ЖУРНАЛ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ ФИЗИКИ, 2020, том 60, № 9, с. 1462–1471

УДК 519.855

НОВЫЙ ВЗГЛЯД НА НЕКОТОРЫЕ ОСНОВОПОЛАГАЮЩИЕ РЕЗУЛЬТАТЫ В ОПТИМИЗАЦИИ

© 2020 г. Ю. Г. Евтушенко[1,2,3,*], А. А. Третьяков[1,4,5,]**

[1] 119333 Москва, ул. Вавилова, 40, ВЦ РАН ФИЦ ИУ РАН, Россия

[2] 141700 Долгопрудный, М.о., Институтский пер., 9, МФТИ, Россия

[3] 125080 Москва, Волоколамское шоссе, 4, МАИ, Россия

[4] 01-447 Warsaw, Newelska 6, System Res. Inst., Polish Acad. Sciences, Poland

[5] 08-110 Siedlce, Siedlce University, Faculty of Sciences, Poland

**e-mail: yuri-evtushenko@yandex.ru*

***e-mail: tret@ap.siedlce.pl*

Поступила в редакцию 18.07.2019 г.
Переработанный вариант 31.08.2019 г.
Принята к публикации 09.04.2020 г.

В статье для некоторых фундаментальных результатов оптимизации предлагаются новые доказательства, которые носят нетрадиционный характер и позволяют по-новому взглянуть на известные положения. При этом используются конструкции теории p-регулярности для обоснования обсуждаемых фактов, а также 2-фактор-метод для решения вырожденных задач. Библ. 8.

Ключевые слова: оптимизация, вырожденность, теорема Куна–Таккера, лемма Фаркаша, замкнутость конуса, 2-фактор-метод, 2-регулярность.

DOI: 10.31857/S0044466920090082

Рассматривается несколько основных результатов теории оптимизации, такие как теорема Лагранжа об условиях оптимальности в задаче с ограничениями равенствами, теорема Куна–Таккера об условиях оптимальности в задаче с ограничениями неравенствами, теорема Люстерника, достаточные условия оптимальности для задачи математического программирования, теорема Фаркаша и теорема о замкнутости конуса – с нетрадиционной точки зрения.

Показано, что во многих случаях доказательство этих центральных для оптимизации результатов можно существенно упростить, не используя идей отделимости, замкнутости конуса, теоремы Фаркаша, выпуклости, предельных переходов или метода штрафных функций, принципа сжимающих отображений или теорему о неявной функции и т.д. и т.п.

1. ЭЛЕМЕНТАРНОЕ ДОКАЗАТЕЛЬСТВО ТЕОРЕМЫ ЛАГРАНЖА ДЛЯ ЗАДАЧИ С РАВЕНСТВАМИ

Рассматривается задача

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

при условиях

$$F(x) = 0_m, \tag{2}$$

где $F = (f_1, \ldots, f_m)^\top$, $\varphi, F \in C^2(\mathbb{R}^n)$, $f_i\colon \mathbb{R}^n \to R$, $i = \overline{1,m}$, и 0_m есть вектор-столбец. Обозначим через x^* решение задачи (1), (2). Тогда имеет место классическая теорема Лагранжа о необходимых условиях оптимальности [1], [2].

Теорема 1. *Пусть* $\varphi, F \in C^2(\mathbb{R}^n)$ *и матрица Якоби*

$$F'(x^*) = \left(\frac{\partial f_i(x^*)}{\partial x_j}\right)_{\substack{i=\overline{1,m}\\ j=\overline{1,n}}}$$

не вырождена. Тогда $\varphi'(x^*) \in \operatorname{Lin}\{f_i'(x^*), i = 1, \ldots, m\}$.

Известные доказательства этой теоремы весьма громоздки и используют такие факты, как теорема о неявной функции [1], предельные переходы [2], [3], теорема Люстреника [1] и т.д. Мы покажем, как можно в доказательстве избежать этих сложностей и получить основной результат только на основе теоремы Вейерштрасса.

Доказательство. Предположим от противного, что $\varphi'(x^*) \notin \operatorname{Lin}\{f_i'(x^*), i = 1,\dots,m\}$. Тогда $\varphi'(x^*) = h + \bar{h}$, где $h \in \operatorname{Ker} F'(x^*)$, $\bar{h} \in \operatorname{Lin}\{f_i'(x^*), i = 1,\dots,m\}$. Тогда для $\alpha \in (0,\varepsilon)$ и $\xi \in \mathbb{R}^m$, $\|\xi\| = 1$, $\left\|F(x^* - \alpha h + \alpha^{3/2}F'(x^*)^\top\xi)\right\|^2 \geq c\alpha^3$, и $\|F(x^* - \alpha h)\|^2 \leq c\alpha^4$, где $\varepsilon > 0$ – достаточно малое и $c > 0$ – независимая константа. Это означает, по теореме Вейерштрасса, что существует $\eta(\alpha)$: $R^1 \to \mathbb{R}^m$ такое, что

$$\min_{\|\eta\|\leq 1}\left\|F(x^* - \alpha h + \alpha^{3/2}F'(x^*)^\top\eta)\right\|^2 = \left\|F(x^* - \alpha h + \alpha^{3/2}F'(x^*)^\top\eta(\alpha))\right\|^2$$

и $\|\eta(\alpha)\| \leq 1$. Это означает, что

$$\left(\left\|F(x^* - \alpha h + \alpha^{3/2}F'(x^*)^\top\eta)\right\|^2\right)'_{\eta=\eta(\alpha)} = 0_n$$

или

$$2F'(x^* - \alpha h + \alpha^{3/2}F'(x^*)^\top\eta(\alpha))\cdot F'(x^*)^\top \cdot F(x^* - \alpha h + \alpha^{3/2}F'(x^*)^\top\eta(\alpha)) = 0_n,$$

что подразумевает равенство

$$F(x^* - \alpha h + \alpha^{3/2}F'(x^*)^\top\eta(\alpha)) = 0_n$$

и, значит,

$$\varphi(x^* - \alpha h + \alpha^{3/2}F'(x^*)^\top\eta(\alpha)) = \varphi(x^*) - \alpha\langle\varphi'(x^*), h\rangle + o(\alpha) < \varphi(x^*),$$

что противоречит локальной минимальности точки x^* и доказывает теорему 1.

