ТЕОРЕМА КУНА–ТАККЕРА В ДИФФЕРЕНЦИАЛЬНОЙ ФОРМЕ*

В. Н. Малозёмов

malv@math.spbu.ru

27 февраля 2010 г.

1°. Рассмотрим гладкую задачу нелинейного программирования:

$$\textit{минимизировать } f(x) \tag{1}$$

при ограничениях

$$\begin{aligned} a_i(x) &\geqslant 0, \quad i \in M_1; \\ a_i(x) &= 0, \quad i \in M_2; \\ x &\in U. \end{aligned}$$

Здесь $U \subset \mathbb{R}^N$ — открытое множество и функции $f(x)$, $a_i(x)$ при всех $i \in \in M_1 \cup M_2$ непрерывно дифференцируемы на U.

Вектор $x \in \mathbb{R}^N$, удовлетворяющий ограничениям задачи (1), называется её *планом*. Множество планов обозначим Ω.

Пусть $x \in \Omega$. Обозначим $M = M_1 \cup M_2$,

$$\begin{aligned} M_1(x) &= \bigl\{ i \in M_1 \mid \ a_i(x) = 0 \bigr\}, \\ I(x) &= M_1(x) \cup M_2. \end{aligned}$$

На рис. 1 схематично изображены введённые индексные множества. Ясно, что

$$M \setminus I(x) = M_1 \setminus M_1(x). \tag{2}$$

В частности,

$$a_i(x) > 0 \quad \text{при} \quad i \in M \setminus I(x). \tag{3}$$

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

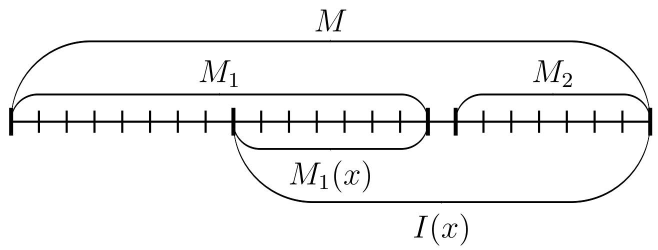

Рис. 1.

Множество $I(x)$ называется *множеством индексов активных ограничений* (ограничение с индексом $i \in I(x)$ активно в том смысле, что выполняется как равенство). Если при $i \in I(x)$ градиенты $a_i'(x)$ линейно независимы, то ограничения в точке x называются *регулярными*.

Множество планов Ω задачи (1) и целевая функция $f(x)$ могут быть довольно сложными. Нас интересуют точки локального минимума.

План x_* называется *точкой локального минимума* в задаче (1), если при некотором $\delta > 0$

$$f(x) \geqslant f(x_*) \qquad \forall x \in \Omega \cap U_\delta(x_*), \tag{4}$$

где $U_\delta(x_*) = \big\{x \in \mathbb{R}^N \mid \|x - x_*\| < \delta\big\}$ — открытая δ-окрестность точки x_*.

Теперь мы можем сформулировать теорему Куна–Таккера в дифференциальной форме [1].

ТЕОРЕМА 1. *Пусть $x_* \in \Omega$ — точка локального минимума в задаче* (1) *и ограничения в этой точке регулярны. Тогда найдётся вектор $u_* = u_*[I(x_*)]$, такой, что*

$$\begin{aligned} f'(x_*) &= \sum_{i \in I(x_*)} u_*[i] a_i'(x_*), \\ u_*[i] &\geqslant 0, \qquad i \in M_1(x_*). \end{aligned} \tag{5}$$

Теорема Куна–Таккера входит в курс «Экстремальные задачи», который я читаю с 1986 г. на математико-механическом факультете СПбГУ для студентов третьего курса отделения прикладной математики и информатики. За основу было взято доказательство этой теоремы из книги [2], с. 33–42. Постепенно доказательство совершенствовалось. В докладе приведён современный вариант доказательства теоремы Куна–Таккера. Обсуждается вопрос о достаточности условий Куна–Таккера.

