ЖУРНАЛ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ ФИЗИКИ, 2003, том 43, № 3, с. 354–375

УДК 519.853.65

# ТЕОРЕМЫ ОБ АЛЬТЕРНАТИВАХ И ИХ ПРИМЕНЕНИЕ В ЧИСЛЕННЫХ МЕТОДАХ[1)]

© 2003 г. А. И. Голиков, Ю. Г. Евтушенко

(119991 Москва, ул. Вавилова, 40, ВЦ РАН)
e-mail: gol@ccas.ru, evt@ccas.ru

Поступила в редакцию 20.09.02 г.

Для систем линейных равенств и неравенств доказаны новые теоремы об альтернативах, которые дали возможность построить ряд эффективных численных методов. Эти методы использованы для нахождения нормальных решений линейных систем равенств и неравенств, для построения разделяющих гиперплоскостей, для коррекции несовместных систем, для решения задач нелинейного программирования (НЛП), существенно упрощая реализацию метода наискорейшего спуска. Из приведенных теорем как частный случай следуют теоремы об альтернативах Фредгольма, Фаркаша, Гейла, Жордана, Штимке и др. Библ. 21.

## 1. ВВЕДЕНИЕ

Существуют различные варианты теорем об альтернативах (см., например, [1]–[10]), применявшиеся в основном для доказательства теорем существования, вывода условий экстремума в задачах оптимизации. В данной работе приводятся конструктивные доказательства новых теорем об альтернативах, ориентированные на построение новых вычислительных методов. Эти доказательства дают возможность вычислять нормальные решения систем линейных равенств и неравенств, находить направление наискорейшего спуска в задачах НЛП, строить разделяющие гиперплоскости, производить коррекцию несобственных задач, конструировать новые алгоритмы решения задач линейного программирования и т.д.

Для заданной линейной системы строится альтернативная система, размерность переменных которой равна числу равенств и неравенств в исходной системе (исключая ограничения на знак переменных). Метод решения исходной разрешимой системы состоит в минимизации невязки альтернативной несовместной системы. По результатам этой минимизации определяется нормальное решение исходной системы (решение с минимальной евклидовой нормой). Благодаря различию размерностей переменных альтернативных систем переход от исходной совместной системы к минимизации невязки альтернативной несовместной системы может оказаться весьма целесообразным. Эта редукция может привести к задаче минимизации по переменным меньшей размерности и дает возможность получить нормальное решение исходной системы.

Из приведенных в работе результатов как частный случай следуют многие известные теоремы об альтернативах (Фредгольма, Фаркаша, Гейла, Жордана, Штимке и др.). Но авторы считают, что главным в статье являются результаты, открывающие новые вычислительные возможности применения теорем об альтернативах. С их помощью в разд. 2 предложен новый метод решения систем линейных равенств и неравенств, находящий нормальное решение, в разд. 3 существенно упрощены расчеты в методе наискорейшего спуска, в разд. 4 получены простые формулы коррекции несобственных систем, в разд. 7 найдены расчетные формулы построения семейства разделяющих гиперплоскостей, а также построены новые методы решения задач линейного программирования (см. [17]).

Приведенные в этой статье доказательства теорем существенно опираются на теорию двойственности. Данная работа является дальнейшим развитием результатов авторов (см. [16]–[19]). Доказательство теорем об альтернативах, основанное на минимизации невязок систем, впервые было предложено в [5]. Отметим также работы [8] и [20], [21], в которых теоремы об альтернативах применяются для построения численных методов.

[1)]Работа выполнена при финансовой поддержке РФФИ (код проекта 01-01-00804), по программе государственной поддержки ведущих научных школ (код проекта 00-15-96080) и при частичной финансовой поддержке Программы Президиума РАН "Математическое моделирование".

## 2. ОСНОВНЫЕ ТЕОРЕМЫ

Пусть $A$ (матрица $m \times n$) задана в следующем виде:

$$A = \begin{bmatrix} A_{11} & A_{22} \\ A_{21} & A_{22} \end{bmatrix}.$$

Здесь прямоугольные матрицы $A_{11}$, $A_{12}$, $A_{21}$, $A_{22}$ имеют, соответственно, размеры: $m_1 \times n_1$, $m_1 \times n_2$, $m_2 \times n_1$, $m_2 \times n_2$. Пусть векторы $x \in \mathbb{R}^n$, $z$, $u$, $b \in \mathbb{R}^m$ имеют разбиение $x^{\mathrm{T}} = [x_1^{\mathrm{T}}, x_2^{\mathrm{T}}]$, $z^{\mathrm{T}} = [z_1^{\mathrm{T}}, z_2^{\mathrm{T}}]$, $u^{\mathrm{T}} = [u_1^{\mathrm{T}}, u_2^{\mathrm{T}}]$, $b^{\mathrm{T}} = [b_1^{\mathrm{T}}, b_2^{\mathrm{T}}]$, где $x_1 \in \mathbb{R}^{n_1}$, $x_2 \in \mathbb{R}^{n_2}$, $n = n_1 + n_2$, $z_1$, $u_1$, $b_1 \in \mathbb{R}^{m_1}$, $z_2$, $u_2$, $b_2 \in \mathbb{R}^{m_2}$, $m = m_1 + m_2$. Предполагается, что матрица $A$ и вектор $b$ ненулевые.

Введем вспомогательные множества $\Pi_x = \{[x_1, x_2] : x_1 \in \mathbb{R}_+^{n_1}, x_2 \in \mathbb{R}^{n_2}\}$, $\Pi_z = \{[z_1, z_2] : z_1 \in \mathbb{R}_+^{m_1}, z_2 \in \mathbb{R}^{m_2}\}$, $\Pi_u = \{[u_1, u_2] : u_1 \in \mathbb{R}_+^{m_1}, u_2 \in \mathbb{R}^{m_2}\}$. Определим вектор $w \in \mathbb{R}^{n+1}$, представимый в виде $w^{\mathrm{T}} = [w_1^{\mathrm{T}}, w_2^{\mathrm{T}}, w_3]$, где $w_1 \in \mathbb{R}^{n_1}$, $w_2 \in \mathbb{R}^{n_2}$, $w_3 \in \mathbb{R}^1$, и вспомогательное множество $\Pi_w = \{[w_1, w_2, w_3] : w_1 \in \mathbb{R}_+^{n_1}, w_2 \in \mathbb{R}^{n_2}, w_3 \in \mathbb{R}^1\}$. Выражение $\|a\|$ обозначает евклидову норму вектора $a$, а выражение $\|a\|_1 = \Sigma_{i=1}^n |a^i|$ – первая норма вектора $a \in \mathbb{R}^n$.

Рассмотрим систему линейных равенств и неравенств

$$A_{11}x_1 + A_{12}x_2 \ge b_1, \quad A_{21}x_1 + A_{22}x_2 = b_2, \quad x_1 \ge 0_{n_1}. \tag{I}$$

Эту систему можно рассматривать как допустимое множество в задаче линейного программирования о минимизации $c^{\mathrm{T}}x$, где $c = 0_n$. Для этой задачи выпишем двойственную, в которой ищется максимум $b^{\mathrm{T}}u$ на допустимом множестве, определяемом линейной системой равенств и неравенств

$$A_{11}^{\mathrm{T}}z_1 + A_{21}^{\mathrm{T}}z_2 \le 0_{n_1}, \quad A_{12}^{\mathrm{T}}z_1 + A_{22}^{\mathrm{T}}z_2 = 0_{n_2}, \quad z_1 \ge 0_{m_1}. \tag{I)'$$

Систему (I)' назовем *сопряженной* к системе (I).

Зададим систему

$$A_{11}^{\mathrm{T}}u_1 + A_{21}^{\mathrm{T}}u_2 \le 0_{n_1}, \quad A_{12}^{\mathrm{T}}u_1 + A_{22}^{\mathrm{T}}u_2 = 0_{n_2}, \quad b_1^{\mathrm{T}}u_1 + b_2^{\mathrm{T}}u_2 = \rho, \quad u_1 \ge 0_{m_1}. \tag{II}$$

Здесь $\rho > 0$ – произвольное фиксированное положительное число. Условие $b_1^{\mathrm{T}}u_1 + b_2^{\mathrm{T}}u_2 = \rho$ исключает тривиальные решения сопряженной системы (I)'.

Для системы (II) сопряженная система имеет вид

$$A_{11}w_1 + A_{12}w_2 - b_1w_3 \ge 0_{m_1}, \quad A_{21}w_1 + A_{22}w_2 - b_2w_3 = 0_{m_2}, \quad w_1 \ge 0_{n_1}. \tag{II)'$$

Множества решений систем (I), (I)', (II) и (II)' обозначим, соответственно, через $X$, $Z$, $U$ и $W$. Если системы (I) или (II) разрешимы, то будем, соответственно, писать $X \ne \varnothing$ или $U \ne \varnothing$. В отличие от (I) и (II), сопряженные системы (I)' и (II)' всегда имеют решения, так как $0_m \in Z$ и $0_{n+1} \in W$. Из вида системы (II) следует, что если она разрешима или не разрешима при некотором $\rho = \rho_1 > 0$, то это свойство сохранится и при любом другом $\rho = \rho_2 > 0$.

Через $\mathrm{pen}(x, X)$ обозначим штраф, вычисленный в точке $x \in \Pi_x$ за нарушение условия $x \in X$. По определению штрафной функции, если $x \in \Pi_x$, то $\mathrm{pen}(x, X) = 0$ тогда и только тогда, когда $x \in X$. Аналогично введем обозначение $\mathrm{pen}(u, U)$. Если $u \in \Pi_u$, то $\mathrm{pen}(u, U) = 0$ тогда и только тогда, когда $u \in U$.

В качестве штрафов будем использовать евклидову норму векторов невязок систем (I) и (II), т.е.

$$\mathrm{pen}(x, X) = \left[\|(b_1 - A_{11}x_1 - A_{12}x_2)_+\|^2 + \|b_2 - A_{21}x_1 - A_{22}x_2\|^2\right]^{1/2},$$

$$\mathrm{pen}(u, U) = \left[\|(A_{11}^{\mathrm{T}}u_1 + A_{21}^{\mathrm{T}}u_2)_+\|^2 + \|(A_{12}^{\mathrm{T}}u_1 + A_{22}^{\mathrm{T}}u_2)\|^2 + (\rho - b_1^{\mathrm{T}}u_1 - b_2^{\mathrm{T}}u_2)^2\right]^{1/2}.$$

Здесь $a_+$ – неотрицательная часть вектора $a$, т.е. $i$-я компонента вектора $a_+$ совпадает с $i$-й компонентой вектора $a$, если она неотрицательна, и равна нулю в противном случае.

Для проверки разрешимости той или иной системы и нахождения соответствующего решения будем использовать методы безусловной минимизации, примененные к любой из следующих задач:

$$I_1 = \min_{x \in \Pi_x} [\mathrm{pen}(x, X)]^2/2, \tag{1}$$

$$I_2 = \min_{u \in \Pi_u} [\mathrm{pen}(u, U)]^2/2. \tag{2}$$

Строго говоря, (1) и (2) не являются задачами безусловной минимизации, так как в них присутствуют ограничения на знаки компонент векторов $x_1$ и $u_1$, однако поскольку большинство методов безусловной минимизации с помощью простейших модификаций могут учитывать наличие ограничений на знак переменных, то сохраним за (1) и (2) этот термин. Задачи (1) и (2) всегда разрешимы, так как квадратичные целевые функции ограничены снизу нулем на непустых допустимых множествах $\Pi_x$ и $\Pi_u$.

Две системы называются *альтернативными*, если разрешима одна и только одна из них. Разрешимость или неразрешимость системы можно характеризовать скалярной величиной – минимальной невязкой системы, полученной из решения задачи (1) или (2). Поэтому альтернативность систем можно определять через решение этих задач. Пусть $x^* \in \Pi_x$ и $u^* \in \Pi_u$ – произвольные решения задач (1) и (2) соответственно, т.е. $I_1 = [\mathrm{pen}(x^*, X)]^2/2$ и $I_2 = [\mathrm{pen}(u^*, U)]^2/2$. Тогда для систем (I) и (II) имеет место следующая

**Лемма 1** (критерий альтернативности). *Системы* (I) *и* (II) *альтернативны тогда и только тогда, когда выполнены следующие условия*:

$$\mathrm{pen}(x^*, X)\mathrm{pen}(u^*, U) = 0, \quad \mathrm{pen}(x^*, X) + \mathrm{pen}(u^*, U) > 0. \tag{3}$$

**Доказательство.** Из первого соотношения в (3) следует, что по крайней мере одна из систем (I) или (II) разрешима. Второе соотношение в (3) показывает, что по крайней мере одна из систем неразрешима. Вместе оба соотношения утверждают, что системы (I) и (II) альтернативны.

**Лемма 2.** *Системы* (I) *и* (II) *не могут быть одновременно разрешимы.*

**Доказательство.** Предположим обратное, пусть существуют $x^*$ – решение системы (I) и $u^*$ – решение системы (II). Тогда подставим в (I) решение $x^*$ и скалярно умножим первое неравенство в (I) на $u_1^*$, второе равенство – на $u_2^*$. После сложения результатов и простейших преобразований получим

$$x_1^{*\mathrm{T}}(A_{11}^{\mathrm{T}}u_1^* + A_{21}^{\mathrm{T}}u_2^*) + x_2^{*\mathrm{T}}(A_{12}^{\mathrm{T}}u_1^* + A_{22}^{\mathrm{T}}u_2^*) \ge b_1^{\mathrm{T}}u_1^* + b_2^{\mathrm{T}}u_2^*.$$

Согласно (II), левая часть этого неравенства неположительна, а правая часть строго положительна, так как $A_{11}^{\mathrm{T}}u_1^* + A_{21}^{\mathrm{T}}u_2^* \le 0_{n_1}$ и $b_1^{\mathrm{T}}u_1^* + b_2^{\mathrm{T}}u_2^* = \rho > 0$. Приходим к противоречию. Следовательно, системы (I), (II) не могут одновременно быть совместны, для них выполнено второе условие в (3). Лемма доказана.