2. НОВОЕ ЭЛЕМЕНТАРНОЕ ДОКАЗАТЕЛЬСТВО ТЕОРЕМЫ КУНА–ТАККЕРА

Рассматривается задача

$$\min \varphi(x) \tag{3}$$

при ограничениях

$$g_i(x) \leq 0, \quad i = 1,\dots,m. \tag{4}$$

Здесь $x \in \mathbb{R}^n$, $\varphi, g_i \in C^1(\mathbb{R}^n)$, $i = 1,\dots,m$. Основной результат для характеризации решения x^* задачи (3), (4) – теорема Куна–Таккера (КТ), которая устанавливает необходимые условия для этой задачи (3), (4).

Теорема 2. *Пусть x^* – решение задачи* (3), (4) *и* $\varphi, g_i \in C^1(\mathbb{R}^n)$, $i = 1,\dots,m$.

Тогда найдутся такие неотрицательные множители $\lambda_0^*, \lambda_1^*, \dots, \lambda_m^*$, что выполнены условия

$$\lambda_0^*\varphi'(x^*) + \sum_{i=1}^m \lambda_i^* g_i'(x^*) = 0_n, \quad \lambda_0^* + \lambda_1^* + \dots + \lambda_m^* = 1. \tag{5}$$

Обозначим множество активных ограничений через $I(x^*) = \{i = 1,\dots,m \mid g_i(x^*) = 0\}$. Традиционные доказательства данной теоремы достаточно сложные и используют в своих конструкциях такие результаты, как теорему о неявной функции, теорему о замкнутости конуса, лемму Фаркаша, теоремы об отделимости, предельные теоремы и т.д. Приведем доказательство, которое не использует вышеупомянутые фундаментальные результаты, используя и опираясь только на тривиальную лемму о свойстве проекции.

Лемма 1. *Пусть $\mathcal{K}$ – замкнутое выпуклое множество в $\mathbb{R}^n$.*

Тогда для проекции $p = \operatorname{Pr}_{\mathcal{K}} 0_n$ нулевого вектора 0_n на множество $\mathcal{K}$ имеет место неравенство

$$\langle z - p, p\rangle \geq 0 \tag{6}$$

для всех $z \in \mathcal{K}$.

Доказательство данной леммы тривиально и мы его не приводим. Пусть

$$\mathcal{K} = \left\{ y \in \mathbb{R}^n \,|\, y = \lambda_0 \varphi'(x^*) + \sum_{i=1}^m \lambda_i g_i'(x^*),\ \lambda_0 \ge 0,\ \lambda_i \ge 0,\ i = 1,\ldots,m,\ \lambda_0 + \lambda_1 + \ldots + \lambda_m = 1 \right\}.$$

Доказательство теоремы КТ. Покажем, что

$$0_n \in \mathcal{K}, \tag{7}$$

так как соотношение (7) эквивалентно (5). Действительно, если это не так, т.е. $0_n \notin \mathcal{K}$, то поскольку $\varphi'(x^*) \in \mathcal{K}$, из леммы 1 следует

$$\langle \varphi'(x^*), -p \rangle \le -\|p\|^2 < 0 \tag{8}$$

и, поскольку $g_i'(x^*) \in \mathcal{K}$, $i = 1,\ldots,m$, имеем

$$\langle g_i'(x^*), -p \rangle \le -\|p\|^2 < 0. \tag{9}$$

Отсюда следует, что

$$\varphi(x^* - \alpha p) < \varphi(x^*)$$

и

$$g_i(x^* - \alpha p) < g_i(x^*) < 0, \quad i = 1,\ldots,m,$$

при $\alpha > 0$ достаточно малых. А это противоречит минимальности точки x^*. Теорема доказана.

3. ТЕОРЕМА ФАРКАША И ТЕОРЕМА О ЗАМКНУТОСТИ КОНУСА

Важную роль при построении условий оптимальности и конструировании численных методов играют так называемые теоремы об альтернативах, одна из которых может быть сформулирована в следующем виде.

Теорема 3 (Фаркаша). *Пусть векторы* $\eta, \xi_i \in \mathbb{R}^n$, $i = 1,\ldots,m$, *и для каждого* $x \in \mathbb{R}^n$ *такого, что* $\langle \xi_i, x \rangle \ge 0$, $i = 1,\ldots,m$, *выполнено неравенство*

$$\langle \eta, x \rangle \ge 0. \tag{10}$$

Назовем это условием Фаркаша.

Тогда

$$\eta \in \operatorname{Cone}\{\xi_i \,|\, i = 1,\ldots,m\} \left\{ y \in \mathbb{R}^n \,|\, y = \sum_{i=1}^m \alpha_i \xi_i,\ \alpha_i \ge 0,\ i = 1,\ldots,m \right\}.$$

Доказательство. Введем в рассмотрение функцию

$$\Phi(x) = \langle \eta, x \rangle + \sum_{i=1}^m \left(\langle \xi_i, x \rangle_- \right)^2,$$

здесь

$$a_- = \begin{cases} a, & a \le 0, \\ 0, & a \ge 0, \quad a \in \mathbb{R}^n. \end{cases}$$

При выполнении условий Фаркаша существует $y^* = \operatorname*{Argmin}_{x \in \mathbb{R}^n} \Phi(x)$. По теореме Ферма $\Phi'(y^*) = 0_n$ или $\eta + \sum_{i=1}^m 2\langle \xi_i, y^* \rangle_- \xi_i = 0_n$. Обозначив через $-\lambda_i^* = 2\langle \xi_i, y^* \rangle_-$, $i = 1,\ldots,m$, получаем $\eta = \sum_{i=1}^m \lambda_i^* x_i^*$, $\lambda_i^* \ge 0$, $i = 1,\ldots,m$, т.е. $\eta \in \operatorname{Cone}\{\xi_i \,|\, i = 1,\ldots,m\}$.