2°. Нам понадобятся два вспомогательных утверждения о линейных и нелинейных системах уравнений.

ЛЕММА 1 ([3], с. 25)**.** *Система линейных уравнений*

$$A[N,M] \times u[M] = c[N]$$

имеет решение $u_*[M]$ *со свойством* $u_*[M_1] \geqslant \mathbb{O}[M_1]$, *где* $M_1 \subset M$, *тогда и только тогда, когда для любого вектора* $g \in \mathbb{R}^N$, *удовлетворяющего условиям*

$$g[N] \times A[N,M_1] \geqslant \mathbb{O}[M_1],$$
$$g[N] \times A[N,M \setminus M_1] = \mathbb{O}[M \setminus M_1],$$

выполняется неравенство $\langle c, g\rangle \geqslant 0$.

Теперь рассмотрим в $\mathbb{R}^N$ систему нелинейных уравнений

$$a_i(x) = 0, \qquad i \in I. \tag{6}$$

ЛЕММА 2 ([2], с. 37–38)**.** *Пусть* x_0 *удовлетворяет системе* (6). *Если при этом все функции* $a_i(x)$ *непрерывно дифференцируемы в окрестности* x_0 *и градиенты* $a_i'(x_0)$ *линейно независимы, то для любого ненулевого вектора* $g_0 \in \mathbb{R}^N$, *удовлетворяющего условию* $\langle a_i'(x_0), g_0\rangle = 0$ *при* $i \in I$, *можно построить в* $\mathbb{R}^N$ *параметрическую кривую* $x = x(t)$, *непрерывно дифференцируемую в окрестности точки* $t = 0$, *со свойствами*

$$x(0) = x_0, \qquad x'(0) = g_0,$$
$$a_i\big(x(t)\big) = 0 \ \text{ при } \ i \in I \ \text{ и малых } \ t.$$

На рис. 2 иллюстрируется содержание этой леммы в случае, когда система (6) состоит из одного уравнения $a(x) = 0$, и $x = (x^1, x^2, x^3)$.

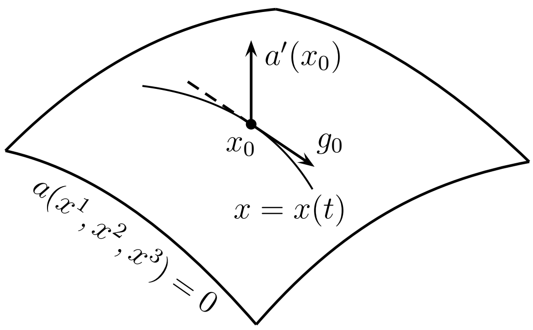

Рис. 2.

Лемму 2 я называю *основной леммой нелинейного программирования.*

3°. Обратимся к доказательству теоремы Куна–Таккера. По существу, нужно проверить, что линейная относительно u_* система (5) совместна.

Воспользуемся леммой 1. Покажем, что для любого вектора $g_0 \in \mathbb{R}^N$, удовлетворяющего соотношениям

$$\begin{aligned} &\langle a_i'(x_*), g_0\rangle \geqslant 0, \quad i \in M_1(x_*); \\ &\langle a_i'(x_*), g_0\rangle = 0, \quad i \in M_2, \end{aligned} \tag{7}$$

выполняется неравенство $\langle f'(x_*), g_0\rangle \geqslant 0$. Отсюда будет следовать заключение теоремы 1.

Зафиксируем соответствующий вектор g_0. Можно считать, что $g_0 \neq \mathbb{O}$. Введем индексное множество

$$I_0 = \bigl\{ i \in M_1(x_*) \mid \langle a_i'(x_*), g_0\rangle = 0 \bigr\} \cup M_2.$$

Согласно (7),

$$\langle a_i'(x_*), g_0\rangle = 0, \qquad i \in I_0, \tag{8}$$

$$\langle a_i'(x_*), g_0\rangle > 0, \qquad i \in I(x_*) \setminus I_0. \tag{9}$$