Первое условие из (3) будет доказано ниже в теореме 3. Среди альтернативных к альтернативной системе (II) содержится исходная задача (I). Действительно, система, альтернативная к (II), имеет вид

$$A_{11}w_1 + A_{12}w_2 - b_1w_3 \ge 0_{m_1}, \quad A_{21}w_1 + A_{22}w_2 - b_2w_3 = 0_{m_2}, \quad \rho w_3 = \rho', \quad w_1 \ge 0_{n_1}, \tag{4}$$

где $\rho' > 0$ – произвольное положительное число. Поэтому $w_3 = \rho'/\rho > 0$. Сделав в (4) замену $x_1 = w_1/w_3$, $x_2 = w_2/w_3$, придем к исходной системе (I).

Введем следующие две задачи квадратичного программирования:

$$I_1^d = \max_{z \in Z}\{b^{\mathrm{T}}z - \|z\|^2/2\}, \tag{5}$$

$$I_2^d = \max_{w \in W}\{\rho w_3 - \|w\|^2/2\}. \tag{6}$$

В отличие от систем (I), (II), которые могут быть совместны или несовместны, задачи (1), (2), (5), (6) всегда имеют решения. Причем задачи (5) и (6) имеют единственные решения, так как в них допустимые множества $Z$ и $W$ непусты и строго вогнутые квадратичные целевые функции ограничены сверху (см. ниже формулу (32)).

Формально задачи безусловной минимизации (1) и (2) не имеют функцию Лагранжа и, следовательно, для них нельзя непосредственно построить двойственные задачи. Тем не менее с помощью введения дополнительных переменных можно построить искусственные ограничения и получить эквивалентные задачи нелинейного программирования, для которых определены двойственные задачи.

**Лемма 3.** *Двойственные к задачам* (5) *и* (6) *являются задачи безусловной минимизации* (1) *и* (2) *соответственно. Задачи* (1) *и* (2) *сводятся к эквивалентным задачам условной минимизации квадратичных функций, которые являются двойственными к* (5) *и* (6) *соответственно.*

**Доказательство.** Первое утверждение следует из традиционных представлений двойственных задач для задач квадратичного программирования (см., например, [6], [12]–[14]).

Второе утверждение несколько нетрадиционно и основано на двухшаговом представлении задач (1) и (2). Такой прием был использован в [14], [15].

Введем вектор дополнительных переменных $y \in \mathbb{R}^m$, $y^{\mathrm{T}} = [y_1^{\mathrm{T}}, y_2^{\mathrm{T}}]$, где $y_1 \in \mathbb{R}^{m_1}$, $y_2 \in \mathbb{R}^{m_2}$ таковы, что

$$y_1 = b_1 - A_{11}x_1 - A_{12}x_2, \quad y_2 = b_2 - A_{21}x_1 - A_{22}x_2.$$

Тогда задача (1) заменяется эквивалентной задачей на условный минимум

$$I_1 = \min_{[x,\, y] \in G} f(y), \tag{7}$$

где целевая функция и допустимое множество $G$ имеют вид

$$f(y) = \|(y_1)_+\|^2/2 + \|y_2\|^2/2,$$

$$G = \{[x, y] : A_{11}x_1 + A_{12}x_2 + y_1 = b_1,\ A_{21}x_1 + A_{22}x_2 + y_2 = b_2,\ x \in \Pi_x\}.$$

В отличие от множества $X$, множество $G$ всегда непусто.

Для задачи квадратичного программирования (7) определим вектор множителей Лагранжа $z \in \Pi_z$ и выпишем функцию Лагранжа:

$$L(x, y, z) = f(y) + z_1^{\mathrm{T}}(b_1 - A_{11}x_1 - A_{12}x_2 - y_1) + z_2^{\mathrm{T}}(b_2 - A_{21}x_1 - A_{22}x_2 - y_2).$$

Преобразуем ее к виду

$$L(x, y, z) = f(y) - x_1^{\mathrm{T}}(A_{11}^{\mathrm{T}}z_1 + A_{21}^{\mathrm{T}}z_2) - x_2^{\mathrm{T}}(A_{12}^{\mathrm{T}}z_1 + A_{22}^{\mathrm{T}}z_2) + z_1^{\mathrm{T}}(b_1 - y_1) + z_2^{\mathrm{T}}(b_2 - y_2). \tag{8}$$

Введем двойственную функцию

$$F(z) = \min_{x \in \Pi_x} \min_{y \in \mathbb{R}^n} L(x, y, z) \tag{9}$$

и рассмотрим двойственную к (7) задачу о нахождении

$$\max_{z \in \Pi_z} F(z).$$

Необходимые и достаточные условия минимума в задаче (9) имеют вид

$$L_{x_1}(x, y, z) = -A_{11}^{\mathrm{T}}z_1 - A_{21}^{\mathrm{T}}z_2 \ge 0_{n_1}, \quad D(x_1)(A_{11}^{\mathrm{T}}z_1 + A_{21}^{\mathrm{T}}z_2) = 0_{n_1}, \quad x_1 \ge 0_{n_1}, \tag{10}$$

$$L_{x_2}(x, y, z) = -A_{12}^{\mathrm{T}}z_1 - A_{22}^{\mathrm{T}}z_2 = 0_{n_2}, \tag{11}$$

$$L_{y_1}(x, y, z) = (y_1)_+ - z_1 = 0_{m_1}, \quad L_{y_2}(x, y, z) = y_2 - z_2 = 0_{m_2}. \tag{12}$$

Здесь и ниже через $D(z)$ обозначена диагональная матрица, у которой $i$-й диагональный элемент есть $i$-я компонента вектора $z$.

При $z \in \Pi_z$ из (12) имеем $z = y$. Подставим это соотношение в (8) и потребуем выполнения условия $z \in Z$. Тогда, согласно определению (9) и условиям (10), (11), двойственная функция принимает вид $F(z) = b^{\mathrm{T}}z - \|z\|^2/2$. Таким образом, приходим к задаче (5), которая является двойственной к (7) и, можно сказать, к (1). Итак, задачи безусловной минимизации (1) и квадратичного программирования (5) можно назвать взаимно двойственными. Аналогично, задачи (2) и (6) в указанном смысле являются двойственными друг к другу. Лемма доказана.

Для введенных задач справедлива теорема двойственности о равенстве оптимальных значений целевых функций

$$I_1 = I_1^d, \quad I_2 = I_2^d. \tag{13}$$

Проекцией точки $\bar{x}$ на непустое замкнутое множество $X$ назовем точку $\bar{x}^* \in X$, ближайшую к точке $\bar{x}$, т.е. решение задачи

$$J = \min_{x \in X} \|\bar{x} - x\| = \|\bar{x} - \bar{x}^*\|. \tag{14}$$

Будем писать $\bar{x}^* = \mathrm{pr}(\bar{x}, X)$, расстояние от точки $\bar{x}$ до множества $X$ обозначим через $\mathrm{dist}(\bar{x}, X) = \|\bar{x}^* - \bar{x}\|$.

**Теорема 1.** *Всякое решение* $x^*$ *задачи* (1) *определяет единственное решение* $z^{*\mathrm{T}} = [z_1^{*\mathrm{T}}, z_2^{*\mathrm{T}}]$ *задачи* (5) *по формулам*

$$z_1^* = (b_1 - A_{11}x_1^* - A_{12}x_2^*)_+, \quad z_2^* = b_2 - A_{21}x_1^* - A_{22}x_2^*, \tag{15}$$

*и справедливы следующие утверждения*:

$$\|z^*\|^2 = b^{\mathrm{T}}z^*, \tag{16}$$

$$z^* \perp Ax^*, \quad z^* \perp (b - z^*), \tag{17}$$

$$z^* = \mathrm{pr}(b, Z), \quad \|z^*\| = \mathrm{pen}(x^*, X), \quad \|b - z^*\| = \mathrm{dist}(b, Z), \tag{18}$$

$$[\mathrm{pen}(x^*, X)]^2 + [\mathrm{dist}(b, Z)]^2 = \|b\|^2. \tag{19}$$

**Доказательство.** В точке $x^*$ необходимые и достаточные условия минимума задачи (1) записываются в виде

$$\begin{gathered} -A_{11}^{\mathrm{T}}(b_1 - A_{11}x_1^* - A_{12}x_2^*)_+ - A_{21}^{\mathrm{T}}(b_2 - A_{21}x_1^* - A_{22}x_2^*) \ge 0_{n_1}, \\ D(x_1^*)[A_{11}^{\mathrm{T}}(b_1 - A_{11}x_1^* - A_{12}x_2^*)_+ + A_{21}^{\mathrm{T}}(b_2 - A_{21}x_1^* - A_{22}x_2^*)] = 0_{n_1}, \quad x_1^* \ge 0_{n_1}, \\ A_{12}^{\mathrm{T}}(b_1 - A_{11}x_1^* - A_{12}x_2^*)_+ + A_{22}^{\mathrm{T}}(b_2 - A_{21}x_1^* - A_{22}x_2^*) = 0_{n_2}. \end{gathered} \tag{20}$$

В формулах (20) введем обозначения

$$z_1^* = (b_1 - A_{11}x_1^* - A_{12}x_2^*)_+, \quad z_2^* = b_2 - A_{21}x_1^* - A_{22}x_2^* \tag{21}$$

и покажем, что $z^{*\mathrm{T}} = [z_1^{*\mathrm{T}}, z_2^{*\mathrm{T}}]$ является решением задачи (5). Необходимые и достаточные условия минимума (20) запишем в виде

$$A_{11}^{\mathrm{T}}z_1^* + A_{21}^{\mathrm{T}}z_2^* \le 0_{n_1}, \quad A_{12}^{\mathrm{T}}z_1^* + A_{22}^{\mathrm{T}}z_2^* = 0_{n_2}, \tag{22}$$

$$D(x_1^*)(A_{11}^{\mathrm{T}}z_1^* + A_{21}^{\mathrm{T}}z_2^*) = 0_{n_1}, \quad x_1^* \ge 0_{n_1}. \tag{23}$$

Из (21) и (22) следует, что $z^* \in Z$. Умножив в (21) первое соотношение скалярно на $z_1^*$, второе – на $z_2^*$, получим

$$\|z_1^*\|^2 = z_1^{*\mathrm{T}}(b_1 - A_{11}x_1^* - A_{12}x_2^*)_+ = z_1^{*\mathrm{T}}(b_1 - A_{11}x_1^* - A_{12}x_2^*) = b_1^{\mathrm{T}}z_1^* - x_1^{*\mathrm{T}}A_{11}^{\mathrm{T}}z_1^* - x_2^{*\mathrm{T}}A_{12}^{\mathrm{T}}z_1^*,$$

$$\|z_2^*\|^2 = z_2^*(b_2 - A_{21}x_1^* - A_{22}x_2^*) = b_2^{\mathrm{T}}z_2^* - x_1^{*\mathrm{T}}A_{21}^{\mathrm{T}}z_2^* - x_2^{*\mathrm{T}}A_{22}^{\mathrm{T}}z_2^*.$$

Складывая эти соотношения, находим

$$\|z^*\|^2 = \|z_1^*\|^2 + \|z_2^*\|^2 = b_1^{\mathrm{T}} z_1^* + b_2^{\mathrm{T}} z_2^* - x_1^{*\mathrm{T}}(A_{11}^{\mathrm{T}} z_1^* + A_{21}^{\mathrm{T}} z_2^*) - x_2^{*\mathrm{T}}(A_{12}^{\mathrm{T}} z_1^* + A_{22}^{\mathrm{T}} z_2^*) = \\ = (b - Ax^*)^{\mathrm{T}} z^* = b^{\mathrm{T}} z^*. \quad (24)$$

В (24) учтено, что, согласно (22) и (23), $x_1^{*\mathrm{T}}(A_{11}^{\mathrm{T}} z_1^* + A_{21}^{\mathrm{T}} z_2^*) + x_2^{*\mathrm{T}}(A_{12}^{\mathrm{T}} z_1^* + A_{22}^{\mathrm{T}} z_2^*) = x^{*\mathrm{T}} A^{\mathrm{T}} z^* = 0$. Таким образом, доказано равенство (16) и показано, что $z^{*\mathrm{T}} A x^* = 0$, т.е. векторы $z^*$ и $Ax^*$ ортогональны. Условие (16) можно переписать в виде $z^{*\mathrm{T}}(z^* - b) = 0$, откуда следует последнее утверждение в (17).

Запишем функцию Лагранжа для задачи (5) в виде

$$L(z, x) = b^{\mathrm{T}} z - \|z\|^2/2 - x_1^{\mathrm{T}}(A_{11}^{\mathrm{T}} z_1 + A_{21}^{\mathrm{T}} z_2) - x_2^{\mathrm{T}}(A_{12}^{\mathrm{T}} z_1 + A_{22}^{\mathrm{T}} z_2) \quad (25)$$

и приведем условия Куна–Таккера

$$L_{z_1}(z, x) = b_1 - z_1 - A_{11} x_1 - A_{12} x_2 \le 0_{m_1}, \quad (26)$$

$$D(z_1)(b_1 - z_1 - A_{11} x_1 - A_{12} x_2) = 0_{m_1}, \quad z_1 \ge 0_{m_1}, \quad (27)$$

$$L_{z_2}(z, x) = b_2 - z_2 - A_{21} x_1 - A_{22} x_2 = 0_{m_2}, \quad (28)$$

$$L_{x_1}(z, x) = -(A_{11}^{\mathrm{T}} z_1 + A_{21}^{\mathrm{T}} z_2) \ge 0_{n_1}, \quad x_1 \ge 0_{n_1}, \quad D(x_1)(A_{11}^{\mathrm{T}} z_1 + A_{21}^{\mathrm{T}} z_2) = 0_{n_1}, \quad (29)$$

$$L_{x_2}(z, x) = -(A_{12}^{\mathrm{T}} z_1 + A_{22}^{\mathrm{T}} z_2) = 0_{n_2}. \quad (30)$$

Сравним необходимые и достаточные условия минимума (21)–(23) для задачи (1) с необходимыми и достаточными условиями оптимальности (26)–(30) для задачи квадратичного программирования (5). Если в условиях Куна–Таккера (26)–(30) в качестве вектора $x$ взять вектор $x^*$, а в качестве $z$ вектор $z^*$, определяемый в (21), то (29), (30) переходят в (22), (23). Легко видеть, что соотношения (21) обеспечивают выполнение условий (26)–(28). Таким образом, седловую точку $[z^*, x^*]$ функции Лагранжа (25) составляют вектор $z^*$ – решение задачи (5) и вектор $x^*$ – решение задачи (1), причем между этими векторами имеется связь, задаваемая формулами (15).