Замечание 1. Из доказательства явно следует вид множителей Фаркаша λ_i^*.

Известные доказательства этой теоремы в большинстве своем полагаются на теоремы отделимости, предельные переходы и, обязательно, на теорему о замкнутости конуса. Однако существующие доказательства теоремы о замкнутости конуса весьма громоздки. Приведем здесь доказательство, которое отличается своей простотой и изяществом. Это доказательство опирается на лемму Фаркаша.

Теорема 4 (о замкнутости конуса). *Множество*

$$\mathcal{K} = \left\{ y \in \mathbb{R}^n \,|\, y = \sum_{i=1}^{m} \lambda_i a_i,\ a_i \in \mathbb{R}^n,\ \lambda_i \geq 0,\ i = 1,\ldots,m \right\}$$

замкнуто.

Доказательство. Пусть последовательность $\{y_i\}$, $i = 1,2,\ldots$, такова, что $y_i \in \mathcal{K}$, $i = 1,2,\ldots$, $\|y_i\| \leq c$, $i = 1,2,\ldots$, и $y_i \to y$ при $i \to \infty$. Нужно показать, что $y \in \mathcal{K}$. Действительно, $\forall x \in \mathbb{R}^n$ и такого, что $\langle a_j, x \rangle \geq 0$, $j = 1,\ldots,m$, следует, что $\langle y_i, x \rangle \geq 0$, $i = 1,2,\ldots$. Но тогда, переходя к пределу при $i \to \infty$, получаем $\forall x \in \mathbb{R}^n$ и такого, что $\langle a_j, x \rangle \geq 0$, $j = 1,\ldots,m$, следует

$$\lim_{i\to\infty} \langle y_i, x \rangle = \langle y, x \rangle \geq 0,$$

и по теореме Фаркаша $y \in \operatorname{Cone}\{a_j \,|\, j = 1,\ldots,m\}$, т.е. $\mathcal{K}$ замкнуто.

Таким образом, теорема Фаркаша и о замкнутом конусе эквивалентны, т.е. одна следует из другой. Это новый взгляд на данные основополагающие результаты в оптимизации.

4. ОБ ИЗБЫТОЧНОСТИ УСЛОВИЙ В ТЕОРЕМЕ ФАРКАША

Основные теоремы об условиях оптимальности в задаче с неравенствами получены с использованием теоремы Фаркаша. Однако оказывается, что требование (10): $\forall x \in \mathbb{R}^n$ такого, что $\langle \xi, x \rangle \geq 0$, $i = 1,\ldots,m$, следует $\langle \eta, x \rangle \geq 0$ избыточно. Достаточно взять некоторый конечный набор элементов z_j и если на нем выполнено (10), а именно $\forall z_j$, $j = 1,\ldots,k$, из того, что $\langle \xi_i, z_j \rangle \geq 0$, $i = 1,\ldots,m \;\Rightarrow\; \langle \eta, z_j \rangle \geq 0$, $j = 1,\ldots,k$, то $\eta = \sum_{i=1}^{m} \alpha_i \xi_i$, $\alpha_i \geq 0$, $i = 1,\ldots,m$. Например, $\xi_1 = (0,1)^\top$, $\xi_2 = (1,0)^\top$, $\eta = (1,1)^\top$, $z_1 = (1,0)^\top$, $z_2 = (0,1)^\top$. Очевидно, что $\langle \xi_i, z_j \rangle \geq 0$, $i = 1,2$ и $\langle \eta, z_j \rangle \geq 0$, $j = 1,2 \Rightarrow$ $\Rightarrow \eta = \alpha_1\xi_1 + \alpha_2\xi_2$, $\alpha_1 = 1$, $\alpha_2 = 1$.

Набор векторов $\{z_j\}$ определяется следующим образом. Пусть $\operatorname{span}\{\xi_i \,|\, i = 1,\ldots,m\} \triangleq \mathbb{R}^p$, $p \leq n$, $\xi_i \in \mathbb{R}^n$, $\eta \in \mathbb{R}^n$. Тогда $\mathbb{R}^n = \mathbb{R}^p \oplus \mathbb{R}^{n-p}$ и пусть $y_1,\ldots,y_{n-p}$ – базис в $\mathbb{R}^{n-p}$. Тогда

1) чтобы η принадлежал $\operatorname{Cone}\{\xi_i \,|\, i = 1,\ldots,m\}$ нужно, чтобы $\eta \in \operatorname{span}\{\xi_i \,|\, i = 1,\ldots,m\}$, а значит, $\langle \eta, y_j \rangle = 0$, $j = 1,\ldots,n-p$. Это можно записать как неравенство: из того, что $\langle \xi_i, \pm y_j \rangle \geq 0 \Rightarrow$ $\Rightarrow \langle \eta, \pm y_j \rangle \geq 0$ $j = 1,\ldots,n-p$;

2) имеем $\eta \in \operatorname{span}\{\xi_i \,|\, i = 1,\ldots,m\}$. Далее в $\mathbb{R}^p$ строим выпуклый конус $\mathcal{K}_p \triangleq$ $\triangleq \{z \in \mathbb{R}^n \cap \mathbb{R}^p \,|\, \langle \xi_i, z \rangle \geq 0,\ i = 1,\ldots,m\}$. Пусть $x_1,\ldots,x_r$ – его крайние (одномерные) ребра. Очевидно, что их число r конечно, как число вершин y многогранника. Полагаем

$$z_1 = x_1,\ldots,z_r = x_r,\quad z_{r+1} = y_1,\quad z_{r+2} = -y_1,\ldots,z_{r+2(n-p)-1} = y_{n-p},\quad z_{r+2(n-p)} = -y_{n-p},$$

т.е. определяем систему $\{z_j\}$, $j = 1,\ldots,k$, $k = r + 2(n-p)$.