Рассмотрим систему нелинейных уравнений

$$a_i(x) = 0, \qquad i \in I_0.$$

Точка x_* ей удовлетворяет, поскольку $I_0 \subset I(x_*)$. На основании (8) и леммы 2 в $\mathbb{R}^N$ можно построить параметрическую кривую $x = x(t)$, непрерывно дифференцируемую в окрестности точки $t = 0$, со свойствами

$$x(0) = x_*, \qquad x'(0) = g_0, \tag{10}$$

$$a_i\bigl(x(t)\bigr) = 0 \text{ при } i \in I_0 \text{ и малых } t. \tag{11}$$

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

При $i \in I_0$ (в частности, при $i \in M_2$) выполняется равенство (11). Пусть $i \in I(x_*) \setminus I_0$. Согласно (10)

$$\begin{aligned} a_i\bigl(x(t)\bigr) &= a_i\bigl(x(0)\bigr) + \bigl\langle a_i'\bigl(x(0)\bigr), x'(0)\bigr\rangle\, t + o(t) = \\ &= t\Bigl[\langle a_i'(x_*), g_0\rangle + \frac{o(t)}{t}\Bigr]. \end{aligned}$$

Последнее выражение при малых $t > 0$ положительно в силу (9).

Осталось рассмотреть индексы $i \in M \setminus I(x_*)$. На них согласно (3) и теореме о стабилизации знака у непрерывной функции при малых t выполняется строгое неравенство $a_i\bigl(x(t)\bigr) > 0$. Таким образом, точка $x(t)$ при малых $t > 0$ удовлетворяет всем ограничениям задачи (1), то есть принадлежит Ω.

По условию теоремы x_* — точка локального минимума в задаче (1), поэтому при малых $t > 0$

$$0 \leqslant f\bigl(x(t)\bigr) - f(x_*) = f\bigl(x(t)\bigr) - f\bigl(x(0)\bigr) = \bigl\langle f'\bigl(x(0)\bigr), x'(0)\bigr\rangle\, t + o(t) =$$
$$= t\Bigl[\langle f'(x_*), g_0\rangle + \frac{o(t)}{t}\Bigr].$$

Поделив это неравенство на $t > 0$ и перейдя к пределу при $t \to +0$, получим $\langle f'(x_*), g_0\rangle \geqslant 0$.

Теорема доказана. □

4°. Условие регулярности ограничений в точке локального минимума существенно для справедливости теоремы Куна–Таккера. Приведём соответствующий пример.

ПРИМЕР 1. На плоскости $x = (u, v)$ введем множество Ω с помощью неравенств

$$a_1(x) := u^3 - v \geqslant 0,$$
$$a_2(x) := -u^4 + v \geqslant 0.$$

Это множество представляет собой лунку (см. рис. 3).

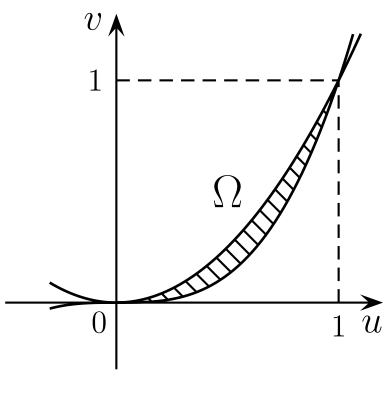

Рис. 3.

Рассмотрим экстремальную задачу

$$f_1(x) := u \to \min_{x\in\Omega}. \tag{12}$$

Очевидно, что единственным решением задачи (12) является $x_* = (0, 0)$ (как единственная точка из Ω с наименьшей первой координатой). Выясним, как в этом случае выглядят соотношения (5).

Имеем

$$f_1'(x) = (1, 0), \qquad a_1'(x) = (3u^2, -1), \qquad a_2'(x) = (-4u^3, 1);$$
$$I(x_*) = M(x_*) = \{1, 2\};$$
$$f_1'(x_*) = (1, 0), \qquad a_1'(x_*) = (0, -1), \qquad a_2'(x_*) = (0, 1).$$

Соотношения (5) принимают вид

$$\begin{gathered} f_1'(x_*) = u_*[1]\, a_1'(x_*) + u_*[2]\, a_2'(x_*), \\ u_*[1] \geqslant 0, \quad u_*[2] \geqslant 0. \end{gathered} \tag{13}$$

Однако равенство в (13) не может выполняться ни при каких $u_*[1]$ и $u_*[2]$, поскольку по первой координате оно переписывается так: $1 = 0$.