Из (5), (13) и (16) следует

$$I_1 = I_1^d = \|z^*\|^2/2 = [\mathrm{pen}(x^*, X)]^2/2. \quad (31)$$

Отсюда получаем второе утверждение в (18).

Задачу (5) преобразуем к эквивалентным задачам

$$I_1^d = \max_{z \in Z}[-\|b - z\|^2 + \|b\|^2/2] = \|b\|^2/2 - \min_{z \in Z}\|b - z\|^2/2. \quad (32)$$

Так как $z^*$ – единственное решение задачи (5), то приходим к выводу, что $z^* = \mathrm{pr}(b, Z)$ и $\|b - z^*\| = \mathrm{dist}(b, Z)$, т.е. доказаны все утверждения (18). С учетом (16) имеем $\|z^*\|^2 + \|b - z^*\|^2 = \|b\|^2$. Откуда следует (19). Теорема доказана.

В теореме 1 утверждение (16) является следствием двойственности задач (1) и (5) (см. (13)). Формулы (15) позволяют выразить оптимальный вектор $z^*$ в задаче (5) через оптимальный вектор $x^*$ задачи (1). Полученный из (15) вектор $z^*$ будем называть *вектором минимальных невязок*. Из (31) следует

**Критерий 1.** Система (I) разрешима тогда и только тогда, когда вектор минимальных невязок $z^*$ равен нулю (задача (5) имеет нулевое решение).

Вектор $x^*$, удовлетворяющий необходимым и достаточным условиям минимума (20) для задачи (1), назовем *псевдорешением системы* (I). Если псевдорешение $x^* \in X$, то этот вектор $x^*$ является решением системы (I). Такая терминология принята в методе наименьших квадратов. Аналогично, решение $u^*$ задачи (2) назовем *псевдорешением системы* (II).

Анализ задач (2) и (6) аналогичен приведенному для задач (1), (5), однако есть некоторые отличия. Поэтому относительно задач (2), (6) приведем теорему, аналогичную теореме 1. Пусть $\hat{A} = [-A, b]$ – матрица $m \times (n + 1)$, вектор $r \in \mathbb{R}^{n+1}$ имеет вид $r^{\mathrm{T}} = [0_n^{\mathrm{T}}, \rho]$.

**Теорема 2.** *Пусть* $u^{*\mathrm{т}} = [u_1^{*\mathrm{т}}, u_2^{*\mathrm{т}}]$ – *произвольное решение задачи* (2). *Тогда решение* $w^{*\mathrm{т}} = [w_1^{*\mathrm{т}}, w_2^{*\mathrm{т}}, w_3^*]$ *задачи* (6) *выражается через решение* $u^*$ *задачи* (2) *по формулам*

$$w_1^* = (A_{11}^{\mathrm{т}}u_1^* + A_{21}^{\mathrm{т}}u_2^*)_+, \quad w_2^* = A_{12}^{\mathrm{т}}u_1^* + A_{22}^{\mathrm{т}}u_2^*, \quad w_3^* = \rho - b_1^{\mathrm{т}}u_1^* - b_2^{\mathrm{т}}u_2^* \tag{33}$$

*и справедливы следующие утверждения*:

$$\|w^*\|^2 = \rho w_3^*, \tag{34}$$

$$w^* \perp \hat{A}^{\mathrm{т}}u^*, \quad w^* \perp (r - w^*), \tag{35}$$

$$w^* = \mathrm{pr}(r, W), \quad \|w^*\| = \mathrm{pen}(u^*, U), \quad \|r - w^*\| = \mathrm{dist}(r, W), \tag{36}$$

$$[\mathrm{pen}(u^*, U)]^2 + [\mathrm{dist}(r, W)]^2 = \|r\|^2, \tag{37}$$

$$\|w^*\| \le \rho, \quad 0 \le w_3^* \le \rho, \quad \|w_1^*\|^2 + \|w_2^*\|^2 \le \rho^2/4. \tag{38}$$

**Доказательство.** Задача строго выпуклого квадратичного программирования (6) является задачей нахождения проекции вектора $[0_n^{\mathrm{т}}, \rho]$ на непустое множество $W$, заданное системой линейных равенств и неравенств (II)'. Эта задача всегда имеет единственное решение, и для нее существует вектор множителей Лагранжа $u \in \mathbb{R}^n$. Приведем функцию Лагранжа для задачи (6):

$$L(w, u) = \rho w_3 - \|w\|^2/2 - u_1^{\mathrm{т}}(b_1 w_3 - A_{11}w_1 - A_{12}w_2) - u_2^{\mathrm{т}}(b_2 w_3 - A_{21}w_1 - A_{22}w_2).$$

Для задачи (6) запишем необходимые и достаточные условия оптимальности (условия Куна–Таккера), вычисленные в седловой точке $[w^*, u^*]$:

$$-w_1^* + A_{11}^{\mathrm{т}}u_1^* + A_{21}^{\mathrm{т}}u_2^* \le 0_{n_1}, \quad D(w_1^*)(-w_1^* + A_{11}^{\mathrm{т}}u_1^* + A_{21}^{\mathrm{т}}u_2^*) = 0_{n_1}, \quad w_1^* \ge 0_{n_1}, \tag{39}$$

$$-w_2^* + A_{12}^{\mathrm{т}}u_1^* + A_{22}^{\mathrm{т}}u_2^* = 0_{n_2}, \tag{40}$$

$$\rho - w_3^* - b_1^{\mathrm{т}}u_1^* - b_2^{\mathrm{т}}u_2^* = 0, \tag{41}$$

$$A_{11}w_1^* + A_{12}w_2^* - b_1 w_3^* \ge 0_{m_1}, \quad D(u_1^*)(A_{11}w_1^* + A_{12}w_2^* - b_1 w_3^*) = 0_{m_1}, \quad u_1^* \ge 0_{m_1}, \tag{42}$$

$$A_{21}w_1^* + A_{22}w_2^* - b_2 w_3^* = 0_{m_2}. \tag{43}$$

Из условий (39)–(41) следует, что $w^* \in W$ и вектор $w^*$ выражается через решение $u^*$ задачи (2) по формулам (33). Подставляя их в условия (42), (43), приходим к соотношениям

$$A_{11}(A_{11}^{\mathrm{т}}u_1^* + A_{21}^{\mathrm{т}}u_2^*)_+ + A_{12}(A_{12}^{\mathrm{т}}u_1^* + A_{22}^{\mathrm{т}}u_2^*) - b_1(\rho - b_1^{\mathrm{т}}u_1^* - b_2^{\mathrm{т}}u_2^*) \ge 0_{m_1}, \quad u_1^* \ge 0_{m_1},$$

$$D(u_1^*)[A_{11}(A_{11}^{\mathrm{т}}u_1^* + A_{21}^{\mathrm{т}}u_2^*)_+ + A_{12}(A_{12}^{\mathrm{т}}u_1^* + A_{22}^{\mathrm{т}}u_2^*) - b_1(\rho - b_1^{\mathrm{т}}u_1^* - b_2^{\mathrm{т}}u_2^*)] = 0_{m_1},$$

$$A_{21}(A_{11}^{\mathrm{т}}u_1^* + A_{21}^{\mathrm{т}}u_2^*)_+ + A_{22}(A_{12}^{\mathrm{т}}u_1^* + A_{22}^{\mathrm{т}}u_2^*) - b_2(\rho - b_1^{\mathrm{т}}u_1^* - b_2^{\mathrm{т}}u_2^*) = 0_{m_2}.$$

Эти соотношения совпадают с необходимыми и достаточными условиями оптимальности задачи (2), вычисленными в точке $u^*$. Итак, из условий Куна–Таккера и условий оптимальности задачи (2) следует, что в рассмотренной седловой точке $[w^*, u^*]$ вектор $w^*$ – решение задачи (6) и $u^*$ – решение задачи (2).

Из равенства целевых функций прямой (6) и двойственной (2) задач с учетом соотношений (33) получаем (34).

Доказательство соотношений (35)–(37) совершенно аналогично доказательству соотношений (17)–(19) в теореме 1.

При $\|w^*\| \neq 0$ из (34) следует, что $w_3^* > 0$. Рассмотрим (34) как квадратное уравнение относительно $w_3^*$:

$$w_3^{*2} - \rho w_3^* + \|w_1^*\|^2 + \|w_2^*\|^2 = 0. \tag{44}$$

Из того, что $\|r\| = \rho$ и вектор $w^*$ – проекция $r$ на $W$, следует первое неравенство в (38). Вторая оценка в (38) следует из условия неотрицательности свободного члена квадратного уравнения: $w_3^*(\rho - w_3^*) = \|w_1^*\|^2 + \|w_2^*\|^2 \geq 0$, а третья – из условия неотрицательности детерминанта квадратного уравнения (44). Теорема доказана.

**Критерий 2.** Система (II) разрешима (неразрешима) тогда и только тогда, когда задача (6) имеет нулевое (ненулевое) решение $w^*$.

**Теорема 3.** *Пусть $x^*$ и $u^*$ – произвольные решения задач* (1) *и* (2) *соответственно, векторы минимальных невязок $z^*$ и $w^*$ определяются по формулам* (15) *и* (33). *Тогда справедливы следующие утверждения*:

1) *системы* (I) *и* (II) *альтернативны, т.е. разрешима одна и только одна из них*;

2) *если система* (I) *несовместна, то нормальное решение $\tilde{u}^*$ системы* (II) *и вектор минимальных невязок $z^*$ системы* (I) *коллинеарны и*

$$\tilde{u}^* = \rho z^*/\|z^*\|^2, \quad z^* = \rho \tilde{u}^*/\|\tilde{u}^*\|^2; \tag{45}$$

3) *если система* (II) *несовместна, то составляющие нормального решения $\tilde{x}^{*\mathrm{T}} = [\tilde{x}_1^{*\mathrm{T}}, \tilde{x}_2^{*\mathrm{T}}]$ системы* (I) *имеют вид*

$$\tilde{x}_1^* = w_1^*/w_3^*, \quad \tilde{x}_2^* = w_2^*/w_3^*. \tag{46}$$

**Доказательство.** Из леммы 2 следует, что системы (I) и (II) не могут быть совместны одновременно. Покажем, что одна из них обязательно совместна. Рассмотрим каждый случай отдельно.

Пусть $X = \varnothing$, тогда $\mathrm{pen}(x^*, X) \neq 0$. Определяемый из (15) вектор $z^* \in Z$ таков, что $\|z^*\| \neq 0$. Умножим в первой формуле (45) левую и правую часть на $b$, учтем (16), получим $b^{\mathrm{T}}\tilde{u}^* = \rho$, поэтому $\tilde{u}^* \in U$, следовательно, $U \neq \varnothing$. Покажем, что $\tilde{u}^*$ есть нормальное решение системы (II), т.е. является решением задачи

$$\min_{u \in U} \|u\|^2/2. \tag{47}$$

Запишем функцию Лагранжа для задачи (47)

$$L(u, \hat{x}) = \|u\|^2/2 + \hat{x}_1^{\mathrm{T}}(A_{11}^{\mathrm{T}}u_1 + A_{21}^{\mathrm{T}}u_2) + \hat{x}_2^{\mathrm{T}}(A_{12}^{\mathrm{T}}u_1 + A_{22}^{\mathrm{T}}u_2) + \hat{x}_3(\rho - b_1^{\mathrm{T}}u_1 - b_2^{\mathrm{T}}u_2).$$

Двойственной к (47) является задача

$$\max_{\hat{x}_1 \in \mathbb{R}_+^{n_1}} \max_{\hat{x}_2 \in \mathbb{R}^{n_2}} \max_{\hat{x}_3 \in \mathbb{R}^1} [\rho\hat{x}_3 - \|(b_1\hat{x}_3 - A_{11}\hat{x}_1 - A_{12}\hat{x}_2)_+\|^2/2 - \|b_2\hat{x}_3 - A_{21}\hat{x}_1 - A_{22}\hat{x}_2\|^2/2]. \tag{48}$$

Приведем условия Куна–Таккера, вычисленные в седловой точке $[u^*, \hat{x}^*]$, где $u^{*\mathrm{T}} = [u_1^{*\mathrm{T}}, u_2^{*\mathrm{T}}]$ – решение задачи (47) и $\hat{x}^{*\mathrm{T}} = [\hat{x}_1^{*\mathrm{T}}, \hat{x}_2^{*\mathrm{T}}, \hat{x}_3^{*\mathrm{T}}]$ – решение задачи (48):

$$u_1^* + A_{11}\hat{x}_1^* + A_{12}\hat{x}_2^* - b_1\hat{x}_3^* \geq 0_{m_1}, \quad D(u_1^*)(u_1^* + A_{11}\hat{x}_1^* + A_{12}\hat{x}_2^* - b_1\hat{x}_3^*) = 0, \quad u_1^* \geq 0_{m_1}, \tag{49}$$

$$u_2^* + A_{21}\hat{x}_1^* + A_{22}\hat{x}_2^* - b_2\hat{x}_3^* = 0_{m_2}, \tag{50}$$

$$A_{11}^{\mathrm{T}}u_1^* + A_{21}^{\mathrm{T}}u_2^* \leq 0_{n_1}, \quad D(\hat{x}_1^*)(A_{11}^{\mathrm{T}}u_1^* + A_{21}^{\mathrm{T}}u_2^*) = 0_{n_1}, \quad \hat{x}_1^* \geq 0_{n_1}, \tag{51}$$

$$A_{12}^{\mathrm{T}}u_1^* + A_{22}^{\mathrm{T}}u_2^* = 0_{n_2}, \tag{52}$$

$$\rho - b_1^{\mathrm{T}}u_1^* - b_2^{\mathrm{T}}u_2^* = 0. \tag{53}$$

Из (49) и (50) получаем, что $u^*$ – решение задачи (47) – связано с решением $\hat{x}^*$ задачи (48) соотношениями

$$u_1^* = (b_1\hat{x}_3^* - A_{11}\hat{x}_1^* - A_{12}\hat{x}_2^*)_+, \quad u_2^* = b_2\hat{x}_3^* - A_{21}\hat{x}_1^* - A_{22}\hat{x}_2^*.$$

Используя эти соотношения, из равенств оптимальных значений целевых функций прямой задачи (47) и двойственной (48) имеем $\|u^*\|^2 = \rho\hat{x}_3^*$. Так как $U \neq \varnothing$ и $u^* \in U$, то $\|u^*\| \neq 0$ в силу условия $b^{\mathrm{T}}u^* = \rho > 0$. Поэтому имеем $\hat{x}_3^* > 0$.