Теорема 5 (обобщенная лемма Фаркаша). *Пусть выполнено* $\langle \xi_i, z_j \rangle \geq 0$, $i = 1,\ldots,m$, $\Rightarrow \langle \eta, z_j \rangle \geq 0$, $j = 1,\ldots,k$, *где* k *определено выше. Тогда*

$$\eta = \sum_{i=1}^{m} \alpha_i \xi_i. \tag{11}$$

Доказательство. Очевидно, что $\eta \in \operatorname{span}\{\xi_i \,|\, i = 1,\ldots,m\}$. Покажем, что $\eta = \sum_{i=1}^{m} \alpha_i \xi_i$, $\alpha_i \geq 0$, $i = 1,\ldots,m$. Действительно, $\forall x \langle \xi_i, x \rangle \geq 0$, $i = 1,\ldots,m \;\Rightarrow\; x \in \mathbb{R}^n \cap \mathbb{R}^p \;\Rightarrow\; x \in \mathcal{K}_p$ и, значит, $x = \beta_1 x_1 + \ldots + \beta_r x_r$, $z_1 = x_1,\ldots,z_r = x_r$, $\beta_j \geq 0$, $j = 1,\ldots,r$. Но $\langle \eta, x_1 \rangle \geq 0,\ldots,\langle \eta, x_r \rangle \geq 0 \;\Rightarrow$ $\Rightarrow \langle \eta, \beta_1 x_1 + \ldots + \beta_r x_r \rangle = \langle \eta, x \rangle \geq 0 \Rightarrow$ выполнены условия теоремы Фаркаша и соотношение (11).

5. О СВОДИМОСТИ ЗАДАЧИ ОПТИМИЗАЦИИ С ОГРАНИЧЕНИЯМИ НЕРАВЕНСТВАМИ К ЗАДАЧЕ ОПТИМИЗАЦИИ С РАВЕНСТВАМИ

При решении оптимизационных задач с ограничениями неравенствами типа (3), (4) привлекательным является сведение ограничений неравенств к ограничениям равенствам путем добавления искусственных переменных. Однако, чтобы условие оптимальности для задачи с неравенствами и задачи с равенствами были эквивалентыми, необходимо вводить условие строгой дополняющей нежесткости (УСДН) [4]. В противном случае условия оптимальности будут не эквивалентны, и поэтому такая редукция, как сказано в [2], не представляет интереса. Поясним это. Введем функцию Лагранжа для задачи (3), (4)

$$L(x,\lambda) = \varphi(x) + \sum_{i=1}^{m} \lambda_i g_i(x).$$

Тогда необходимые условия минимума Куна–Таккера выглядят следующим образом:

$$L'_x(x,\lambda) = \varphi'(x) + \sum_{i=1}^{m} \lambda_i g'_i(x) = 0 \tag{12}$$

плюс условие дополняющей нежесткости

$$\lambda_i g_i(x) = 0, \quad i = 1,\ldots,m, \tag{13}$$

$$\lambda_i \geq 0, \quad i = 1,\ldots,m. \tag{14}$$

Достаточные условия для задачи (3), (4) будут следующими:

– если в допустимой точке x^* $(g_i(x^*) \leq 0, i = 1,\ldots,m)$ выполнены необходимые условия (12)–(14) при некотором λ^* и

$$\left\langle L''_{xx}(x^*\lambda^*)h, h\right\rangle > 0, \tag{15}$$

$\forall h \neq 0$ такого, что $\left\langle g'_i(x^*), h\right\rangle \leq 0$, $i \in I(x^*) = \left\{i \in \{1,\ldots,m\} \,|\, g_i(x^*) = 0\right\}$, то x^* – точка локального минимума задачи (3), (4) (см., например, [5]).

Существенным недостатком этих условий от достаточных условий оптимальности для задачи с равенствами является то, что точка x^* должна быть допустима. Уравнения (12)–(14) не гарантируют этого, в отличие от необходимых условий для задачи с равенствами. Поэтому модифицируем задачу (3), (4) в виде задачи с равенствами путем введения новой искусственной переменной $s = (s_1,\ldots,s_m)^\top$.

$$\begin{gathered} \min \varphi(x) \\ g_i(x) + s_i^2 = 0, \quad i = 1,\ldots,m. \end{gathered} \tag{16}$$

Задачи (3), (4) и (16) эквивалентны [6]. Необходимые условия оптимальности для задачи (16) выглядят следующим образом:

$$\varphi'(x) + \sum_{i=1}^{m} \lambda_i g'_i(x) = 0_n, \tag{17}$$

$$\lambda_1 \begin{pmatrix} 2s_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \ldots + \lambda_m \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 2s_m \end{pmatrix} = 0_m = 2\lambda s, \tag{18}$$

$$g_i(x) + s_i^2 = 0, \quad i = 1,\ldots,m, \tag{19}$$

где $\lambda s = (\lambda_1 s_1, \ldots, \lambda_m s_m)^\top$. Заметим, что из систем (17)–(19) нельзя определить знак величины $\lambda_i^* \geq 0$, $i = 1, \ldots, m$. Но уже не требуется допустимости решения x^* – она следует из системы (19). Обозначим $\bar{g}_i(x, s) = g_i(x) + s_i^2$, $i = 1, \ldots, m$, и введем функцию Лагранжа для задачи (16)

$$\bar{L}(x, s, \lambda) = \varphi(x) + \sum_{i=1}^{m} \lambda_i (g_i(x) + s_i^2). \tag{20}$$