Казалось бы, получено противоречие с теоремой 1. Но никакого противоречия нет. В точке x_* ограничения задачи (12) нерегулярны. Действительно, $a_1'(x_*) + a_2'(x_*) = 0$, так что градиенты $a_1'(x_*)$ и $a_2'(x_*)$ линейно зависимы. А без условия регулярности ограничений справедливость теоремы 1 не гарантируется.

Впрочем, без гарантии условия (5) могут выполняться и при отсутствии регулярности ограничений. В качестве примера рассмотрим задачу минимизации функции $f_2(x) := v$ на той же лунке Ω. Её единственным решением является $x_* = (0, 0)$, причем, как отмечалось выше, ограничения в этой точке нерегулярны. Вместе с тем, $f_2'(x_*) = (0, 1)$ и $f_2'(x_*) = a_2'(x_*)$. Видим, что условия (5) выполняются при $u_*[1] = 0$, $u_*[2] = 1$.

5°. В соотношения (5) входят первые производные функций, участвующих в постановке задачи (1), поэтому их называют необходимыми условиями оптимальности первого порядка. Выведем достаточные условия первого порядка для точки строгого локального минимума.

Напомним, что план x_* называется *точкой строгого локального минимума* в задаче (1), если при некотором $\delta > 0$

$$f(x) > f(x_*) \qquad \forall x \in \Omega \cap \dot{U}_\delta(x_*),$$

где $\dot{U}_\delta(x_*) = U_\delta(x_*) \setminus \{x_*\}$ — проколотая δ-окрестность точки x_*.

Пусть в точке $x_* \in \Omega$ выполнено необходимое условие оптимальности (5). Дополнительно обозначим (см. рис. 4)

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

Отметим, что

$$u_*[i] = 0, \qquad i \in I(x_*) \setminus I^+(x_*). \tag{14}$$

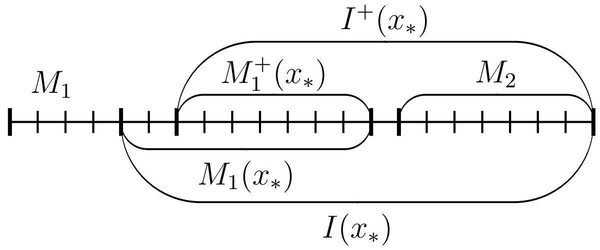

Рис. 4.

Рассмотрим систему линейных соотношений

$$\begin{aligned} &\langle a_i'(x_*), g\rangle = 0, \qquad && i \in I^+(x_*); \\ &\langle a_i'(x_*), g\rangle \geqslant 0, \qquad && i \in I(x_*) \setminus I^+(x_*). \end{aligned} \tag{15}$$

Множество векторов g, удовлетворяющих (15), является конусом в $\mathbb{R}^N$. Обозначим его G_*.

ТЕОРЕМА 2. *Пусть в точке $x_* \in \Omega$ выполняются условия* (5). *Если при этом $G_* = \{\mathbb{O}\}$, то x_* — точка строгого локального минимума в задаче* (1).

Доказательство. Допустим противное. Тогда найдётся последовательность планов $\{y_k\}$ со свойствами

$$y_k \neq x_* \text{ при всех } k; \quad y_k \to x_* \text{ при } k \to \infty;$$
$$f(y_k) \leqslant f(x_*) \text{ при всех } k.$$

Представим y_k в виде $y_k = x_* + \lambda_k g_k$, где $g_k = \dfrac{y_k - x_*}{\|y_k - x_*\|}$ и $\lambda_k = \|y_k - x_*\|$. Ясно, что $\lambda_k > 0$ и $\lambda_k \to +0$ при $k \to \infty$. По определению $\|g_k\| = 1$ при всех k, поэтому можно считать, что $g_k \to g_*$. В силу непрерывности нормы $\|g_*\| = 1$. Покажем, что $g_* \in G_*$.