Сделав в (49)–(52) замену

$$u^* = \hat{x}_3^* z^*, \quad \hat{x}_1^* = \hat{x}_3^* x_1^*, \quad \hat{x}_2^* = \hat{x}_3^* x_2^*$$

и сократив эти выражения на положительную величину $\hat{x}_3^*$, придем к условиям Куна–Таккера (26)–(30) для задачи (5), выполненным в точке $[z^*, x^*]$. Подставив $u^* = \hat{x}_3^* z^*$ в (53), с учетом (16) получим

$$\rho/\hat{x}_3^* - b^{\mathrm{T}}z^* = \rho/\hat{x}_3^* - \|z^*\|^2.$$

Отсюда при $\hat{x}_3^* = \rho/\|z^*\|^2$ следует, что $u^* = \hat{x}_3^* z^* = \rho z^*/\|z^*\|^2 = \tilde{u}^*$, т.е. нормальное решение системы (II) выражается по первой формуле в (45).

Из первой формулы в (45) следует $\rho = \|\tilde{u}^*\|\,\|z^*\|$, $z^* = \tilde{u}^*\,\|z^*\|^2/\rho = \tilde{u}^*\rho^2/(\rho\|\tilde{u}^*\|^2) = \rho\tilde{u}^*/\|\tilde{u}^*\|^2$, т.е. приходим ко второй формуле в (45).

Пусть теперь $U = \varnothing$, тогда $\mathrm{pen}(u^*, U) \neq 0$. Определяемый из (33) вектор $w^* \in W$ таков, что $\|w^*\| \neq 0$. При этом в силу условия (34) имеем $w_3^* > 0$. Из условий (42), (43) следует, что вектор $x^* = [x_1^*, x_2^*]$ с компонентами

$$x_1^* = w_1^*/w_3^*, \quad x_2^* = w_2^*/w_3^* \tag{54}$$

является решением системы (I).

Пусть $\tilde{x}^*$ – нормальное решение системы (I), т.е. является решением задачи

$$\min_{x \in X}\|x\|^2. \tag{55}$$

Приведем функцию Лагранжа для этой задачи:

$$L(x, \mu) = \|x\|^2/2 + \mu_1^{\mathrm{T}}(b_1 - A_{11}x_1 - A_{12}x_2) + \mu_2^{\mathrm{T}}(b_2 - A_{21}x_1 - A_{22}x_2),$$

и условия Куна–Таккера, вычисленные в седловой точке $[\tilde{x}^*, \mu^*]$, где $\tilde{x}^{*\mathrm{T}} = [\tilde{x}_1^{*\mathrm{T}}, \tilde{x}_2^{*\mathrm{T}}]$ – решение задачи (55), $\mu^{*\mathrm{T}} = [\mu_1^{*\mathrm{T}}, \mu_2^{*\mathrm{T}}]$ – оптимальный вектор множителей Лагранжа в задаче (55):

$$\tilde{x}_1^* - A_{11}^{\mathrm{T}}\mu_1^* - A_{21}^{\mathrm{T}}\mu_2^* \geq 0_{n_1}, \quad D(\tilde{x}_1^*)(\tilde{x}_1^* - A_{11}^{\mathrm{T}}\mu_1^* - A_{21}^{\mathrm{T}}\mu_2^*) = 0_{n_1}, \quad \tilde{x}_1^* \geq 0_{n_1},$$

$$\tilde{x}_2^* - A_{12}^{\mathrm{T}}\mu_1^* - A_{22}^{\mathrm{T}}\mu_2^* = 0_{n_2},$$

$$b_1 - A_{11}\tilde{x}_1^* - A_{12}\tilde{x}_2^* \leq 0_{m_1}, \quad D(\mu_1^*)(b_1 - A_{11}\tilde{x}_1^* - A_{12}\tilde{x}_2^*) = 0_{m_1}, \quad \mu_1^* \geq 0_{m_1},$$

$$b_2 - A_{21}\tilde{x}_1^* - A_{22}\tilde{x}_2^* = 0_{m_2}.$$

Эти условия Куна–Таккера для задачи (55) можно получить из соотношений (39), (40), (42), (43), входящих в условия Куна–Таккера для задачи (6), если соотношения разделить на $w_3^*$ и обозначить $x_1^* = w_1^*/w_3^*$, $x_2^* = w_2^*/w_3^*$, $\mu_1^* = u_1^*/w_3^*$, $\mu_2^* = u_3^*/w_3^*$. Таким образом, получаем, что вектор $\tilde{x}^*$, определяемый составляющими (54), является нормальным решением системы (I). Теорема доказана.

Итак, теорема 3 сводит вопрос о разрешимости системы (I) или (II) к решению задачи минимизации невязки любой из этих систем. При этом если норма минимальной невязки оказалась ненулевой, то эта система несовместна и по этой невязке по простым формулам находится нормальное решение совместной системы.

Если в задачах (1) и (2) использовать взвешенные невязки, то введенные при невязках коэффициенты должны быть учтены в обеих альтернативных системах и формулы (45), (46) пригодны для вычисления нормальных решений в этих видоизмененных системах.

Возможны различные варианты представления альтернативной системы (II). Как следует из приведенных теорем, система, альтернативная к (I), получается из сопряженной системы (I)' путем добавления условия, исключающего возможность тривиального решения задачи (5). Например, можно потребовать для решений сопряженной системы (I)' выполнения условия $b^{\mathrm{T}}u > 0$ (как в лемме Фаркаша) либо условия $b^{\mathrm{T}}u = 1$ (как в теореме Гейла), и т.д. Если в системе (I) отсутствуют ограничения типа неравенств, то в теореме 3 на $\rho$ достаточно наложить условие $\rho \neq 0$.

## 3. ОПРЕДЕЛЕНИЕ НАПРАВЛЕНИЯ НАИСКОРЕЙШЕГО СПУСКА В МЕТОДЕ ВОЗМОЖНЫХ НАПРАВЛЕНИЙ

Ниже будут рассмотрены различные системы, являющиеся частными случаями систем (I) и (II). За этими системами сохраним обозначения (I) и (II), добавляя индексы, указывающие номера соответствующих разделов.

Пусть не равен нулю вектор $z^*$ – решение задачи (5). Введем нормированные векторы $z_n = z/\|z^*\|$ и $z_n^* = z^*/\|z^*\|$. Определим допустимое множество нормированных векторов

$$Z_n = \{z_n \in \mathbb{R}^m : z_n \in Z, \|z_n\| = 1\},$$

где $Z$ – сопряженное множество (I)'.

Рассмотрим следующую вспомогательную задачу:

$$I_3 = \max_{z_n \in Z_n} b^{\mathrm{T}} z_n. \tag{56}$$

**Теорема 4.** *Пусть $x^*$ – произвольное решение задачи* (1), *$z^*$ – решение задачи* (5), *определяемое по формуле* (15); *кроме того,* $\|z^*\| \neq 0$. *Тогда решением задачи* (56) *является вектор* $z_n^* = z^*/\|z^*\|$, *причем*

$$I_3 = b^{\mathrm{T}} z_n^* = \|z^*\|. \tag{57}$$

**Доказательство.** Из равенства оптимальных значений целевых функций взаимно двойственных задач (1) и (5) следует, что

$$1/2 = \max_{z_n \in Z} [b^{\mathrm{T}} z_n/\|z^*\| - \|z_n\|^2/2].$$

Так как $Z_n \subset Z$, то имеем

$$1/2 = \max_{z_n \in Z} [b^{\mathrm{T}} z_n/\|z^*\| - \|z_n\|^2/2] \geq \max_{z_n \in Z_n} [b^{\mathrm{T}} z_n/\|z^*\| - \|z_n\|^2/2] = \max_{z_n \in Z_n} [b^{\mathrm{T}} z_n/\|z^*\| - 1/2].$$

Отсюда следует

$$\|z^*\| \geq \max_{z_n \in Z_n} b^{\mathrm{T}} z_n. \tag{58}$$

Если в качестве $z_n$ взять вектор $z^*/\|z^*\|$, который принадлежит допустимому множеству $Z_n$, то, согласно (16), неравенство (58) переходит в равенство. Теорема доказана.

Задача вида (56) возникает в методе возможных направлений при решении задачи НЛП

$$\min_{p \in P} f(p), \quad P = \{p \in \mathbb{R}^m : h(p) \leq 0_{n_1}, g(p) = 0_{n_2}\}. \tag{59}$$

Здесь $f: \mathbb{R}^m \longrightarrow \mathbb{R}^1$, $h: \mathbb{R}^m \longrightarrow \mathbb{R}^{n_1}$, $g: \mathbb{R}^m \longrightarrow \mathbb{R}^{n_2}$, функции $f(p)$, $h(p)$, $g(p)$ непрерывно дифференцируемы, множество $P$ не пусто, задача (59) имеет решение.

Предположим, что фиксирована произвольная допустимая точка $p \in P$. Введем вектор множителей Лагранжа $x \in \mathbb{R}^n$, $x^{\mathrm{T}} = [x_1^{\mathrm{T}}, x_2^{\mathrm{T}}]$, где $x_1 \in \mathbb{R}_+^{n_1}$, $x_2 \in \mathbb{R}^{n_2}$, $n = n_1 + n_2$, и определим функцию Лагранжа

$$L(p, x) = f(p) + h^{\mathrm{T}}(p)x_1 + g^{\mathrm{T}}(p)x_2.$$

Введем условия дополняющей нежесткости

$$x_1^i h^i(p) = 0, \quad 1 \le i \le n_1. \tag{60}$$

Компонента $h^i(p)$ вектора $h(p)$ называется *активной* в точке $p \in P$, если $h^i(p) = 0$. В силу (60), все компоненты вектора $x_1$, соответствующие неактивным компонентам вектора $h(p)$, равны нулю. Для простоты будем считать, что в функцию Лагранжа входит вектор $h(p)$, все компоненты которого активны. Приведем условия Куна–Таккера для задачи (59), вычисленные в точке $[p, x]$, где $p \in P$:

$$L_p(p, x) = f_p(p) + h_p(p)x_1 + g_p(p)x_2 = 0_m, \quad x_1 \ge 0_{n_1}. \tag{I$_3$}$$

При фиксированном $p \in P$ эти условия относительно $x$ можно рассматривать как частный случай системы (I).

Введем вектор $p' = p + \tau z$, где $\tau$ – длина шага, направление спуска $z \in \mathbb{R}^m$, $\|z\| = 1$. В задаче (59) линеаризуем целевую функцию и функции, задающие ограничения. Величину $\tau$ считаем малой; отбрасывая члены высшего порядка малости, приходим к следующей задаче отыскания направления наискорейшего спуска:

$$I_4 = \min_{z \in \hat{Z}_n} z^{\mathrm{T}} f_p(p), \quad \hat{Z}_n = \{z \in \mathbb{R}^m : h_p^{\mathrm{T}}(p)z \le 0_{n_1}, g_p^{\mathrm{T}}(p)z = 0_{n_2}, \|z\| = 1\}. \tag{61}$$

Если задача (61) имеет решение $z_n^*$ и при этом $I_4 < 0$, то такое направление назовем *направлением наискорейшего спуска*. Это значит, что по крайней мере в линейном приближении можно "улучшить" точку $p$, взяв новый вектор $p'$. Тогда при достаточно малом шаге $\tau$ вектор $p'$ остается в допустимом множестве $P$ и $f(p') < f(p)$. Если задача (61) не имеет такого решения, то точку $p$ невозможно локально улучшить.

Чтобы воспользоваться ранее полученными результатами, будем считать, что $h_p^{\mathrm{T}}(p) = A_{21}^{\mathrm{T}}$, $g_p^{\mathrm{T}}(p) = A_{22}^{\mathrm{T}}$, $-f_p(p) = b_2$, остальные подматрицы матрицы $A$ и вектор $b_1$ нулевые. Тогда система (II), альтернативная к (I), запишется в виде

$$u^{\mathrm{T}} h_p(p) \le 0_{n_1}^{\mathrm{T}}, \quad u^{\mathrm{T}} g_p(p) = 0_{n_2}^{\mathrm{T}}, \quad -u^{\mathrm{T}} f_p(p) = \rho > 0. \tag{II$_3$}$$

Если система (II)$_3$ разрешима, то, согласно теореме 3, ее нормальное решение имеет вид $\tilde{u}^* = \rho z^*/\|z^*\|^2$, где $z^* = -L_p(p, x^*)$, и вектор $x^*$ находится из решения задачи безусловной минимизации (1), которая в данном случае записывается в виде

$$I_1 = \min_{x_1 \in \mathbb{R}_+^{n_1}} \min_{x_2 \in \mathbb{R}^{n_2}} \|L_p(p, x)\|^2/2. \tag{62}$$

Нормируем вектор $\tilde{u}^*$, получим $\tilde{u}_n^* = z^*/\|z^*\| = z_n^*$. Вектор $z_n^*$ принадлежит $\hat{Z}_n$, и согласно теореме 4 (см. формулу (57)) имеем $I_4 = -I_3 = -\|z^*\|$. Т.е. направление $z_n^*$ является направлением наискорейшего спуска в линеаризованной задаче (61). Это направление существует тогда и только тогда, когда система (I)$_3$ при фиксированном $p \in P$ не может быть разрешима относительно множителей Лагранжа $x_1 \in \mathbb{R}_+^{n_1}$, $x_2 \in \mathbb{R}^{n_2}$ задачи (59), т.е. $I_1 > 0$.

Таким образом, для определения наискорейшего спуска нет необходимости решать задачу условной минимизации (61). Это направление находится из решения задачи безусловной минимизации (62).

Отметим, что такой подход особенно эффективен, когда $n$ – количество активных ограничений в точке $p \in P$ – много меньше размерности вектора $p$, так как минимизация в задаче (62) ведется в $n$-мерном пространстве. В точке $p'$ переопределяется набор активных ограничений, который по-прежнему обозначается через $h(p)$.