Пусть $g(x) = (g_1(x), \ldots, g_m(x))^\top$, $s^2 = (s_1^2, \ldots, s_m^2)^\top$. Считаем, что в решении x^* выполнено условие регулярности ограничений (УРО), если векторы $\{g_i'(x^*)\}$, $i \in I(x^*)$, линейно независимы. Тогда достаточные условия оптимальности означают положительную определенность матрицы $\bar{L}''_{(x,s)}(x^*, s^*, \lambda^*)$ на ядре $\operatorname{Ker} \bar{g}'(x^*, s^*)$, т.е.

$$\left\langle \bar{L}''_{(x,s)}(x^*, \lambda^*, s^*)h, h \right\rangle \geq \alpha \|h\|^2, \quad \alpha > 0, \tag{21}$$

для любых $h \in \operatorname{Ker} \bar{g}(x^*, s^*)$. Отсюда, кстати, следует неотрицательность $\lambda_i^* \geq 0$, $i = 1, \ldots, m$. Запишем, для наглядности, что

$$\bar{L}''_{(x,s)}(x^*, \lambda^*, s^*) = \begin{pmatrix} L_{xx}(x^*, \lambda^*) & 0_{nm} \\ 0_{mn} & 2\lambda^* \end{pmatrix}.$$

Однако матрица $\bar{L}''(x^*, s^*, \lambda^*)$ не будет положительно определена на ядре $\bar{g}'(x^*, s^*)$, если не выполнено условие строгой дополняющей нежесткости (УСДН), т.е. λ_i^* и s_i^* не равны нулю одновременно при $i = 1, \ldots, m$. То есть достаточные условия (15) и (21) не равноценны. Это связано с тем, что можно ослабить необходимые условия оптимальности 2-го порядка для задачи с равенствами, а значит, и построить более слабые достаточные условия. Поясним следующее. Пусть существует такое $h_2 \in \operatorname{Ker} F'(x^*)$ для задачи (1), (2), что

$$\left\langle L''_{xx}(x^*, \lambda^*)h_2, h_2 \right\rangle = 0 \quad \left(L(x, \lambda) = \varphi(x) + \sum_{i=1}^{m} \lambda_i f_i(x) \right) \tag{22}$$

и, значит, достаточные условия 2-го порядка не выполнены. Пусть $\operatorname{Ker} F'(x^*) = \mathbb{R}_1^n$ и $\mathbb{R}_2^n = (\mathbb{R}_1^n)^\perp$. Тогда $\mathbb{R}^n = \mathbb{R}_1^n + \mathbb{R}_2^n$. Будет верна следующая

Теорема 6 (необходимые и достаточные условия оптимальности специального вида). *Пусть* $\varphi, F \in C^3(\mathbb{R}^n)$ *и для любого* $h_2 \in \operatorname{Ker} F'(x^*)$ $(\|h\| = 1)$ *такого, что* $L''(x^*\lambda^*)[h_2]^2 = 0$ *или* $(L''(x^*\lambda^*)h_2 = 0)$ *выполняются условия*

1)

$$\left\| F''(x^*)[h_2]^2 \right\| \geq \alpha > 0 \tag{23}$$

2) *и существуют* $h_1 \in \mathbb{R}_2^r$ и $C(h_2) \geq c > 0$, $\|h_1\| = 1$, *удовлетворяющие соотношению*

$$F'(x^*)[h_1] + \frac{1}{2} F''(x^*)[c(h_2)h_2]^2 = 0. \tag{24}$$

Тогда (необходимость), если x^* – решение (1), (2), *то* $\exists \lambda^*$ *такое, что* $L'(x^*, \lambda^*) = 0$ *и*

$$L''(x^*, \lambda^*)[h_1]^2 \geq 0. \tag{25}$$

Более того (достаточность), если $\forall h_1, h_2$, *удовлетворяющих* (23), (24), *существует* $\beta > 0$ *такое, что* $L'(x^*, \lambda^*) = 0$ *и*

$$L''_{xx}(x^*, \lambda^*)[h_1]^2 \geq \beta > 0, \tag{26}$$

то x^* – локальный минимум задачи (1), (2).

Замечание 2. Элемент h_1 может не принадлежать $\operatorname{Ker} F'(x^*)$.

Доказательство. *Необходимость.* Покажем, что $\forall t \in (0, \varepsilon)$, $\varepsilon > 0$ достаточно малое, существует $\omega(t)\mathbb{R} \to \mathbb{R}^n$ такое, что дуга

$$\gamma(t) = x^* + c_2 t h_2 + t^2 h_1 + \omega(t) \in M(x^*) = \{x \in \mathbb{R}^n \,|\, F(x) = F(x^*) = 0\}$$

и $\|\omega(t)\| = o(t^2)$. Действительно,

$$\begin{aligned} F(x^* + c_2 t h_2 + t^2 h_1) &= F(x^*) + F'(x^*)[c_2 t h_2 + t^2 h_1] + \frac{1}{2} F''(x^*)[c_2 t h_2 + t^2 h_1]^2 + o(t^2) = \\ &= t^2 F'(x^*) h_1 + \frac{1}{2} F''(x^*)[c_2 t h_2]^2 + o(t^2) = o(t^2). \end{aligned} \tag{27}$$

В силу невырожденности $F'(x^*)$ можем применить принцип сжимающих отображений [1] к отображению

$$\Phi_t(x) = x - \left(F'(x^*)\right)^{-1} F(x^* + c_2 t h_2 + t^2 h_1 + x), \tag{28}$$

из которого следует существование $\omega(t)$ такого, что

$$F(x^* + c_2 t h_2 + t^2 h_1 + \omega(t)) = 0 \tag{29}$$

и

$$\|\omega(t)\| = o(t^2). \tag{30}$$

Далее применим стандартную технику к соотношению $L(x^* + c_2 t h_2 + t^2 h_1 + \omega(t), \lambda^*) \geq L(x^*, \lambda^*)$ получим $L''_{xx}(x^*, \lambda^*)[h_1]^2 \geq 0$. Необходимость доказана.

