“Динамика неоднородных систем” под ред. Ю.С. Попкова, 2004, с. 107–135.

Голиков А.И., Евтушенко Ю.Г., Ячимович М.
Вычислительный центр Российской академии наук
Москва, Россия

НОРМАЛЬНОЕ РЕШЕНИЕ И РЕГУЛЯРИЗАЦИЯ ЗАДАЧ ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ[1]

1. Введение

Известно, что нахождение нормального решения задачи ЛП тесно связано с методом регуляризации (см., например, [1]–[7]) и методом квадратичной штрафной функции [8]. По-видимому, первой работой, в которой рассматривался вопрос о совпадении решения возмущенной задачи ЛП, возникающей в методе регуляризации, с решением исходной задачи была статья Х. Удзавы, опубликованная в 1958 г. (русский перевод в [9]). Исходная задача ЛП имела вид

$$\max_{x\in X} c^\top x, \qquad X = \{x : Ax \le b\}, \tag{1}$$

а возмущенная задача была представлена в виде

$$\max_{x\in X} \left(c^\top x - \frac{\varepsilon}{2} \|x\|^2 \right). \tag{2}$$

При условии ограниченности множества X и существовании строго внутренней точки в [9] была доказана теорема о существовании положительного числа ε_* такого, что решение $x(\varepsilon)$ задачи (2) при всяком $0 < \varepsilon \le \varepsilon_*$ совпадает с решением исходной задачи ЛП (1). Также была приведена оценка числа ε_*, существенно использовавшая предположения теоремы. В [3] условия на X были заменены условием существования решения задачи ЛП (1). В [10] и [6, 7] независимо от статьи [9] при условии существования решения задачи (1) было установлено совпадение решений возмущенной задачи (2) и исходной (1), приведены новые оценки величины ε_*. В [3, 7, 11] было показано, что двойственной к регуляризованной задаче (2) является вспомогательная задача

$$\min_{u\ge 0} \left[b^\top u + \frac{1}{2\varepsilon}, \|c - A^\top u\|^2 \right], \tag{3}$$

возникающая при применении квадратичного штрафа к двойственной задаче ЛП

$$\min_{u\in U} b^\top u, \qquad U = \{u : A^\top u = c,\ u \ge 0\}.$$

Указана связь между решениями задач (2) и (3):

$$x(\varepsilon) = \frac{c - A^\top u(\varepsilon)}{\varepsilon}.$$

В разд. 2 данной статьи приводятся нетрадиционные способы задания задач ЛП, получены соответствующие необходимые и достаточные условия оптимальности, отличные от

[1]Работа выполнена при финансовой поддержке РФФИ (код проекта 03-01-00464), по программе поддержки ведущих научных школ (код проекта НШ-1737.2003.1) и программе сотрудничества между РАН и Черногорской академией наук и искусства.

условий Куна–Таккера. В качестве основных переменных использованы переменные прямой задачи и вектор невязок ограничений двойственной задачи. Эти условия оптимальности позволяют находить нормальное решение прямой задачи и нормальный вектор невязок двойственной задачи ЛП с помощью однократной безусловной максимизации вогнутой функции с числом переменных на единицу больше, чем число переменных прямой задачи.

В разд. 3 с помощью теории двойственности задачи квадратичного программирования для нахождения нормального решения задачи ЛП сводятся к задачам безусловной максимизации гладких вогнутых функций в пространствах меньшей размерности, чем размерность вектора прямых переменных. Изучается множество решений двойственной задачи безусловной максимизации, что позволяет в следующем разделе получить оценки для параметра регуляризации.

В разд. 4 рассмотрены задачи, в которых проведена нетрадиционная регуляризация, отличная от регуляризации по А.Н. Тихонову. Указывается значение параметра регуляризации, начиная с которого решение регуляризованной задачи совпадает с решением исходной задачи ЛП. Минимальное значение этого параметра для некоторых задач ЛП может быть равно нулю или быть отрицательным. Двойственная задача к нетрадиционной регуляризованной задаче приводит к новому варианту метода квадратичных штрафных функций для решения задачи ЛП.