## 4. ЗАДАЧИ ПРОЕКТИРОВАНИЯ И КОРРЕКЦИИ

Приведем геометрическую интерпретацию полученных результатов. Из (16) следует, что вектор невязки $z^*$ ортогонален вектору $b - z^*$. Поэтому три точки: начало координат в $\mathbb{R}^m$, точки $z^*$ и $b$ – образуют прямоугольный треугольник, у которого вектор $b$ определяет гипотенузу, вектор $z^*$ – катет длиной, равной $\operatorname{pen}(x^*, X)$, вектор $b - z^*$ – второй катет длиной $\operatorname{dist}(b, Z)$. Из теоремы Пифагора следует (19). Пусть $b^\perp$ – проекция вектора $b$ на множество $Z$. Тогда ортогональный к нему вектор $b^\parallel = b - b^\perp$ образован из двух составляющих: $b_1^\parallel = b_1 - (b_1 - A_{11}x_1^* - A_{12}x_2^*)_+$ и $b_2^\parallel = A_{21}x_1^* + A_{22}x_2^*$. Из (16), (17) следует, что для $b_1^\perp \perp b_1^\parallel$ и $b_2^\perp \perp b_2^\parallel$ формулы (18) примут вид

$$z^* = b^\perp = \operatorname{pr}(b, Z), \quad \|b^\perp\| = \operatorname{pen}(x^*, X), \quad \|b^\parallel\| = \operatorname{dist}(b, Z). \tag{63}$$

В этих обозначениях соотношения (16) и (19) становятся совершенно очевидными: $\|b^\perp\|^2 = \|b^\perp\|^2$, $\|b^\perp\|^2 + \|b^\parallel\|^2 = \|b\|^2$. Если $Z$ – линейное подпространство, то $b^\perp$ – проекция вектора $b^\parallel$ на $Z$, а $b^\parallel$ – проекция $b$ на ортогональное дополнение к $Z$.

Аналогично, в теореме 3 имеем

$$w^* = r^\perp = \operatorname{pr}(r, W), \quad \|r^\perp\| = \operatorname{pen}(u^*, U), \quad r^\parallel = r - r^\perp, \quad \|r^\parallel\| = \operatorname{dist}(r, W). \tag{64}$$

Из теорем 2 и 3, формул (63) и (64) следует, что условия (3) критерия альтернативности можно представить в виде

$$\|b^\perp\|\|r^\perp\| = \|z^*\|\|w^*\| = 0, \quad \|b^\perp\| + \|r^\perp\| = \|z^*\| + \|w^*\| > 0.$$

Таким образом, системы (I) и (II) альтернативны.

Из теоремы 1 следует, что $\|z^*\| \le \|b\|$, $b^\mathrm{T}z^* \ge 0$. Вектор $z^*$ лежит в полусфере с центром в начале координат пространства $\mathbb{R}^m$ и с радиусом, равным $\|b\|$, причем $z^*$ лежит в той половине сферы, где вектор $z^*$ образует острый угол с вектором $b$. Очевидно, что векторы $r^\perp$ и $r^\parallel$ лежат в сфере радиуса $\rho$ с центром в начале координат пространства $\mathbb{R}^{n+1}$.

Рассмотрим задачу (14) о проектировании точки $\bar{x} \in \mathbb{R}^n$ на непустое множество $X$. Согласно теореме 3, нормальное решение системы (I) есть проекция начала координат на множество $X$: $\tilde{x}^* = \operatorname{pr}(0_n, X)$. Сделав в (14) замену переменных $y = x - \bar{x}$, сведем ее к задаче проектирования начала координат на "сдвинутое" множество

$$\bar{X} = \{y \in \mathbb{R}^n : A_{11}y_1 + A_{12}y_2 \ge \bar{b}_1,\ A_{21}y_1 + A_{22}y_2 = \bar{b}_2,\ y_1 \ge -\bar{x}_1\}, \tag{65}$$

где $\bar{b}_1 = b_1 - A_{11}\bar{x}_1 - A_{12}\bar{x}_2$ и $\bar{b}_2 = b_2 - A_{21}\bar{x}_1 - A_{22}\bar{x}_2$.

Задача (14) примет вид

$$J = \min_{y \in \bar{X}} \|y\| = \|y^*\| = \|\operatorname{pr}(0_n, \bar{X})\|. \tag{66}$$

Между решениями задач (14) и (66) есть простая связь:

$$\bar{x}^* = \operatorname{pr}(\bar{x}, X) = \bar{x} + y^*. \tag{67}$$

Аналогично можно показать, что если $U \ne \varnothing$, то нормальное решение системы (II) есть $\tilde{u}^* = \operatorname{pr}(0_m, U)$ и

$$\bar{u}^* = \operatorname{pr}(\bar{u}, U) = \bar{u} + v^*, \tag{68}$$

где $v^* = \operatorname{pr}(0_n, \bar{U})$, $\bar{U} = \{v \in \mathbb{R}^m : A_{11}^\mathrm{T} v_1 + A_{21}^\mathrm{T} v_2 \le d_1,\ A_{12}^\mathrm{T} v_1 + A_{22}^\mathrm{T} v_2 = d_2,\ b_1^\mathrm{T} v_1 + b_2^\mathrm{T} v_2 = d_3,\ v_1 \ge -\bar{u}_1\}$ и $d_1 = -A_{11}^\mathrm{T}\bar{u}_1 - A_{21}^\mathrm{T}\bar{u}_2$, $d_2 = -A_{12}^\mathrm{T}\bar{u}_1 - A_{22}^\mathrm{T}\bar{u}_2$, $d_3 = \rho - b_1^\mathrm{T}\bar{u}_1 - b_2^\mathrm{T}\bar{u}_2$.

В [11], [12] ставятся задачи оптимальной коррекции линейных несовместных (несобственных) систем. Применительно к системе (I) речь идет о нахождении такого вектора $\tilde{b}$, чтобы при замене вектора $b$ на $b - \tilde{b}$ несовместная система (I) становилась совместной (собственной) и при этом евклидова норма вектора $\tilde{b}$ была минимальна.

**Теорема 5.** *Пусть $x^*$ – произвольное решение задачи* (1), *$z^*$ – вектор минимальных невязок, вычисленный в точке $x^*$. Тогда оптимальная коррекция системы* (I) *состоит в замене вектора $b$ на вектор $b - z^*$. Псевдорешение $x^*$ системы* (I) *является решением скорректированной системы*

$$A_{11}x_1 + A_{12}x_2 \ge b_1 - z_1^*, \quad A_{21}x_1 + A_{22}x_2 = b_2 - z_2^*, \quad x_1 \ge 0_{n_1}. \tag{69}$$

Согласно (63), составляющие вектора $z^*$ представимы в виде $z_1^* = b_1^{\perp}$, $z_2^* = b_2^{\perp}$, тогда $b_1^{\parallel} = b_1 - b_1^{\perp}$, $b_1^{\parallel} \le b_1$, $b_2^{\parallel} = b_2 - b_2^{\perp}$ и (69) можно записать в виде

$$A_{11}x_1 + A_{12}x_2 \ge b_1^{\parallel}, \quad A_{21}x_1 + A_{22}x_2 = b_2^{\parallel}, \quad x_1 \ge 0_{n_1}. \tag{70}$$

При замене $b$ на $b^{\parallel}$ альтернативная система (II) становится несовместной и решение задачи минимизации ее невязки согласно теореме 3 позволяет определить нормальное решение $\tilde{x}^*$ скорректированной системы (70). Итак, коррекция несовместной задачи (I) и нахождение нормального решения скорректированной системы (70) состоит в решении двух задач безусловной минимизации в $\mathbb{R}^n$ и $\mathbb{R}^m$. Для нахождения $m$-мерного вектора $\tilde{b} = z^*$ с минимальной евклидовой нормой решается задача безусловной минимизации (1) с $n$ переменными, а для нахождения $n$-мерного нормального вектора $\tilde{x}^*$ – решения скорректированной системы (70) приходится решать задачу безусловной минимизации (1) с $n$ переменными.

Можно предложить иной способ коррекции несовместной системы (I), при котором вместо приведенных двух задач безусловной минимизации решается только одна задача безусловной минимизации в пространстве $\mathbb{R}^m$. Задачу коррекции системы (I) с одновременным нахождением решения представим как задачу нахождения нормального решения совместной системы, которая получена из системы (I) с помощью дополнительных переменных $\tilde{b}_1 \in \mathbb{R}_+^{m_1}$, $\tilde{b}_2 \in \mathbb{R}^{m_2}$:

$$A_{11}x_1 + A_{12}x_2 + \tilde{b}_1 \ge b_1, \quad A_{21}x_1 + A_{22}x_2 + \tilde{b}_2 = b_2, \quad x_1 \ge 0_{n_1}, \quad \tilde{b}_1 \ge 0_{m_1}. \tag{71}$$

Эта система всегда совместна. Поэтому альтернативная к ней система всегда несовместна и имеет вид

$$A_{11}^{\mathrm{T}}u_1 + A_{21}^{\mathrm{T}}u_2 \le 0_{n_1}, \quad A_{12}^{\mathrm{T}}u_1 + A_{22}^{\mathrm{T}}u_2 = 0_{n_2}, \quad u_1 \le 0_{m_1}, \quad u_2 = 0_{m_2},$$

$$b_1^{\mathrm{T}}u_1 + b_2^{\mathrm{T}}u_2 = \rho > 0, \quad u_1 \ge 0_{m_1}.$$

Решив задачу минимизации невязки этой альтернативной системы

$$\min_{u_1 \in \mathbb{R}_+^{m_1}} \min_{u_2 \in \mathbb{R}^{m_2}} [\|(A_{11}^{\mathrm{T}}u_1 + A_{21}^{\mathrm{T}}u_2)_+\|^2 + \|A_{12}^{\mathrm{T}}u_1 + A_{22}^{\mathrm{T}}u_2\|^2 + \|u_1\|^2 + \|u_2\|^2 + (\rho - b_1^{\mathrm{T}}u_1 - b_2^{\mathrm{T}}u_2)^2]/2,$$

в соответствии с формулой (46) получим нормальное решение системы (71):

$$\tilde{x}_1^* = (A_{11}^{\mathrm{T}}u_1^* + A_{21}^{\mathrm{T}}u_2^*)_+/w_3^*, \quad \tilde{x}_2^* = A_{12}^{\mathrm{T}}u_1^* + A_{22}^{\mathrm{T}}u_2^*/w_3^*,$$

$$\tilde{b}_1^* = u_1^*/w_3^*, \quad \tilde{b}_2^* = u_2^*/w_3^*, \quad w_3^* = \rho - b_1^{\mathrm{T}}u_1^* - b_2^{\mathrm{T}}u_2^*.$$

## 5. СИСТЕМЫ ЛИНЕЙНЫХ УРАВНЕНИЙ

Рассмотрим частный случай, когда в (I) и (II) отсутствуют неравенства. Тогда система (I), определяющая множество $X$, имеет вид

$$Ax = b, \tag{I$_5$}$$

а альтернативная система, определяющая множество $U$, записывается в виде

$$A^{\mathrm{T}}u = 0_n, \quad b^{\mathrm{T}}u = \rho \neq 0.$$

Эту систему удобно представлять в виде

$$\hat{A}^{\mathrm{T}}u = r, \tag{II$_5$}$$

где $\hat{A} = [-A, b]$ и $r^{\mathrm{T}} = [0_n^{\mathrm{T}}, \rho]$, $\rho \neq 0$.

Из теоремы 3 следует альтернатива Фредгольма: всегда разрешима одна и только одна система: либо $(\mathrm{I})_5$, либо $(\mathrm{II})_5$.

Вектор $b$ разложим в сумму ортогональных векторов $b = b^{\|} + b^{\perp}$, где $b^{\|} = \mathrm{pr}(b, \mathrm{im}A)$, $b^{\perp} = (b, \ker A^{\mathrm{T}})$, $Z = \ker A^{\mathrm{T}}$. Если $\|b^{\perp}\| = 0$, то $X \neq \varnothing$ и $U = \varnothing$. Если $\|b^{\perp}\| \neq 0$, то $X = \varnothing$ и $U \neq \varnothing$.

Для выявления разрешимости той или иной системы и одновременого нахождения решения разрешимой системы достаточно найти хотя бы один вектор $x^*$ или $u^*$, решив любую из следующих задач безусловной минимизации:

$$\min_{x \in \mathbb{R}^n} \|b - Ax\|^2/2 = \|b - Ax^*\|^2/2, \quad \min_{u \in \mathbb{R}^m} \|r - \hat{A}^{\mathrm{T}}u\|^2/2 = \|r - \hat{A}^{\mathrm{T}}u^*\|^2/2. \tag{72}$$

Эти задачи можно рассматривать как метод наименьших квадратов, примененный, соответственно, к системам $(\mathrm{I})_5$ и $(\mathrm{II})_5$. Из теорем 1–3 получим $z^* = b - Ax^*$, $w^{*\mathrm{T}} = [w_2^{*\mathrm{T}}, w_3^*] = [r - \hat{A}^{\mathrm{T}}u^*]^{\mathrm{T}}$, $w_2^* = A^{\mathrm{T}}u^*$, $w_3^* = \rho - b^{\mathrm{T}}u^*$. Необходимые и достаточные условия минимума в задачах (72), обычно называемые “нормальными уравнениями”, имеют вид

$$A^{\mathrm{T}}(b - Ax^*) = 0_n, \quad \hat{A}(r - \hat{A}^{\mathrm{T}}u^*) = 0_m. \tag{73}$$

Отсюда следует, что векторы невязок $z^* = b^{\perp} \in \ker A^{\mathrm{T}}$ и $w^* = r^{\perp} \in \ker \hat{A}$. Векторы $x^*$, $u^*$, удовлетворяющие (73), являются псевдорешениями систем $(\mathrm{I})_5$ и $(\mathrm{II})_5$ соответственно.