Достаточность. Предполагая противное, т.е. что существует последовательность точек $\{x_k\} \in M(x^*)$ и $x_k \to x^*$ при $k \to \infty$ такая, что $\varphi(x_k) < \varphi(x^*)$, используя доказательство необходимости и стандартную технику доказательства достаточности условий, приходим к противоречию с соотношением (26).

Пример 1. Рассмотрим задачу (16) в виде

$$\min \varphi(y) \tag{31}$$

при $F(x) = g(y) + s^2 = 0_m$. Здесь роль x играет (y, s) (чтобы не было путаницы с последними записями, мы заменили в задаче (3), (4) x на y). Если точка (y^*, s^*), $s^* = 0$ – решение задачи (3), (4), то для $h_2 = (0, \bar{s})^\top$ очевидно существует $h_1 = (\bar{y}, 0)^\top$ такое, что

$$F'(x^*)[h_1] + \frac{1}{2} F''(x^*)[h_2]^2 = 0 \Leftrightarrow g'(y^*)\bar{y} + \frac{1}{2} 2\bar{s}^2 = 0. \tag{32}$$

Поскольку условия оптимальности носят локальный характер, мы рассмотрим только активные ограничения, т.е. $i \in I(x^*)$ и, не нарушая общности, считаем $r = m$. Очевидно, что условие (23) выполнено, если $\bar{s} \neq 0$. А условие (24) выполнено в силу регулярности ограничений $\{g_i(x)\}$, $i = 1, \ldots, m$, в точке x^*, при этом $c_2 = 1$. Таким образом, к задаче (16) применима теорема 6, так как $L''_{xx}(x^*, \lambda^*)[h_2]^2 = 0$, хотя $h_2 = (0, \bar{s})^\top \in \operatorname{Ker} F'(x^*) = \{(\bar{y}, \bar{s}) \,|\, g'(y^*)\bar{y} + 0 \cdot \bar{s} = 0\}$. Более того будет справедлива следующая теорема об эквивалентности.

Теорема 7 (об эквивалентности условий оптимальности). *Пусть* $\varphi, g \in C^3(\mathbb{R}^n)$. *Тогда условия оптимальности* (12), (13), (15) *и* (26) *эквивалентны.*

Доказательство. Очевидно нужно проверить только достаточную часть условий оптимальности. Пусть выполнены условия (12)–(15). Тогда, беря $\bar{s} \neq 0$, получаем $\bar{y} = \{g'(y^*)\}^{-1}(-\bar{s}^2)$ из условия (24). При этом $L''_{xx}(y^*, s^* \lambda^*)[y]^2 \geq \beta \|\bar{y}\|^2$, так как выполнено достаточное условие (15). Здесь $\bar{y}$ играет роль h_1 в теореме 6 и поэтому имеет место (26), а значит, выполнены достаточные условия теоремы 6. В одну сторону теорема доказана.

Пусть теперь выполнены достаточные условия (26) для задачи (31). Тогда $\forall h$ такого, что $g'(y^*)h < 0_m$ будет $\bar{s} \neq 0$ и, значит, выполнено (26) при $h_1 = h$, т.е. $L''_{xx}(y^*, 0, \lambda^*)[h]^2 \geq \beta \|h\|^2$, а это значит, что $L''_{xx}(y^*,\lambda^*)[h]^2 \geq \beta\|h\|^2$ (где $L(y,\lambda)$ играет роль функции $L(x,\lambda) = \varphi(x) + \sum_{i=1}^{m} \lambda_i g_i(x)$ в задаче (3), (4)), и окончательно выполнены достаточные условия (15) для задачи (3), (4). Теорема доказана.

6. 2-ФАКТОР-МЕТОД РЕШЕНИЯ ВЫРОЖДЕННЫХ СИСТЕМ НЕЛИНЕЙНЫХ УРАВНЕНИЙ

Для численного решения систем уравнений, в частности возникающих из необходимых условий оптимальности, основной конструкцией для построения различного рода итерационных процессов является метод Ньютона [4]. Однако если для системы нелинейных уравнений

$$\Phi(x) = 0_n \tag{33}$$

в решении x^* матрица $\Phi'(x^*)$ вырождена, то применение метода Ньютона, итерационная схема которого имеет вид

$$x_{k+1} = x_k - \{\Phi'(x_k)\}^{-1}\Phi(x_k), \quad k = 0,1,\ldots, \tag{34}$$

становится неэффективным или вообще невозможным. В этом случае применяется так называемый 2-фактор-метод, обеспечивающий сходимость, и, более того, квадратичную скорость сходимости. В общем случае, схема 2-фактор-метода решения вырожденной системы нелинейных уравнений (33) при $x \in \mathbb{R}^n$ имеет следующий вид:

$$x_{k+1} = x_k - \{\Phi'(x_k) + P\Phi''(x_k)h\}^{-1}(\Phi(x_k) + P\Phi'(x_k)h), \quad k = 0,1,\ldots, \tag{35}$$

где P – матрица ортогонального проектирования на $(\operatorname{Im}\Phi'(x^*))^{\perp}$, а вектор h, $\|h\| = 1$ – некоторый фиксированный вектор такой, что матрица $\{\Phi'(x^*) + P\Phi''(x^*)h\}$ невырожденна, т.е. отображение $\Phi(\cdot)$ является 2-регулярным на векторе h в точке решения x^* системы (33). В этом случае справедлива следующая

Теорема 8. *Пусть* $\Phi \in C^3(\mathbb{R}^n)$ *и* $\Phi(\cdot)$ 2-*регулярно в решении* x^* *на векторе* h. *Тогда для* $x_0 \in U_\varepsilon(x^*)$ *последовательность* $\{x_k\}$, *определяемая схемой* (35), *сходится к решению* x^*, *причем справедлива оценка скорости сходимости*

$$\|x_{k+1} - x^*\| = c\|x_k - x^*\|^2, \quad k = 0,1,\ldots, \tag{36}$$

где $\varepsilon > 0$ – *достаточно малое*.

