УСЛОВИЯ ОПТИМАЛЬНОСТИ ВТОРОГО ПОРЯДКА В НЕЛИНЕЙНОМ ПРОГРАММИРОВАНИИ*

В. Н. Малозёмов

malv@math.spbu.ru

16 октября 2010 г.

На современном уровне излагается теория условий оптимальности второго порядка в нелинейном программировании (ср. с [1], с. 42–52).

Данный доклад примыкает к докладу [2], посвящённому условиям оптимальности первого порядка.

1°. Рассмотрим задачу нелинейного программирования

$$\begin{gathered} f(x) \to \inf, \\ a_i(x) \geqslant 0, \quad i \in M_1; \\ a_i(x) = 0, \quad i \in M_2; \\ x \in U, \end{gathered} \tag{1}$$

где $U \subset \mathbb{R}^N$ — открытое множество и f, a_i при $i \in M_1 \cup M_2$ — дважды непрерывно дифференцируемые на U функции. Множество планов задачи (1) обозначим Ω.

Зафиксируем $x_* \in \Omega$ и введём индексные множества

$$\begin{gathered} M_1(x_*) = \big\{ i \in M_1 \mid \ a_i(x_*) = 0 \big\}, \\ I(x_*) = M_1(x_*) \cup M_2. \end{gathered}$$

Говорят, что в точке x_* ограничения задачи (1) *регулярны*, если градиенты $a_i'(x_*)$ при $i \in I(x_*)$ линейно независимы.

---

*Семинар по дискретному гармоническому анализу и геометрическому моделированию «DHA & CAGD»: `http://www.dha.spb.ru/`

Обозначим $M = M_1 \cup M_2$. Говорят, что в точке x_* выполняются *условия Куна–Таккера*, если существует вектор $u_* = u_*[M]$ со свойствами

$$f'(x_*) = \sum_{i \in M} u_*[i]\, a_i'(x_*); \tag{2}$$

$$u_*[i]\, a_i(x_*) = 0, \qquad i \in M_1; \tag{3}$$

$$u_*[i] \geqslant 0, \qquad i \in M_1. \tag{4}$$

С помощью функции Лагранжа

$$\mathcal{L}(x, u) = f(x) - \sum_{i \in M} u[i]\, a_i(x)$$

условие (2) можно переписать в виде

$$\mathcal{L}_x'(x_*, u_*) = \mathbb{O}. \tag{5}$$

С двойственным вектором u_* свяжем два индексных множества (см. рис. 1)

$$M_1^+(x_*) = \{i \in M_1(x_*) \mid u_*[i] > 0\},$$
$$I^+(x_*) = M_1^+(x_*) \cup M_2.$$

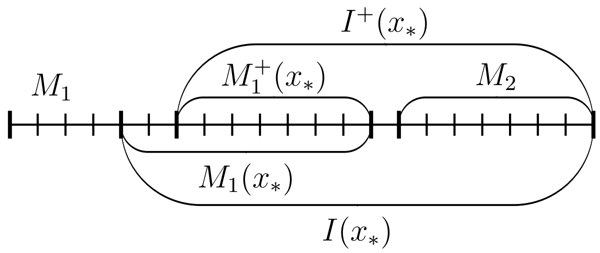

Рис. 1

Отметим, что в силу условия дополнительности (3) и условия неотрицательности (4)

$$u_*[i] = 0, \qquad i \in M \setminus I^+(x_*). \tag{6}$$

Рассмотрим систему линейных соотношений

$$\begin{aligned} \langle a_i'(x_*), g \rangle &= 0, \qquad i \in I^+(x_*); \\ \langle a_i'(x_*), g \rangle &\geqslant 0, \qquad i \in I(x_*) \setminus I^+(x_*). \end{aligned} \tag{7}$$

Множество векторов g, удовлетворяющих (7), является конусом в $\mathbb{R}^N$. Обозначим его G_*.