Если $\|b^{\perp}\| = 0$, то $X \neq \varnothing$, $U = \varnothing$, $b = b^{\|}$, $z^* = 0_m$, $w_3^* \neq 0$ и нормальное решение системы $(\mathrm{I})_5$ выражается формулой

$$\tilde{x}^* = A^{\mathrm{T}}u^*/(\rho - b^{\mathrm{T}}u^*), \tag{74}$$

где $u^*$ – псевдорешение системы $(\mathrm{II})_5$. Подставляя эту формулу в $(\mathrm{I})_5$, после простых преобразований получаем

$$(AA^{\mathrm{T}} + bb^{\mathrm{T}})u^* = \rho b. \tag{75}$$

Если $\|b^{\perp}\| \neq 0$, то $X = \varnothing$, $U \neq \varnothing$ и нормальное решение системы $(\mathrm{II})_5$ представимо в виде

$$\tilde{u}^* = \rho(b - Ax^*)/\|b - Ax^*\|^2, \tag{76}$$

где $x^*$ – псевдорешение системы $(\mathrm{I})_5$. Система $(\mathrm{I})_5$, скорректированная по второй формуле в (69), представима в виде $Ax = Ax^* = b^{\|}$. Ее нормальное решение $x^*$ определяется по формуле (74), в которой $b = b^{\|}$, а $u^*$ находится из второй задачи в (72) при $b = b^{\|}$. Если у матрицы $A$ ранг равен $n$, то скорректированная система имеет единственное решение $x^*$, которое совпадает с ее нормальным решением $\tilde{x}^*$.

Ниже потребуем, чтобы ранг матрицы $A$ размера $m \times n$ был максимальным. Через $A^+$ обозначим $(n \times m)$-матрицу, псевдообратную к $A$. Рассмотрим два специальных случая, когда задачи (1), (2), (5), (6) решаются аналитически.

**Случай 1.** Пусть $\mathrm{rank}A = m$.

Тогда $n \geq m$, $X \neq \varnothing$, $\|b^{\perp}\| = 0$ и $U = \varnothing$, $A^+ = A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}$, $AA^+ = I_m$, строки матрицы $A$ линейно независимы, $(A^{\mathrm{T}})^{\|} = A^+A$ – матрица $n \times n$ проектирования на $\mathrm{im}A^{\mathrm{T}}$, $I_n - (A^{\mathrm{T}})^{\|} = (A^{\mathrm{T}})^{\perp}$ – матрица проектирования на $\ker A$. По теореме Кронекера–Капелли, $\operatorname{rank} A = \operatorname{rank}\hat{A} < \operatorname{rank}[\hat{A}^{\mathrm{T}}, r] = m + 1$. Нормальное решение $\tilde{x}^*$ системы $(\mathrm{I})_5$ можно представить в различных формах:

$$\tilde{x}^* = A^+b = A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}b = \operatorname{pr}(0_n, X) = (A^{\mathrm{T}})^{\|}x, \tag{77}$$

где $x$ – произвольный вектор из $X$.

Матрица $\hat{A}^{\mathrm{T}}$ имеет количество строк большее, чем количество столбцов. Поэтому $(\hat{A}^{\mathrm{T}})^+ = (\hat{A}\hat{A}^{\mathrm{T}})^{-1}\hat{A}$, $(\hat{A}^{\mathrm{T}})^+\hat{A}^{\mathrm{T}} = I_m$, $(\hat{A}^{\mathrm{T}})^{\|} = \hat{A}^{\mathrm{T}}(\hat{A}^{\mathrm{T}})^+$ – квадратная матрица порядка $n + 1$ проектирования на $\operatorname{im}\hat{A}^{\mathrm{T}}$, $(\hat{A}^{\mathrm{T}})^{\perp} = I_{n+1} - (\hat{A}^{\mathrm{T}})^{\|}$. Псевдорешение системы $(\mathrm{II})_5$ запишется в виде

$$u^* = (\hat{A}^{\mathrm{T}})^+r = (\hat{A}\hat{A}^{\mathrm{T}})^{-1}\hat{A}r. \tag{78}$$

Вектор $u^*$ удовлетворяет скорректированной системе $\hat{A}^{\mathrm{T}}u^* = r^{\|}$, где $r^{\|}$ – проекция вектора $r$ на подпространство $\operatorname{im}\hat{A}^{\mathrm{T}}$:

$$r^{\|} = \operatorname{pr}(r, \operatorname{im}\hat{A}^{\mathrm{T}}) = (\hat{A}^{\mathrm{T}})^{\|}r = \hat{A}^{\mathrm{T}}(\hat{A}\hat{A}^{\mathrm{T}})^{-1}\hat{A}r.$$

Вектор $w^*$ – проекция $r$ на ортогональное дополнение $\ker\hat{A}$ к подпространству $\operatorname{im}\hat{A}^{\mathrm{T}}$. Действительно, после ряда очевидных преобразований можно показать, что

$$w^* = (\hat{A}^{\mathrm{T}})^{\perp}r = r^{\perp}.$$

Определим квадратную порядка $m$ невырожденную матрицу $\Phi = (\hat{A}\hat{A}^{\mathrm{T}})^{-1} = (AA^{\mathrm{T}} + bb^{\mathrm{T}})^{-1}$; тогда из (78) следует

$$u^* = \rho\Phi b. \tag{79}$$

Эта формула следует также и из (75), если в (75) предположить, что матрица $\Phi$ существует. Подставляя (79) в формулу (74), получаем

$$\tilde{x}^* = A^{\mathrm{T}}\Phi b/(1 - b^{\mathrm{T}}\Phi b). \tag{80}$$

Так как нормальное решение единственно, то можно приравнять выражения (77) и (80). В результате получим матричное тождество

$$A^{\mathrm{T}}\Phi b/(1 - b^{\mathrm{T}}\Phi b) = A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}b. \tag{81}$$

С помощью оптимального вектора множителей Лагранжа $\mu^* \in \mathbb{R}^m$ в задаче нахождения нормального решения совместной системы $(\mathrm{I})_5$ представим (77) в виде

$$\tilde{x}^* = A^{\mathrm{T}}\mu^*, \quad \mu^* = (AA^{\mathrm{T}})^{-1}b. \tag{82}$$

Согласно формуле (80) имеем

$$\mu^* = \Phi b/(1 - b^{\mathrm{T}}u^*). \tag{83}$$

Из (82) и (83) получаем тождество

$$\Phi b = (1 - b^{\mathrm{T}}\Phi b)(AA^{\mathrm{T}})^{-1}b. \tag{84}$$

Обозначим через $\gamma$ наименьшее собственное значение матрицы $AA^{\mathrm{T}}$. Так как $\operatorname{rank} A = m$, то $\gamma > 0$ и из (82) получим

$$\begin{gathered}\gamma\|\mu^*\|^2 \le \|\tilde{x}^*\|^2 = \mu^{*\mathrm{T}}AA^{\mathrm{T}}\mu^* = \mu^{*\mathrm{T}}b \le \|b\|\|\mu^*\|,\\ \|\mu^*\| \le \|b\|/\gamma, \quad \|\tilde{x}^*\| \le \|b\|/\sqrt{\gamma}.\end{gathered} \tag{85}$$

Пусть вектор $\bar{x}^*$ – проекция $\bar{x}$ на $X$. Тогда из (67) имеем $\bar{x}^* - \bar{x} = \operatorname{pr}(0_n, \bar{X})$. Согласно (65), множество $\bar{X}$ представимо в виде $\bar{X} = \{y \in \mathbb{R}^n : Ay = \bar{b}\}$, где $\bar{b} = b - A\bar{x}$. Вместе с тем $\operatorname{pr}(0_n, \bar{X})$ – вектор из $\overline{X}$ с наименьшей евклидовой нормой, поэтому, как в (77), получим

$$\bar{x}^* - \bar{x} = A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}\bar{b}. \tag{86}$$

Аналогично (82), (85) имеем

$$\bar{x}^* = A^+\bar{b} = A^{\mathrm{T}}\bar{\mu}^*, \quad \bar{\mu}^* = (AA^{\mathrm{T}})^{-1}\bar{b},$$

$$\gamma\|\bar{\mu}^*\|^2 \le \|\bar{x}^*\|^2 = \bar{\mu}^{*\mathrm{T}}AA^{\mathrm{T}}\bar{\mu}^* = \bar{\mu}^{*\mathrm{T}}\bar{b} \le \|\bar{\mu}^*\|\|\bar{b}\|, \quad \|\bar{\mu}^*\| \le \|\bar{b}\|/\gamma,$$

$$\mathrm{dist}(\bar{x}, X) = \|\bar{x}^* - \bar{x}\| = (\bar{\mu}^{\mathrm{T}}\bar{b})^{1/2} \le \|\bar{b}\|/\sqrt{\gamma} = \mathrm{pen}(\bar{x}, X)/\sqrt{\gamma}. \tag{87}$$

Формулы (86) и (87) записываются особенно просто, если $m = 1$, в этом случае $b \in \mathbb{R}^1$, матрица $A$ является $n$-мерной вектор-строкой $a^{\mathrm{T}}$. Формула (86) задает проекцию вектора $\bar{x}$ на плоскость $a^{\mathrm{T}}x = b$, формула (87) определяет расстояние от $\bar{x}$ до этой плоскости, т.е.

$$\bar{x}^* = \bar{x} + a(b - a^{\mathrm{T}}\bar{x})/\|a\|^2, \quad \mathrm{dist}(\bar{x}, X) = |b - a^{\mathrm{T}}\bar{x}|/\|a\|. \tag{88}$$

Если $\bar{x} = 0_n$, то $\bar{x}^* = ab/\|a\|^2$, $\|\bar{x}^*\| = |b|/\|a\|$.

Случай 2. Пусть $\mathrm{rank}A = n$.

Тогда $n \le m$. Пусть $U \ne \varnothing$, $\|b^{\perp}\| = 0$, тогда $X = \varnothing$. Так как столбцы матрицы $A$ линейно независимы, то $A^+ = (A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}$, $A^+A = I_n$, $A^{\|} = AA^+$ – матрица $m \times m$ проектирования на $\mathrm{im}A$, $I_m - A^{\|} = A^{\perp}$ – матрица проектирования на ортогональное подпространство, т.е. на $\ker A^{\mathrm{T}}$. Из теоремы Кронекера–Капелли следует, что $\mathrm{rank}\hat{A} = n + 1$. Псевдорешение системы $(\mathrm{I})_5$ и нормальное решение системы $(\mathrm{II})_5$ запишутся в виде

$$x^* = A^+b, \quad A^+ = (A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}, \quad \tilde{u}^* = \hat{A}(\hat{A}^{\mathrm{T}}\hat{A})^{-1}r = \mathrm{pr}(0_m, U) = (\hat{A})^{\|}u, \tag{89}$$

где $u \in U$. Вектор $x^*$ удовлетворяет скорректированной совместной системе $Ax^* = b^{\|}$, где $b^{\|} = \mathrm{pr}(b, \mathrm{im}A) = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}b$ – проекция вектора $b$ на подпространство $\mathrm{im}A$. Вектор $z^*$ – проекция вектора $b$ на $\ker A^{\mathrm{T}}$ – ортогональное дополнение к $\mathrm{im}A$. Действительно, можно показать, что $z^* = A^{\perp}b = b^{\perp}$.

В (89) матрица $\hat{A}^{\mathrm{T}}\hat{A}$ квадратная порядка $n + 1$, невырожденна и представима в блочном виде:

$$\hat{A}^{\mathrm{T}}\hat{A} = \begin{bmatrix} A^{\mathrm{T}}A & -A^{\mathrm{T}}b \\ -b^{\mathrm{T}}A & b^{\mathrm{T}}b \end{bmatrix}.$$

С помощью формулы Фробениуса определим обратную матрицу и из (89) получим

$$(\hat{A}^{\mathrm{T}}\hat{A})^{-1}r = \rho\beta\begin{bmatrix} -HA^{\mathrm{T}}b \\ 1 + \beta b^{\mathrm{T}}AHA^{\mathrm{T}}b \end{bmatrix},$$
$$\tilde{u}^* = \beta\rho[(1 + \beta b^{\mathrm{T}}AHA^{\mathrm{T}}b)b - AHA^{\mathrm{T}}b], \tag{90}$$

где $\beta = 1/\|b\|^2$, $H = (A^{\mathrm{T}}A - \beta A^{\mathrm{T}}bb^{\mathrm{T}}A)^{-1}$ – квадратная матрица порядка $n$. Приравнивая формулы для $\tilde{u}^*$ из (76) и (90), получаем второе матричное тождество:

$$\frac{[I_m - A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}]b}{\|[I_m - A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}]b\|^2} = \beta[(1 + \beta b^{\mathrm{T}}AHA^{\mathrm{T}}b)b - AHA^{\mathrm{T}}b]. \tag{91}$$

После того как матричные тождества (81), (84) и (91) установлены, их справедливость можно доказать и непосредственно, без привлечения теорем 1–3.

Обозначим через $\eta$ наименьшее собственное значение матрицы $\hat{A}^{\mathrm{T}}\hat{A}$, тогда из (89) следует

$$\tilde{u}^* = \hat{A}\xi^*, \quad \xi^* = (\hat{A}^{\mathrm{T}}\hat{A})^{-1}r,$$

$$\eta\|\xi^*\|^2 \le \|\tilde{u}^*\|^2 = \xi^{*\mathrm{T}}\hat{A}^{\mathrm{T}}\hat{A}\xi^* \le \xi^{*\mathrm{T}}r \le \|\xi^*\|/\rho,$$

$$\|\xi^*\| < 1/(\rho\eta), \quad \|\tilde{u}^*\| \le 1/(\rho\sqrt{\eta}),$$

$$\operatorname{dist}(\bar{u}, U) = \|\bar{u}^* - \bar{u}\| \le \|r - \hat{A}^{\mathrm{T}}\bar{u}\|/\sqrt{\eta}. \tag{92}$$

Используя (68) и последнюю формулу в (89), получаем выражение для проекции вектора $\bar{u}$ на непустое множество $U$:

$$\bar{u}^* = \operatorname{pr}(\bar{u}, U) = \bar{u} + \operatorname{pr}(0_m, \bar{U}) = \bar{u} + \hat{A}(\hat{A}^{\mathrm{T}}\hat{A})^{-1}\bar{r}.$$

## 6. СИСТЕМЫ ЛИНЕЙНЫХ НЕРАВЕНСТВ

Пусть система, определяющая множество $X$, имеет вид

$$Ax \ge b, \tag{I$_6$}$$

где $A$ – матрица $m \times n$, $b \in \mathbb{R}^m$, $\|b\| \ne 0$. Альтернативная система, определяющая множество $U$, записывается в виде

$$\hat{A}^{\mathrm{T}}u = r, \quad u \ge 0_m, \tag{II$_6$}$$

где, как и в разд. 5, расширенная матрица имеет вид $\hat{A} = [-A, b]$ и $r^{\mathrm{T}} = [0_m^{\mathrm{T}}, \rho]$, $\rho > 0$.