Доказательство см., например, в [7]. Покажем, как 2-фактор-метод можно применять для численного решения задач оптимизации с ограничениями неравенствами при использовании искусственных переменных.

7. ПОНЯТИЕ 2-РЕГУЛЯРНОСТИ И 2-ФАКТОР-МЕТОДА НЬЮТОНА ДЛЯ РЕШЕНИЯ ВЫРОЖДЕННОЙ СИСТЕМЫ КУНА–ТАККЕРА

При решении оптимизационных задач с ограничениями типа неравенства путем введения искусственных переменных необходимо вводить условия строго дополняющей нежесткости (УСДН) для того, чтобы система КТ была невырожденной. В противном случае нельзя гарантировать применимость высокоэффективных методов типа Ньютона. В данном разделе мы рассмотрим эффективность указанного подхода с точки зрения применения 2-фактор-метода для решения системы вырожденных условий оптимальности и докажем, что система КТ 2-регулярна.

Итак, имеем задачу (3), (4)

$$\min \varphi(x)$$

при

$$g_i(x) \leq 0, \ldots 1 = 1,\ldots,m,$$

которую заменяем на задачу (16)

$$\min \varphi(x)$$

$$g_i(x) + s_i^2 = 0, \ldots 1 = 1,\ldots,m.$$

Введем функцию Лагранжа

$$L(x,s,\lambda) = \varphi(x) + \left\langle g(x) + s^2, \lambda \right\rangle, \quad \lambda \in \mathbb{R}^m,$$

и обозначим $u = (x, s, \lambda)$. Тогда имеем

$$L'(u) = \begin{pmatrix} L_x(u) \\ 2\lambda s \\ g(x) + s^2 \end{pmatrix}, \tag{37}$$

$$L''(u) = \begin{bmatrix} L_{xx}(u) & 0_{nm} & (g_x(x))^\top \\ 0_{mn} & 2\lambda & 2s \\ g_x(x) & 2s & 0_{mm} \end{bmatrix}. \tag{38}$$

Введем систему из ℓ нелинейных уравнений, где $\ell = n + 2m$

$$L'(u) = 0_\ell. \tag{39}$$

Пусть $u^* = (x^*, s^*, \lambda^*)^\top \in \mathbb{R}^\ell$ ее решение. Если $\lambda^* \geq 0_m$, $L'(u^*) = 0_\ell$, то в точке (x^*, λ^*) выполнены необходимые условия минимума для задачи (3), (4). Если при этом добавить достаточные условия (15), то точка x^* будет локальным решением задачи (3), (4). Однако наличие неравенства $\lambda^* \geq 0_m$ не позволяет получить полную систему равенств. Поэтому заменим $\lambda^* \geq 0_m$ на систему равенств

$$\lambda^\top - y^2 = 0_m, \tag{40}$$

где $y \in \mathbb{R}^m$, $y^2 = (y_1^2, \ldots, y_m^2)^\top$ и рассмотрим расширенную систему

$$\Psi(z) = \begin{pmatrix} L'(u) = 0_\ell \\ \lambda^\top - y^2 = 0_m \end{pmatrix} = 0_r, \tag{41}$$

где $z = (x, s, \lambda, y)^\top$, $r = n + 3m$. Таким образом, мы задачу оптимизации с неравенствами свели к системе равенств. Причем в решении (x^*, λ^*) выполнены необходимые условия КТ, и теперь $\lambda^* \geq 0$. Пусть

$$\Psi'(z) = \begin{pmatrix} L_{xx}(u) & 0_{nm} & (g_x(x))^\top & 0_{nm} \\ 0_{mn} & 2\lambda & 2s & 0_{mm} \\ g_x(x) & 2s & 0_{mm} & 0_{mm} \\ 0_{mn} & I_m & 0_{mm} & 2y \end{pmatrix}. \tag{42}$$

Здесь I_m – единичная матрица размерности $m \times m$. Метод Ньютона для решения системы (42) имеет вид

$$z_{k+1} = z_k - \left(\Psi'(z_k)\right)^{-1} \Psi(z_k). \tag{43}$$

Если в решении z^* не выполнено УСДН, то матрица $\Psi'(z^*)$ вырождена и метод (43) не гарантирует сходимость. Построим так называемый 2-фактор-метод Ньютона для решения системы (41), обеспечивающий квадратичную скорость сходимости. Следуя [7], [8], введем отображение

$$\Phi(z) = \Psi(z) + P\Psi'(z)h \tag{44}$$

и 2-фактор-оператор

$$\Phi'(z) = \Psi'(z) + P\Psi''(z)h, \tag{45}$$

где $r \times r$ матрица P проектирует вектор $\Psi''(z^*)h$ на $\operatorname{Im} \Psi'(z^*)^\perp$, r-мерный вектор h, имеющий единичную норму, выбирается из условия невырожденности $\Phi'(z^*)$. Отметим, что матрицу P можно выбирать, вообще говоря, просто из условия невырожденности 2-фактор-оператора $\Phi'(z^*)$.

