойственной задачи.

Теория двойственности

Лемма 5 (Слабая теорема двойственности).

$$\forall x \in Q\ \forall \lambda \geq 0\ (h(\lambda) \leq f(x)).$$

Доказательство. Т.к. $\lambda_i \varphi_i(x) \leq 0, \forall i$ и $\forall \lambda \geq 0$, то

$$h(\lambda) = \inf_{x \in R^n} L(x, \lambda) \leq L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i \varphi_i(x) \leq f(x).$$

Лемма 6. Если $\overline{x} \in Q$ и $\overline{\lambda} \geq 0$ и $f(\overline{x}) = h(\overline{\lambda})$, то $\overline{x}$ и $\overline{\lambda}$ — оптимальные решения задачи P и D, соответственно.

Доказательство. Из леммы 5 имеем

$$f(\overline{x}) = h(\overline{\lambda}) \leq f(x)\ \forall x \in Q \Rightarrow \overline{x} \text{ — оптимальное решение } P,$$

$$h(\lambda) \leq f(\overline{x}) = h(\overline{\lambda})\ \forall \lambda \geq 0 \Rightarrow \overline{\lambda} \text{ — оптимальное решение } D.$$

Теория двойственности

Теорема 10. Если $\overline{x}, \overline{\lambda}$ — допустимые решения прямой и двойственной задачи, то $f(\overline{x}) = h(\overline{\lambda})$ тогда и только тогда, когда пара $(\overline{x}, \overline{\lambda})$ — седловая точка функции Лагранжа, причем $L(\overline{x}, \overline{\lambda}) = f(\overline{x}) = h(\overline{\lambda})$.

Доказательство. Необходимость. (14):

$$L(\overline{x}, \lambda) \le f(\overline{x}) = h(\overline{\lambda}) \le L(\overline{x}, \overline{\lambda}).$$

Первое неравенство: $f(\overline{x}) = \sup_{\lambda \ge 0} L(\overline{x}, \lambda)$.

Второе: $h(\lambda) = \inf\limits_{x \in R^n} L(x, \lambda)$.

Аналогично (15):

$$L(\overline{x}, \overline{\lambda}) \le f(\overline{x}) = h(\overline{\lambda}) \le L(x, \overline{\lambda}) \Longrightarrow$$

$(\overline{x}, \overline{\lambda})$ — седловая точка функции L и $L(\overline{x}, \overline{\lambda}) = f(\overline{x}) = h(\overline{\lambda})$.

Теория двойственности

Достаточность. Пусть $(\overline{x}, \overline{\lambda})$ — седловая точка функции L и $L(\overline{x}, \overline{\lambda}) = f(\overline{x}) = h(\overline{\lambda})$.

(14)$\Longrightarrow \forall \lambda \geq 0 : L(\overline{x}, \lambda) \leq L(\overline{x}, \overline{\lambda}) \Longrightarrow$

$$\sup_{\lambda \geq 0} L(\overline{x}, \lambda) \overset{(*)}{=} L(\overline{x}, \overline{\lambda}) \overset{(**)}{=} f(\overline{x}).$$

Равенство $(*)$ — тривиально, а $(**)$ следует из условия $\overline{x} \in Q$ (если $\varphi_i(\overline{x}) < 0 \Rightarrow \overline{\lambda}_i = 0$).

(15)$\Longrightarrow \forall x \in R^n : L(\overline{x}, \overline{\lambda}) \leq L(x, \overline{\lambda}) \Longrightarrow$

$$L(\overline{x}, \overline{\lambda}) = \inf_{x} L(x, \overline{\lambda}) = h(\overline{\lambda}).$$

$\overline{x}, \overline{\lambda}$ — допустимые решения $\overset{\text{л. 5}}{\Longrightarrow}$

$$L(\overline{x}, \overline{\lambda}) = h(\overline{\lambda}) \leq f(\overline{x}) = L(\overline{x}, \overline{\lambda}). \blacksquare$$