Согласно теореме 3, совместна одна и только одна из систем $(\mathrm{I})_6$ или $(\mathrm{II})_6$. Это утверждение для $\rho = 1$ известно как теорема Гейла.

Введенные ранее множества принимают вид

$$Z = \{z \in \mathbb{R}^m_+, A^{\mathrm{T}}z = 0_n\}, \quad W = \{w \in \mathbb{R}^{n+1} : \hat{A}w \le 0_m\},$$

$$\bar{X} = \{y \in \mathbb{R}^n : Ay = \bar{b} = b - A\bar{x}\}, \quad \bar{U} = \{v \in \mathbb{R}^m : \hat{A}^{\mathrm{T}}v = \bar{r} = \bar{r} - \hat{A}^{\mathrm{T}}\bar{u}\},$$

$$\operatorname{pen}(x, X) = \|(b - Ax)_+\|, \quad \operatorname{pen}(u, U) = \|r - \hat{A}^{\mathrm{T}}u\|.$$

Векторы $x^*$, $u^*$ определяются из задач

$$\min_{x \in \mathbb{R}^n} \|(b - Ax)_+\|^2/2 = \|(b - Ax^*)_+\|^2/2, \quad \min_{u \in \mathbb{R}^m_+} \|r - \hat{A}^{\mathrm{T}}u\|^2/2 = \|r - \hat{A}^{\mathrm{T}}u^*\|^2/2. \tag{93}$$

Необходимые и достаточные условия минимума в задачах (93) имеют вид

$$A^{\mathrm{T}}(b - Ax^*)_+ = 0_n, \quad \hat{A}(r - \hat{A}^{\mathrm{T}}u^*) \le 0_m, \quad D(u^*)[\hat{A}(r - \hat{A}^{\mathrm{T}}u^*)] = 0_m, \quad u^* \ge 0_m.$$

Отсюда следует, что векторы минимальных невязок $z^* \in Z$ и $w^* \in W$.

Из теорем 1–3 получим $z^* = (b - Ax^*)_+$, $w^{*\mathrm{T}} = [w_2^{*\mathrm{T}}, w_3^*] = [r - \hat{A}^{\mathrm{T}}u^*]^{\mathrm{T}}$, $w_2^* = A^{\mathrm{T}}u^*$, $w_3^* = \rho - b^{\mathrm{T}}u^*$; если $X \ne \varnothing$, то $w_3^* > 0$ и нормальное решение системы $(\mathrm{I})_6$ выражается следующим образом:

$$\tilde{x}^* = A^{\mathrm{T}}u^*/(\rho - b^{\mathrm{T}}u^*). \tag{94}$$

При $\rho = 1$ эта формула была получена в [8]. Так как $\tilde{x}^* \in X$, то из (94) вытекает аналог формулы (75): $(AA^{\mathrm{T}} + bb^{\mathrm{T}})u^* \ge \rho b$.

Если $X = \varnothing$, то $z^* > 0_m$ и нормальное решение системы $(\mathrm{II})_6$ представимо в виде

$$\tilde{u}^* = \rho(b - Ax^*)_+/\|(b - Ax^*)_+\|^2 = \rho b^{\perp}/\|b^{\perp}\|^2. \tag{95}$$

Ниже потребуем, чтобы ранг матрицы $A$ размера $m \times n$ был максимальным. Через $A^+$ обозначим $(n \times m)$-матрицу, псевдообратную к $A$.

**Случай 1.** Пусть $X \ne \varnothing$.

Тогда $U = \varnothing$. Решив задачу (2) (вторую задачу в (93)), из (94) определим нормальный вектор $\tilde{x}^* \in X$, поэтому $A\tilde{x}^* \geq b$. Предположим, что в этой системе неравенств в точке $\tilde{x}^*$ первые $s$ условий выполняются как равенства, остальные $c = m - s$ – как строгие неравенства. В соответствии с этим разобьем матрицу $A$ и векторы $b$, $u$ : $A^{\mathrm{T}} = [A_s^{\mathrm{T}}, A_c^{\mathrm{T}}]$, $b^{\mathrm{T}} = [b_s^{\mathrm{T}}, b_c^{\mathrm{T}}]$, $u^{\mathrm{T}} = [u_s^{\mathrm{T}}, u_c^{\mathrm{T}}]$. Рассмотрим систему

$$A_s x = b_s. \tag{96}$$

Пусть $s \leq n$ и ранг $A_s$ равен $s$, тогда система (96) разрешима и ее нормальное решение, согласно (77), можно представить в виде

$$\tilde{x} = A_s^+ b_s = \mathrm{pr}(0_n, X_s),$$

где $A_s^+ = A_s^{\mathrm{T}}(A_s A_s^{\mathrm{T}})^{-1}$, $X_s = \{x \in \mathbb{R}^n : A_s x = b_s\}$. Так как $\tilde{x}^* = \tilde{x}$, то получаем

$$A^{\mathrm{T}}u^*/(\rho - b^{\mathrm{T}}u^*) = A_s^{\mathrm{T}}u_s^*/(\rho - b_s^{\mathrm{T}}u_s^*) = A_s^{\mathrm{T}}(A_s A_s^{\mathrm{T}})^{-1}b_s.$$

Если все ограничения в точке $\tilde{x}^*$ оказались активными, то приходим к формулам предыдущего раздела. Если $b \leq 0_m$, то $\tilde{x}^* = 0_n$. При $b < 0_m$ в точке $\tilde{x}^*$ нет активных ограничений; кроме того, если $\mathrm{rank}A = m$, то $u^* = 0_m$.

Рассмотрим задачу нахождения нормального решения $\tilde{x}^* = \mathrm{pr}(0_n, X)$. Выпишем для этой задачи функцию Лагранжа $L(x, \mu) = \|x\|^2/2 + \mu^{\mathrm{T}}(b - Ax)$ и условия Куна–Таккера

$$\tilde{x}^* = A^{\mathrm{T}}\mu^*, \quad A\tilde{x}^* \geq b, \quad \mu^{*\mathrm{T}}(A\tilde{x}^* - b) = 0, \quad \mu^* \geq 0_m.$$

Представим оптимальный вектор множителей Лагранжа как $\mu^{*\mathrm{T}} = [\mu_s^{*\mathrm{T}}, \mu_c^{*\mathrm{T}}]$, где $\mu_s^* = (A_s A_s^{\mathrm{T}})^{-1}b_s$, $\mu_c^* = 0_c$. Здесь для простоты считается, что первые $s$ ограничений активны, т.е. $A_s\tilde{x}^* = \bar{b}_s$, $A_c\bar{x}^* < \bar{b}_c$. Тогда $\mu^{*\mathrm{T}}AA^{\mathrm{T}}\mu^* = \mu_s^{*\mathrm{T}}A_s A_s^{\mathrm{T}}\mu_s^*$. Пусть $\bar{\gamma}$ – наименьшее собственое значение матрицы $A_s A_s^{\mathrm{T}}$. Очевидно, что $\gamma \leq \bar{\gamma}$, где $\gamma$ – наименьшее собственное значение матрицы $AA^{\mathrm{T}}$. Согласно условию $\mathrm{rank}A = m$ имеем $\bar{\gamma} \geq \gamma > 0$ и

$$\bar{\gamma}\|\mu_s^*\|^2 \leq \mu_s^{*\mathrm{T}}A_s A_s^{\mathrm{T}}\mu_s^* = \mu_s^{*\mathrm{T}}b_s \geq 0, \quad \|\mu_s^*\| \leq \|b\|/\bar{\gamma},$$
$$\|\tilde{x}^*\|^2 = \|A_s^{\mathrm{T}}\mu_s^*\|^2 = \mu_s^{*\mathrm{T}}b_s \leq \|b_s\|^2/\bar{\gamma}.$$

Для вычисления проекции вектора $\bar{x}$ на множество $X$ воспользуемся результатами разд. 4: учтем, что здесь $\mu^* = (AA^{\mathrm{T}})^{-1}\bar{b}$, $\bar{b} = b - A\bar{x}$, и получим

$$\begin{aligned}[\mathrm{dist}(\bar{x}, X)]^2 &= \|\bar{x}^* - \bar{x}\|^2 = [\mathrm{pr}(\bar{x}, X)]^2 = [\mathrm{pr}(0_n, \bar{X})]^2 = \mu^{*\mathrm{T}}(b - A\bar{x}) \leq \\ &\leq \|\mu_s^*\|\|(b_s - A_s\bar{x})_+\| \leq \|(b_s - A_s\bar{x})_+\|/\bar{\gamma} \leq \|(b - A\bar{x})_+\|/\bar{\gamma}.\end{aligned} \tag{97}$$

Оценки (87), (92) и (97) аналогичны неравенствам Хоффмана (см., например, [2], [3]). Заметим, что в полученных формулах есть существенные отличия. Во-первых, конкретизированы коэффициенты $\gamma$, $\eta$ и $\bar{\gamma}$, во-вторых, даны точные формулы для вычисления расстояния между $\bar{x}^*$ и $\bar{x}$, между $\bar{u}^*$ и $\bar{u}$. В приведенном изложении совершенно очевидны утверждения о том, что $\gamma$, $\bar{\gamma}$ не зависят от $b$, а $\eta$ не зависит от $r$ (см. теорему 10.1 в [3], лемму 35.5 в [2]).

**Случай 2.** Пусть $X = \varnothing$.

Тогда $\|z^*\| \neq 0$, система $(\mathrm{I})_6$ несовместна, среди компонент вектора $Ax^* - b$ есть отрицательные. Скорректированная система будет иметь вид

$$Ax \geq b^{\parallel} = b - (b - Ax^*)_+.$$

Следовательно, если $(b - Ax^*)^i \le 0$, то $(b^{||})^i = b^i$, т.е. $i$-я компонента вектора $b$ не меняется. Если $(b - Ax^*)^i > 0$, то $(b^{||})^i = (Ax^*)^i$ и вместо $b^i$ берется $(Ax^*)^i$, что обеспечивает допустимость вектора $x^*$ в скорректированной задаче.

## 7. ЗАДАЧА О РАЗДЕЛЯЮЩИХ ГИПЕРПЛОСКОСТЯХ

Представим $A$, $b$, $u$ и $z$ в виде

$$A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, \quad u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \quad z = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix},$$

где $A_1$, $A_2$ – матрицы, соответственно, размеров $k \times n$, $l \times n$; векторы $b_1, u_1, z_1 \in \mathbb{R}^k$ и $b_2, u_2, z_2 \in \mathbb{R}^l$; $k + l = m$. Считаем, что множество $X$ состоит из двух непустых множеств

$$X_1 = \{x \in \mathbb{R}^n : A_1 x \ge b_1\}, \quad X_2 = \{x \in \mathbb{R}^n : A_2 x \ge b_2\}$$

таких, что $X_1 \cap X_2 = \varnothing$. Поставим задачу об отыскании гиперплоскости, строго разделяющей $X_1$ и $X_2$.

Введем скалярный параметр $\alpha \in [0, 1]$.

**Теорема 6** (о семействе параллельных разделяющих гиперплоскостей). *Пусть полиэдры $X_1$ и $X_2$ непусты, их пересечение $X = X_1 \cap X_2 = \varnothing$; решением первой задачи в* (93) *является $x^*$, вектор минимальных невязок $z^* = (b - Ax^*)_+$ имеет своими составляющими $z_1^* = (b_1 - A_1x^*)_+$ и $z_2^* = (b_2 - A_2x^*)_+$. Тогда верно следующее*:

1) *семейство параллельных гиперплоскостей, разделяющих множества $X_1$ и $X_2$, может быть представлено двумя эквивалентными формулами:*

$$(z_1^*)^{\mathrm{T}}(A_1x - b_1) + \alpha\|z^*\|^2 = 0, \tag{98}$$

$$(z_2^*)^{\mathrm{T}}(b_2 - A_2x) + (\alpha - 1)\|z^*\|^2 = 0, \tag{99}$$

*причем если $0 < \alpha < 1$, то эти гиперплоскости строго разделяют $X_1$ и $X_2$;*

2) *если в качестве $\alpha$ взять*

$$\alpha^* = \|z_1^*\|^2/\|z^*\|^2, \tag{100}$$

*то точка $x^*$ лежит на разделяющей гиперплоскости, соответствующей этому значению $\alpha$;*

3) *расстояние $d$ между разделяющими гиперплоскостями, соответствующими $\alpha = 0$ и $\alpha = 1$, определяется по формуле*

$$d = \|z^*\|^2/\|A_1^{\mathrm{T}}z_1^*\|.$$

**Доказательство.** Из теоремы 1 следует, что

$$A^{\mathrm{T}}z = 0_n, \quad b^{\mathrm{T}}z^* = \|z^*\|^2.$$

Умножая первое соотношение слева на произвольный вектор $x \in \mathbb{R}^n$ и вычитая из него второе, получаем

$$z^{*\mathrm{T}}(Ax - b) + \|z^*\|^2 = 0. \tag{101}$$

Определим линейную функцию $\varphi(x, \alpha)$ от переменной $x \in \mathbb{R}^n$ и параметра $\alpha$ двумя эквивалентными способами:

$$\varphi(x, \alpha) = (z_1^*)^{\mathrm{T}}(A_1x - b_1) + \alpha\|z^*\|^2, \tag{102}$$

$$\varphi(x, \alpha) = (z_2^*)^{\mathrm{T}}(b_2 - A_2x) + (\alpha - 1)\|z^*\|^2. \tag{103}$$

Условие $\varphi(x, \alpha) = 0$ при каждом фиксированном $\alpha \in [0, 1]$ определяет гиперплоскость, разделяющую множества $X_1$ и $X_2$, так как если $\alpha \ge 0$ и $x \in X_1$, то, согласно (102), $\varphi(x, \alpha) \ge 0$, а если $\alpha \le 1$, $x \in X_2$, то, согласно (103), имеем $\varphi \le 0$. Таким образом, получено семейство параллельных гиперплоскостей, которые при $0 < \alpha < 1$ строго разделяют $X_1$ и $X_2$.