ТЕОРЕМА 1 (необходимое условие оптимальности второго порядка). *Пусть $x_* \in \Omega$ — точка локального минимума в задаче* (1) *и ограничения в ней регулярны. Тогда в x_* выполняются условия Куна–Таккера, то есть найдётся вектор u_* со свойствами* (2)–(4). *Более того, для всех $g \in G_*$ выполняется неравенство*

$$\left\langle \mathcal{L}''_{xx}(x_*, u_*)g, g \right\rangle \geqslant 0. \tag{8}$$

Доказательство. Выполнение условий Куна–Таккера считаем известным фактом [2]. Проверим выполнение условий второго порядка (8).

Зафиксируем ненулевой вектор $g \in G_*$ и введём индексное множество

$$I_g(x_*) = \{i \in I(x_*) \mid \left\langle a'_i(x_*), g \right\rangle = 0\}.$$

Согласно определениям

$$M_2 \subset I^+(x_*) \subset I_g(x_*) \subset I(x_*) \tag{9}$$

и

$$\left\langle a'_i(x_*), g \right\rangle > 0, \quad i \in I(x_*) \setminus I_g(x_*). \tag{10}$$

Рассмотрим систему нелинейных уравнений

$$a_i(x) = 0, \quad i \in I_g(x_*). \tag{11}$$

Согласно (9) точка x_* удовлетворяет (11), градиенты $a'_i(x_*)$ при $i \in I_g(x_*)$ линейно независимы и ненулевой вектор g ортогонален $a'_i(x_*)$ при всех $i \in I_g(x_*)$. По основной лемме нелинейного программирования [2] существует параметрическая кривая $x = x(t)$, непрерывно дифференцируемая в окрестности точки $t = 0$, такая, что

$$x(0) = x_*, \quad x'(0) = g, \tag{12}$$

$$a_i\big(x(t)\big) = 0 \ \text{ при } \ i \in I_g(x_*) \ \text{ и малых } \ t. \tag{13}$$

Покажем, что $x(t) \in \Omega$ при малых $t > 0$.

При $i \in I_g(x_*)$ выполняется равенство (13). В частности, оно выполняется при $i \in M_2$. При $i \in I(x_*) \setminus I_g(x_*)$ согласно (12) имеем

$$a_i\big(x(t)\big) = a_i\big(x(0)\big) + \Big\langle a'_i\big(x(0)\big), x'(0) \Big\rangle\, t + o(t) =$$
$$= a_i(x_*) + \left\langle a'_i(x_*), g \right\rangle t + o(t) = t\left[\left\langle a'_i(x_*), g \right\rangle + \frac{o(t)}{t}\right].$$

На основании (10) заключаем, что $a_i\big(x(t)\big) > 0$ при малых $t > 0$. Остаётся рассмотреть индексы $i \in M \setminus I(x_*)$, на которых $a_i(x_*) > 0$. Очевидно, что и на этих индексах будет выполняться неравенство $a_i\big(x(t)\big) > 0$ при малых t. Таким образом, действительно $x(t)$ является планом задачи (1) при малых $t > 0$.

Обратимся к функции Лагранжа. В силу (6), (13) и включения $I^+(x_*) \subset \subset I_g(x_*)$ имеем

$$\mathcal{L}\big(x(t), u_*\big) = f\big(x(t)\big) - \sum_{i \in I^+(x_*)} u_*[i]\, a_i\big(x(t)\big) = f\big(x(t)\big).$$

К этому нужно добавить, что $\mathcal{L}(x_*, u_*) = f(x_*)$.

Воспользуемся тем, что x_* — точка локального минимума. При малых $t > 0$ получим

$$0 \leqslant f\big(x(t)\big) - f(x_*) = \mathcal{L}\big(x(t), u_*\big) - \mathcal{L}(x_*, u_*) =$$
$$= \big\langle \mathcal{L}'_x(x_*, u_*), x(t) - x_* \big\rangle + \frac{1}{2}\Big\langle \mathcal{L}''_{xx}\big(\xi(t), u_*\big)\big(x(t) - x_*\big), x(t) - x \Big\rangle,$$

где $\xi(t) \in [x_*, x(t)]$. Отметим, что точка $\xi(t)$ может не принадлежать Ω. Нам важно, что $\xi(t) \to x_*$ при $t \to +0$. С учётом (5) перепишем последнее неравенство в эквивалентном виде

$$\Big\langle \mathcal{L}''_{xx}\big(\xi(t), u_*\big) \frac{x(t) - x(0)}{t}, \frac{x(t) - x(0)}{t} \Big\rangle \geqslant 0. \tag{14}$$

Согласно (12)

$$\lim_{t \to +0} \frac{x(t) - x(0)}{t} = g.$$

Кроме того, $\mathcal{L}''_{xx}\big(\xi(t), u_*\big) \to \mathcal{L}''_{xx}(x_*, u_*)$ при $t \to +0$ в силу дважды непрерывной дифференцируемости функции Лагранжа по x. Переходя к пределу в (14) при $t \to +0$, получаем (8).