При $i \in M_1(x_*)$ по теореме о среднем имеем

$$0 \leqslant a_i(y_k) - a_i(x_*) = \langle a_i'(\xi_{ik}), g_k\rangle \lambda_k, \tag{16}$$

где точка ξ_{ik} принадлежит отрезку $[x_*, y_k]$. Аналогично при $i \in M_2$

$$0 = a_i(y_k) - a_i(x_*) = \langle a_i'(\xi_{ik}), g_k\rangle \lambda_k, \tag{17}$$

где ξ_{ik} также принадлежит отрезку $[x_*, y_k]$. Поделив соотношения (16) и (17) на $\lambda_k > 0$ и перейдя к пределу при $k \to \infty$, получим

$$\begin{aligned} &\langle a_i'(x_*), g_*\rangle \geqslant 0, \qquad && i \in M_1(x_*); \\ &\langle a_i'(x_*), g_*\rangle = 0, \qquad && i \in M_2. \end{aligned} \tag{18}$$

Кроме того, из неравенства

$$0 \geqslant f(y_k) - f(x_*) = \langle f'(\eta_k), g_k \rangle \lambda_k,$$

где η_k принадлежит отрезку $[x_*, y_k]$, следует, что

$$\langle f'(x_*), g_* \rangle \leqslant 0. \tag{19}$$

Покажем, что на самом деле

$$\langle a_i'(x_*), g_* \rangle = 0, \qquad i \in M_1^+(x_*). \tag{20}$$

Согласно (19), (5), (14) и (18) имеем

$$0 \geqslant \langle f'(x_*), g_* \rangle = \sum_{i \in I(x_*)} u_*[i] \langle a_i'(x_*), g_* \rangle =$$
$$= \sum_{i \in I^+(x_*)} u_*[i] \langle a_i'(x_*), g_* \rangle = \sum_{i \in M_1^+(x_*)} u_*[i] \langle a_i'(x_*), g_* \rangle.$$

В последней сумме все слагаемые неотрицательны, а сумма неположительна. Значит, все слагаемые равны нулю. Если учесть, что $u_*[i] > 0$ при $i \in M_1^+(x_*)$, то приходим к равенствам (20).

На основании (18) и (20) заключаем, что $g_* \in G_*$. Но это противоречит условию $G_* = \{\mathbb{O}\}$.

Теорема доказана. $\square$

СЛЕДСТВИЕ. *Пусть в точке $x_* \in \Omega$ выполнено условие* (5). *Если при этом $|I^+(x_*)| = |N|$ и градиенты $a_i'(x_*)$ при $i \in I^+(x_*)$ линейно независимы, то x_* — точка строгого локального минимума в задаче* (1).

Действительно, в данном случае $G_* = \{\mathbb{O}\}$, что следует из условия

$$\langle a_i'(x_*), g \rangle = 0, \quad i \in I^+(x_*),$$

входящего в определение конуса G_*.

ПРИМЕР 2. Рассмотрим задачу минимизации функции $f_3(x) = -u$ на той же лунке Ω, что и в примере 1. Её единственным решением является $x_* = (1, 1)$ (как единственная точка из Ω с наибольшей первой координатой). В данном случае

$$I(x_*) = \{1, 2\}; \qquad f_3'(x_*) = (-1, 0), \quad a_1'(x_*) = (3, -1), \quad a_2'(x_*) = (-4, 1),$$

так что

$$f_3'(x_*) = a_1'(x_*) + a_2'(x_*).$$

Условия (5) выполнены с $u_*[1] = u_*[2] = 1$. Более того, $I^+(x_*) = \{1, 2\}$ и градиенты $a_1'(x_*)$, $a_2'(x_*)$ линейно независимы. По следствию из теоремы 2 план x_* является точкой строгого локального минимума. Это соответствует отмеченному выше факту оптимальности x_*.