Теорема 9. *Пусть* $\varphi, g \in C^3(\mathbb{R}^n)$, *существует удовлетворяющий системе* (41) *вектор* z^*, *в точке* (x^*, λ^*) *выполнены достаточные условия оптимальности* (15) *и УРО. Тогда* 2*-фактор-матрица* $\Phi'(z^*)$ *невырожденна*.

Доказательство. Покажем, что отображение $\Psi(z)$ будет 2-регулярно в точке z^* на векторе $h^\top = (0_n^\top, 0_m^\top, e_0^\top, e_1^\top)$, где $e_0 \in \mathbb{R}^m$ и каждая i-я компонента вектора e_0 равна единице, если $i \in I_0(x^*) = \{i \in I(x^*) | \lambda_i^* = 0, s_i^* = 0\}$, и равно нулю в противном случае, а $e_1 \in \mathbb{R}^m$ и каждая i-я компонента вектора e_1 равна единице, если $\lambda_i^* = 0$, и равна нулю в противном случае. Определим матрицу

$$\Phi'(z^*) = \Psi'(z^*) + P\Psi''(z^*)h = \begin{pmatrix} L_{xx}(z^*) & 0_{nm} & (g_x(x^*))^\top & 0_{nm} \\ 0_{mn} & 2\lambda^* + 2J_0(x^*) & 2s^* & 0_{mm} \\ (g_x(x^*))^\top & 2s^* & 0_{mm} & 0_{mm} \\ 0_{mn} & I_m & 0_{mm} & 2y^* + 2J_1(x^*) \end{pmatrix}, \qquad (46)$$

где $r \times r$-диагональная матрица P имеет i-й диагональный элемент, равный нулю, если $1 \le i \le k$ или $k + m + 1 \le i \le \ell$, а при $k + 1 \le i \le n + m$, i-й диагональный элемент равен единице, если $i - n \in I_0(x^*)$, и равен нулю в противном случае. При $\ell + 1 \le i \le r$, i-й диагональный элемент равен единице, если $i - \ell \in I_0(x^*)$, и равен нулю в противном случае. При этом $m \times m$-диагональная матрица $J_0(x^*)$ имеет j-й диагональный элемент, равный либо нулю, либо единице, если $j \in I_0(x^*)$. В свою очередь $m \times m$-диагональная матрица $J_1(x^*)$ имеет j-й диагональный элемент, равный единице, если $j \in I_0(x^*)$, и нулю в противном случае. Отсюда следует, что при $h = (0^\top, 0^\top, e_0^\top, e_1^\top)^\top$ получаем

$$P\Psi''(z^*)h = \begin{pmatrix} 0_{nn} & 0_{nm} & 0_{nm} & 0_{nm} \\ 0_{mn} & 2J_0(x^*) & 0_{nm} & 0_{mm} \\ 0_{mn} & 0_{mm} & 0_{mm} & 0_{mm} \\ 0_{mn} & 0_{mn} & 0_{mm} & 2J_1(x^*) \end{pmatrix}$$

и справедливость формулы (46), а также и невырожденность матрицы $\Phi'(z^*)$.

Очевидно, что 2-фактор-метод, примененный к системе (41), имеет вид

$$z_{k+1} = z_k - \left(\Psi'(z_k) + P\Psi''(z_k)h\right)^{-1}\left(\Psi(z_k) + P\Psi'(z_k)h\right), \quad k = 0,1,\ldots. \qquad (47)$$

Теорема 10. *Пусть* $\Psi \in C^3(\mathbb{R}^r)$ *и для* $h \in \mathbb{R}^r$, $\|h\| \neq 0$, *существует* $\left(\Psi'(z^*) + P\Psi''(z^*h)\right)^{-1}$. *Тогда* 2-фактор-метод (47) *сходится к решению* z^* системы (41) *с квадратичной скоростью*, *т.е.*

$$\|z_{k+1} - z^*\| \le c\|z_k - z^*\|^2, \quad k = 0,1,\ldots,$$

где $z_0 \in U_\varepsilon(z^*)$, $\varepsilon > 0$ — *достаточно малое, а* $c > 0$ *независимая константа.*

Доказательство. Следует из того факта, что 2-фактор-метод (47) есть обычный метод Ньютона, примененный к системе

$$\Psi(z) + P\Psi'(z)h = 0, \qquad (48)$$

так как точка z^* — решение системы (41), будет одновременно решением системы (48) и для нее выполнены все условия сходимости метода Ньютона с квадратичной скоростью.

СПИСОК ЛИТЕРАТУРЫ

1. *Иоффе А.Д., Тихомиров В.М.* Теория экстремальных задач. М.: Наука, 1974.
2. *Bertsekas D.P.* Nonlinear programming. Belmont: Athena scientific, 1999. P. 191–276.
3. *Ашманов С.А., Тимохов А.В.* Теория оптимизации в задачах и упражнениях. М.: Физматлит, 1991. 448 с.
4. *Евтушенко Ю.Г.* Методы решения экстремальных задач и их применение в системах оптимизации. М.: Наука, 1982. 432 с.
5. *Поляк Б.Т.* Введение в оптимизацию. М.: Физматлит, 1983.
6. *Евтушенко Ю.Г., Третьяков А.А.* Аппроксимация p-го порядка множества решений нелинейных уравнений // Ж. вычисл. матем. и матем. физ. 2013. Т. 53. № 12. С. 1951.
7. *Белаш К.Н., Третьяков А.А.* Методы решения вырожденных задач // Ж. вычисл. матем. и матем. физ. 1988. Т. 28. № 7. С. 1097–1102.
8. *Brezhneva O.A., Tret'yakov A.A.* The p-factor-Lagrange methods for degenerate nonlinear programming // Numerical Functional Analysis and Optimization. 2007. V. 28. № 9–10. P. 1051–1086.