Теорема доказана. □

2°. Обратимся к достаточным условиям строгого локального минимума. В [2] установлены достаточные условия первого порядка:

Если в точке $x_* \in \Omega$ выполнены условия Куна–Таккера и $G_* = \{\mathbb{O}\}$, то x_* — точка строгого локального минимума в задаче (1).

Рассмотрим случай, когда G_* состоит не только из нулевого вектора.

ТЕОРЕМА 2 (достаточные условия оптимальности второго порядка). *Пусть в точке $x_* \in \Omega$ выполнены условия Куна–Таккера и для любого ненулевого вектора g из G_**

$$\big\langle \mathcal{L}''_{xx}(x_*, u_*) g, g \big\rangle > 0. \tag{15}$$

Тогда x_* — точка строгого локального минимума в задаче (1).

Д о к а з а т е л ь с т в о проведём от противного. Предположим, что x_* не является точкой строгого локального минимума. В этом случае найдётся последовательность $\{y_k\}$ точек из Ω, отличных от x_*, со свойствами

$$y_k \to x_* \ \text{ при } \ k \to \infty, \tag{16}$$

$$f(y_k) \leqslant f(x_*) \ \text{ при всех } k. \tag{17}$$

Представим y_k в виде $y_k = x_* + \lambda_k g_k$, где $\lambda_k = \|y_k - x_*\|$ и $g_k = (y_k - x_*)/\lambda_k$. Очевидно, что $\|g_k\| = 1$. Из ограниченной последовательности $\{g_k\}$ можно выделить сходящуюся подпоследовательность. Сходящуюся подпоследовательность снова обозначим $\{g_k\}$. Имеем $g_k \to g_*$ при $k \to \infty$. По непрерывности нормы $\|g_*\| = 1$. Как показано в [2], $g_* \in G_*$. Согласно (15)

$$\big\langle \mathcal{L}''_{xx}(x_*, u_*)g_*, g_* \big\rangle > 0. \tag{18}$$

Вместе с тем, в силу (17), (3) и (4)

$$\mathcal{L}(y_k, u_*) - \mathcal{L}(x_*, u_*) = f(y_k) - f(x_*) - \sum_{i \in M_1} u_*[i]\big(a_i(y_k) - a_i(x_*)\big) \leqslant$$
$$\leqslant - \sum_{i \in M_1} u_*[i]\, a_i(y_k) \leqslant 0.$$

Воспользуемся формулой Тейлора и тем, что $y_k - x_* = \lambda_k\, g_k$. Получим

$$0 \geqslant \mathcal{L}(y_k, u_*) - \mathcal{L}(x_*, u_*) = \big\langle \mathcal{L}'_x(x_*, u_*), g_k \big\rangle \lambda_k +$$
$$+ \frac{1}{2}\big\langle \mathcal{L}''_{xx}(x_* + \theta_k\, \lambda_k\, g_k, u_*)g_k, g_k \big\rangle \lambda_k^2,$$

где $\theta_k \in (0, 1)$. Отсюда и из (5) следует, что

$$\big\langle \mathcal{L}''_{xx}(x_* + \theta_k\, \lambda_k\, g_k, u_*)g_k, g_k \big\rangle \leqslant 0.$$

Учитывая, что согласно (16) $\lambda_k \to +0$ при $k \to \infty$ и что функция Лагранжа дважды непрерывно дифференцируема по x, в пределе при $k \to \infty$ приходим к неравенству

$$\big\langle \mathcal{L}''_{xx}(x_*, u_*)g_*, g_* \big\rangle \leqslant 0.$$

Но это противоречит (18).