6°. Условия оптимальности (5) можно представить в другом виде.

Будем говорить, что в точке $x_* \in \Omega$ выполняются *условия Куна–Таккера*, если найдется вектор $u_* = u_*[M]$ со свойствами

$$f'(x_*) = \sum_{i \in M} u_*[i]\, a_i'(x_*); \tag{21}$$

$$u_*[i]\, a_i(x_*) = 0, \qquad i \in M_1; \tag{22}$$

$$u_*[i] \geqslant 0, \qquad i \in M_1. \tag{23}$$

ЛЕММА 3. *Условия оптимальности* (5) *равносильны условиям Куна–Таккера.*

Доказательство. Пусть выполнены условия (5). Дополним вектор $u_*[I(x_*)]$ компонентами $u_*[i] = 0$ при $i \in M \setminus I(x_*)$. Получим вектор $u_*[M]$. Очевидно, что он обладает свойствами (21) и (23). Что касается свойства (22), то следует рассмотреть два случая: $i \in M_1(x_*)$ и $i \in M_1 \setminus M_1(x_*) = M \setminus I(x_*)$ (см. (2)). В первом случае равенство выполняется, поскольку $a_i(x_*) = 0$. Во втором случае нужно учесть, что $u_*[i] = 0$.

Наоборот, пусть выполнены условия Куна–Таккера. В силу (22)

$$u_*[i] = 0 \quad \text{при} \quad i \in M_1 \setminus M_1(x_*) = M \setminus I(x_*).$$

Очевидно, что вектор $u_*[I(x_*)]$ удовлетворяет условиям (5).

Лемма доказана. □

С учетом леммы 3 теорему 1 можно переформулировать так.

Пусть $x_* \in \Omega$ — точка локального минимума в задаче (1) *и ограничения в ней регулярны. Тогда в x_* выполняются условия Куна–Таккера.*

Отметим, что условия (21), (22), (23) называются соответственно *условием Лагранжа*, *условием дополнительности* и *условием неотрицательности.*

7°. Обратимся к задаче нелинейного программирования с линейными ограничениями:

$$\begin{gathered} f(x) \to \inf, \\ A[M_1, N] \times x[N] \geqslant b[M_1], \\ A[M_2, N] \times x[N] = b[M_2]. \end{gathered} \tag{24}$$

Множество планов, как и раньше, обозначим Ω. Предположим, что целевая функция $f(x)$ дифференцируема на некотором открытом множестве, содержащем Ω.

ТЕОРЕМА 3. *Пусть $x_* \in \Omega$ — точка локального минимума в задаче (24). Тогда найдётся вектор $u_* = u_*[M]$, где $M = M_1 \cup M_2$, со свойствами*

$$\begin{gathered} f'(x_*)[N] = u_*[M] \times A[M,N]; \\ u_*[i] \times \big(A[i,N] \times x_*[N] - b[i]\big) = 0, \qquad i \in M_1; \\ u_*[i] \geqslant 0, \qquad i \in M_1. \end{gathered} \tag{25}$$

Условие регулярности ограничений в точке x_* здесь не требуется.

Простое доказательство теоремы 3, опирающееся на критерий оптимальности в линейном программировании, приведено в [3], с. 88.

ТЕОРЕМА 4. *Предположим, что целевая функция $f(x)$ задачи (24) выпукла и дифференцируема на некотором открытом выпуклом множестве, содержащем Ω. Если в точке $x_* \in \Omega$ выполняются условия (25), то x_* — точка глобального минимума $f(x)$ на Ω.*

По поводу доказательства см. [3], с. 91.

ЛИТЕРАТУРА

1. Kuhn H. W. *Nonlinear programming: a historical view* / In: SIAM-AMS Proceedings. Vol. IX. AMS, Providence, 1976. P. 1–26.

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

3. Гаврин М. К., Малоземов В. Н. *Экстремальные задачи с линейными ограничениями.* Л.: Изд-во ЛГУ, 1984. 176 с.