Проекция $\bar{x}^*$ точки $x^*$ на разделяющую гиперплоскость (98) согласно (88) определяется по формуле

$$\bar{x}^* = x^* + A_1^{\mathrm{T}} z_1^* [(z_1^*)^{\mathrm{T}}(b_1 - A_1 x^*) - \alpha \|z^*\|^2] / \|A_1^{\mathrm{T}} z_1^*\|^2.$$

Здесь $(z_1^*)^{\mathrm{T}}(b_1 - A_1 x^*) = \|z_1^*\|^2$. Поэтому если $\alpha$ задается формулой (100), то $\bar{x}^* = x^*$, т.е. $x^*$ принадлежит разделяющей гиперплоскости (98).

Аналогично, подставляя $\alpha^*$ из (100) в (99), учитывая, что $1 - \alpha^* = \|z_2^*\|^2 / \|z^*\|^2$, получаем, что $x^*$ принадлежит разделяющей гиперплоскости (99).

Обозначим через $\mathrm{pr}(\alpha)$ проекцию начала координат на гиперплоскость (98). Тогда согласно (88) имеем

$$\mathrm{pr}(\alpha) = A_1^{\mathrm{T}} z_1^* [(z_1^*)^{\mathrm{T}} b_1 - \alpha \|z^*\|^2] / \|A_1^{\mathrm{T}} z_1^*\|^2.$$

После простейших вычислений имеем $d = \|\mathrm{pr}(1) - \mathrm{pr}(0)\| = \|z^*\|^2 / \|A_1^{\mathrm{T}} z_1^*\|$. Теорема доказана.

Доказательство п. 1) теоремы 6 близко к доказательству теоремы Еремина (см. [2, теорема 10.1]), которое основано на использовании теоремы об альтернативах. В теореме из [2] разделяющая гиперплоскость в обозначениях настоящей работы представлена в следующих эквивалентных видах:

$$(u_1^*)^{\mathrm{T}}(A_1 x - b_1) + \rho/2 = 0, \quad (u_2^*)^{\mathrm{T}}(b_2 - A_2 x) - \rho/2 = 0,$$

где $u_1^*$ и $u_2^*$ – произвольное решение системы вида

$$A_1^{\mathrm{T}} u_1 + A_2^{\mathrm{T}} u_2 = 0_n, \quad b_1^{\mathrm{T}} u_1 + b_2^{\mathrm{T}} u_2 = \rho > 0, \quad u_1 \ge 0_k, \quad u_2 \ge 0_l. \tag{104}$$

Таким образом, для построения разделяющей гиперплоскости согласно теореме 6 требуется решить задачу безусловной минимизации невязки несовместной системы (I) в пространстве $\mathbb{R}^n$, а согласно теореме из [2] требуется решить совместную систему (104) с $m$ неизвестными.

## 8. ТЕОРЕМЫ ЖОРДАНА И ШТИМКЕ ОБ АЛЬТЕРНАТИВАХ

Теорема об альтернативах Жордана утверждает, что разрешима одна и только одна из следующих систем: либо

$$Ax > 0_m, \tag{I$_8$}$$

либо

$$A^{\mathrm{T}} u = 0_n, \quad u \ge 0_m, \quad \|u\|_1 > 0. \tag{II$_8$}$$

Этот результат непосредственно не следует из теоремы 3. Преобразуем системы $(\mathrm{I})_6$ и $(\mathrm{II})_6$, положив $b = \rho e_m$, где $e_m$ – единичный $m$-мерный вектор. Тогда системы $(\mathrm{I})_6$ и $(\mathrm{II})_6$ примут вид

$$Ax \ge \rho e_m, \tag{I$_8^*$}$$

$$A^{\mathrm{T}} u = 0_n, \quad u \ge 0_m, \quad \rho \|u\|_1 = \rho. \tag{II$_8^*$}$$

Система $(\mathrm{I})_8$ разрешима одновременно с системой $(\mathrm{I})_8^*$, т.е. если разрешима система $(\mathrm{I})_8$, то по ее решению $x'$ можно определить параметр $\rho$, равный минимальной компоненте вектора $Ax'$. При этом $\rho$ вектор $x'$ удовлетворяет системе $(\mathrm{I})_8^*$. Обратно: если система $(\mathrm{I})_8^*$ имеет какое-либо решение, то очевидно, что это решение будет удовлетворять системе $(\mathrm{I})_8$. Аналогично, если $u'$ – решение системы $(\mathrm{II})_8$, то вектор $u = u'/\|u'\|_1$ есть решение системы $(\mathrm{II})_8^*$. Справедливо и обратное. Итак, альтернатива в теореме Жордана может быть заменена на альтернативу, которая представлена системами $(\mathrm{I})_8^*$ и $(\mathrm{II})_8^*$, задающими замкнутые множества. Поэтому, в отличие от $(\mathrm{I})_8$ и $(\mathrm{II})_8$, для альтернативных систем $(\mathrm{I})_8^*$ и $(\mathrm{II})_8^*$ могут существовать нормальные решения и их можно искать на основании результатов теоремы 3.

Заметим, что в последнем равенстве системы $(\mathrm{II})_8^*$ можно разделить обе части на ρ, однако при этом надо учесть сомножитель $\rho^2$ в выражении $\mathrm{pen}(x, X) = [\|A^\mathrm{T}u\|^2 + \rho^2(1 - \|u\|_1)^2]^{1/2}$, чтобы обеспечить выполнение утверждения (46). Только после этого все результаты разд. 6 переносятся на случай систем $(\mathrm{I})_8^*$ и $(\mathrm{II})_8^*$. Формулы (94), (95), в частности, принимают вид

$$\tilde{x}^* = A^\mathrm{T}u^*/[\rho(1-\|u^*\|_1)], \quad \tilde{u}^* = \rho(\rho e_m - Ax^*)_+/\|(\rho e_m - Ax^*)_+\|^2.$$

Близкая по форме теорема об альтернативах Штимке утверждает, что разрешима одна и только одна из следующих систем: либо $Ax \ge 0_m$, $\|Ax\| > 0$, либо $A^\mathrm{T}u = 0_n$, $u > 0_m$.

Как и в случае теоремы Жордана, альтернативные системы в теореме Штимке можно заменить системами, для которых может существовать нормальное решение. Введем вектор дополнительных переменных $\xi \in \mathbb{R}^m$, тогда первую систему в теореме Штимке можно переписать в виде

$$Ax - \xi = 0_m, \quad \|\xi\|_1 = \rho, \quad \xi \ge 0_m. \qquad (\mathrm{I})_8^{**}$$

Эта система разрешима одновременно с первой системой из теоремы Штимке. Система $(\mathrm{I})_8^{**}$ является частным случаем общей системы (I) из разд. 2 и поэтому альтернативная к ней система примет вид

$$A^\mathrm{T}u = 0_n, \quad -u + e_m\sigma \le 0_m, \quad \rho\sigma = \rho > 0.$$

Эту систему можно переписать более кратко:

$$A^\mathrm{T}u = 0_n, \quad u \ge e_m. \qquad (\mathrm{II})_8^{**}$$

Система $(\mathrm{II})_8^{**}$ разрешима одновременно со второй альтернативной системой в теореме Штимке. К альтернативным системам $(\mathrm{I})_8^{**}$ и $(\mathrm{II})_8^{**}$ теперь может быть применена теорема 3.

## 9. СИСТЕМЫ ЛИНЕЙНЫХ УРАВНЕНИЙ С НЕОТРИЦАТЕЛЬНЫМИ ПЕРЕМЕННЫМИ

Пусть система, определяющая множество $X$, имеет вид

$$Ax = b, \quad x \ge 0_n, \qquad (\mathrm{I})_9$$

где $A$ – матрица $m \times n$, $b \in \mathbb{R}^m$, $\|b\| \ne 0$. Альтернативная система, определяющая множество $U$, записывается в виде

$$A^\mathrm{T}u \le 0_n, \quad b^\mathrm{T}u = \rho > 0, \qquad (\mathrm{II})_9$$

где ρ – произвольная фиксированная константа.

Согласно теореме 3, совместна одна и только одна из систем $(\mathrm{I})_9$ или $(\mathrm{II})_9$. Это утверждение при использовании в $(\mathrm{II})_9$ записи $b^\mathrm{T}u > 0$ известно как лемма Фаркаша.

Вектор $x^*$ или $u^*$ найдем из задач

$$\min_{x \in \mathbb{R}^m_+} \|b - Ax\|^2/2 = \|b - Ax^*\|^2/2, \qquad (105)$$

$$\min_{u \in \mathbb{R}^m} [\|(A^\mathrm{T}u)_+\|^2 + (\rho - b^\mathrm{T}u)^2]/2 = [\|(A^\mathrm{T}u^*)_+\|^2 + (\rho - b^\mathrm{T}u^*)^2]/2. \qquad (106)$$

Необходимые и достаточные условия минимума в задачах (105), (106) имеют вид

$$-A^\mathrm{T}(b - Ax^*) \ge 0_n, \quad D(x^*)[A^\mathrm{T}(b - Ax^*)] = 0_n, \quad x^* \ge 0_n, \quad A(A^\mathrm{T}u^*)_+ - b(\rho - b^\mathrm{T}u^*) = 0_m.$$

Из теорем 1–3 и результатов разд. 4 имеем $b^\perp = z^* = b - Ax^*$, $b^\parallel = Ax^*$, $w^{*\mathrm{T}} = [w_1^{*\mathrm{T}}, w_3^*]$, $w_1^* = (A^\mathrm{T}u^*)_+$, $w_3^* = \rho - b^\mathrm{T}u^*$.

Если $X \neq \varnothing$, то $\|b^\perp\| = 0$, $U = \varnothing$, $w_3^* > 0$ и нормальное решение системы $(\mathrm{I})_9$ выражается через решение задачи (106) следующим образом:

$$\tilde{x}^* = (A^\mathrm{T} u^*)_+/(\rho - b^\mathrm{T} u^*).$$

При этом из условия $\tilde{x}^* \in X$ получаем аналог формулы (75):

$$A(A^\mathrm{T} u^*)_+ + bb^\mathrm{T} u^* = \rho b.$$

Если $\|b^\perp\| \neq 0$, то $X = \varnothing$, $U \neq \varnothing$, $\|z^*\| \neq 0_m$ и нормальное решение системы $(\mathrm{II})_9$ имеет вид

$$\tilde{u}^* = \rho(b - Ax^*)/\|b - Ax^*\|^2 = \rho b^\perp/\|b^\perp\|^2,$$

несовместная система $(\mathrm{I})_9$ в результате оптимальной коррекции примет вид $Ax = b^\parallel$, $x \geq 0_n$ и тогда альтернативная к ней несовместная система запишется следующим образом: $A^\mathrm{T} u \leq 0_n$, $(x^*)^\mathrm{T} A^\mathrm{T} u = \rho > 0$.

## СПИСОК ЛИТЕРАТУРЫ


1. *Черников С.Н.* Линейные неравенства. М.: Наука, 1968.
2. *Еремин И.И.* Теория линейной оптимизации. Екатеринбург: УрО РАН, 1998.
3. *Васильев Ф.П., Иваницкий А.Ю.* Линейное программирование. М.: Факториал, 1998.
4. *Гейл Д.* Теория линейных экономических моделей. М.: Изд-во иностр. лит., 1963.
5. *Разумихин Б.С.* Физические модели и методы теории равновесия в программировании и экономике. М.: Наука, 1975.
6. *Mangasarian O.L.* Nonlinear programming. Philadelphia: SIAM, 1994.
7. *Схрейвер А.* Теория линейного и целочисленного программирования. М.: Мир, 1991.
8. *Dax A.* The relationship between theorems of the alternative, least norm problems, steepest descent directions, and degeneracy: A review // Ann. Operat. Res. 1993. V. 46. № 1. P. 11–60.
9. *Giannessi F.* Theorems of the alternative and optimization // Encyclopedia of Optimization. Dordrecht etc.: Kluwer Acad. Publs., 2001. V. 5. P. 437–444.
10. *Broyden C.G.* On theorems of alternative // Optimizat. Meth. and Software. 2001. V. 16. № 1–4. P. 101–111.
11. *Еремин И.И., Мазуров В.Д., Астафьев Н.Н.* Несобственные задачи линейного и выпуклого программирования. М.: Наука, 1983.
12. *Еремин И.И.* Противоречивые модели оптимального планирования. М.: Наука, 1983.
13. *Еремин И.И.* Двойственность в линейной оптимизации. Екатеринбург: УрО РАН, 2001.
14. *Еремин И.И.* О квадратичных задачах и полноквадратичных задачах выпуклого программирования // Изв. вузов. Матем. 1998. № 12. С. 22–28.
15. *Evtushenko Yu.* Computation of exact gradients in distributed dynamic systems // Optimizat. Meth. and Software. 1998. V. 9. № 1–3. P. 45–75.
16. *Голиков А.И., Евтушенко Ю.Г.* Двойственный подход к решению систем линейных равенств и неравенств // Тр. XII Байкальской междунар. конф. Методы оптимизации и прилож. Пленарные доклады. Иркутск, 2001. С. 91–99.
17. *Голиков А.И., Евтушенко Ю.Г.* Новый метод решения систем линейных равенств и неравенств // Докл. РАН. 2001. Т. 381. № 4. С. 444–447.
18. *Голиков А.И., Евтушенко Ю.Г.* Отыскание нормальных решений в задачах линейного программирования // Ж. вычисл. матем. и матем. физ. 2000. Т.40. № 12. С. 1766–1786.
19. *Голиков А.И., Евтушенко Ю.Г.* Применение теорем об альтернативах к нахождению нормальных решений линейных систем // Изв. вузов. Математика. 2001. № 12 (475). С. 21–31.
20. *Войтов О.Н., Зоркальцев В.И., Филатов А.Ю.* Исследование систем неравенств алгоритмами внутренних точек на задачах поиска допустимых режимов электроэнергетических систем: Препринт № 10. Иркутск: ИСЭМ СО РАН, 1997. 28 с.
21. *Зоркальцев В.И.* Теорема Фаркаша и теория двойственности в линейной оптимизации: Препринт № 9. Иркутск: ИСЭМ СО РАН, 2001. 15 с.