Теорема доказана. $\square$

3°. Приведём пример на использование условий оптимальности второго порядка.

ПРИМЕР. Рассмотрим задачу нелинейного программирования

$$f(x) := (x_1+1)^2 + x_2^2 \to \inf,$$
$$a(x) := x_1 - \alpha\, x_2^2 \geqslant 0,$$

где $\alpha \in \mathbb{R}$ — параметр. Рис. 2 поясняет, что единственным решением этой задачи (при $\alpha > 0$) является точка $x_* = (0,0)$. Проверим, как в точке x_* выполняются условия оптимальности.

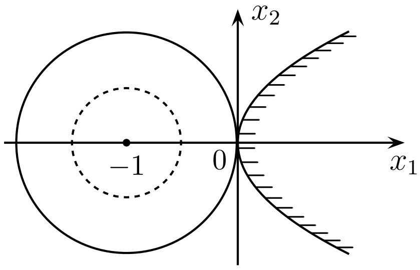

Рис. 2

Имеем

$$f'(x) = (2x_1+2, 2x_2), \qquad a'(x) = (1, -2\alpha\, x_2);$$
$$f'(x_*) = (2,0), \qquad a'(x_*) = (1,0);$$
$$f'(x_*) = 2a'(x_*).$$

Положим $u_* = 2$. Видим, что условия Куна–Таккера (2)–(4) выполняются при всех $\alpha \in \mathbb{R}$.

Конус G_* в данном случае определяется условием $\big\langle a'(x_*), g\big\rangle = 0$ или $g_1 = 0$. Таким образом, G_* состоит из векторов вида $(0, g_2)$, где $g_2 \in \mathbb{R}$.

Далее,

$$f''(x_*) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}, \qquad a''(x_*) = \begin{pmatrix} 0 & 0 \\ 0 & -2\alpha \end{pmatrix},$$

так что

$$\mathcal{L}''_{xx}(x_*, u_*) = f''(x_*) - u_* a''(x_*) = \begin{pmatrix} 2 & 0 \\ 0 & 2+4\alpha \end{pmatrix}.$$

При $g \in G_*$

$$\big\langle \mathcal{L}''_{xx}(x_*, u_*)g, g\big\rangle = (2+4\alpha)g_2^2.$$

Условия $g \in G_*$, $g \neq \mathbb{O}$ означают, что $g_2 \neq 0$. Ясно, что при $\alpha > -\frac{1}{2}$ и всех ненулевых g из G_* выполняется неравенство (15). По теореме 2 x_* — точка строгого локального минимума.

Вместе с тем, при $\alpha < -\frac{1}{2}$ неравенство (8) нарушается на любом ненулевом векторе g из G_*. По теореме 1 x_* не является даже точкой локального минимума. Рис. 3 поясняет эту ситуацию.

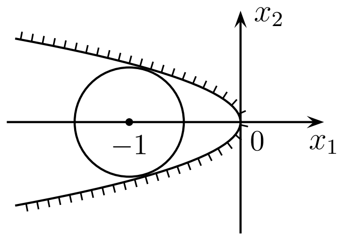

Рис. 3

Остается неопределённым случай $\alpha = -\frac{1}{2}$. Рассмотрим его. Перепишем ограничение задачи в эквивалентном виде

$$2x_1 + x_2^2 \geqslant 0.$$

Для любого плана x имеем

$$f(x) = x_1^2 + 2x_1 + 1 + x_2^2 \geqslant x_1^2 + 1 \geqslant 1.$$

Равенство $f(x) = 1$ достигается при выполнении условий $2x_1 + x_2^2 = 0$ и $x_1 = 0$, т. е. в точке $x_* = (0, 0)$. Таким образом, при $\alpha = -\frac{1}{2}$ точка x_* является единственным решением исходной задачи.

ЛИТЕРАТУРА

1. Фиакко А., Мак-Кормик Г. *Нелинейное программирование.* Пер. с англ. М.: Мир, 1972. 240 с.

2. Малоземов В. Н. *Теорема Куна–Таккера в дифференциальной форме* // Семинар «DHA & CAGD». Избранные доклады. 27 февраля 2010 г. (`http://dha.spb.ru/reps10.shtml#0227`)