Методы решения задач ЛП из разд. 3 и 4 требуют знаний либо оптимального значения целевой функции, либо значения ε_*, начиная с которого методы регуляризации и квадратичного штрафа позволяют найти нормальное решение. Предлагаемый метод в разд. 2 свободен от этих требований. Нормальное решение прямой задачи ЛП и нормальный вектор оптимальных невязок двойственной задачи ЛП вычисляются по простым формулам в результате однократного решения задачи безусловной максимизации гладкой вогнутой кусочно-квадратичной функции. Число переменных в этой задаче на единицу больше числа переменных прямой задачи ЛП.

В разд. 5 для последовательности задач регуляризации систем линейных уравнений приводится последовательность двойственных задач, у которых число переменных равно числу уравнений. Показано, как по простым формулам по решениям последовательности двойственных задач строится последовательность решений задач регуляризации, сходящаяся к нормальному решению исходной системы уравнений.

2. Нахождение нормального решения задачи ЛП с помощью безусловной максимизации

Рассмотрим прямую и двойственную задачи линейного программирования (ЛП)

$$\min_{x\in X} c^\top x, \qquad X = \{x \in \mathbb{R}^n : Ax = b,\ x \geq 0_n\}, \tag{P}$$

$$\max_{u\in U} b^\top u, \qquad U = \{u \in \mathbb{R}^m : v = c - A^\top u \geq 0_n\}. \tag{D}$$

Здесь A — матрица $m \times n$ ранга m, $m \leq n$, ее дефект $\nu = n - m$; векторы $c, v, x \in \mathbb{R}^n$; $u, b \in \mathbb{R}^m$; через 0_i обозначен нулевой i-мерный вектор. Предполагаем, что эти задачи имеют решение. Множества решений задач (P) и (D) обозначим соответственно через X_* и U_*, а оптимальные векторы через x_* и u_*.

Необходимые и достаточные условия оптимальности (условия Куна–Таккера) для задач (P) и (D) имеют вид

$$Ax - b = 0_m, \qquad c - A^\top u \geq 0_n, \qquad D(x)(c - A^\top u) = 0_n, \qquad x \geq 0_n. \tag{4}$$

Здесь и ниже через $D(z)$ обозначена диагональная матрица, у которой i-й диагональный элемент является i-й компонентой z^i вектора z.

Помимо традиционных необходимых и достаточных условий оптимальности (4) и форм записи (P) и (D) взаимно двойственных задач ЛП будем использовать иные представления, следуя [12]. Для этого введем матрицу K размера $\nu \times n$. Считаем, что строки матрицы K линейно независимы и принадлежат нуль-пространству матрицы A, поэтому натянутое на них пространство $\mathrm{im}\ K^\top$ совпадает с нуль-пространством (ядром) матрицы A. В качестве K можно использовать любую матрицу, ν строк которой образуют базис нуль-пространства матрицы A. Таким образом, $\mathrm{im}\ K^\top$ является ортогональным дополнением к пространству $\mathrm{im}\ A^\top$. Поэтому

$$\mathrm{im}\ K^\top = \ker A, \qquad AK^\top = 0_{m\nu}, \qquad \mathbb{R}^n = \mathrm{im}\ A^\top \oplus \mathrm{im}\ K^\top. \tag{5}$$

Здесь через 0_{ij} обозначена $(i \times j)$-матрица с нулевыми элементами, $W_1 \oplus W_2$ — прямая сумма подпространств W_1 и W_2.

Если матрицу A представить в блочном виде $A = [B \mid N]$, где B не вырождена, то матрицу K можно записать в следующем виде: $K = [-N^\top (B^{-1})^\top \mid I_\nu]$. Здесь и ниже I_s обозначает единичную s-мерную матрицу. Если с помощью преобразований Гаусса–Жордана матрицу A привести к виду $A = [I_m \mid N]$, то матрица K представима в виде $K = [-N^\top \mid I_\nu]$.

В ряде случаев построение матрицы K упрощается. Например, пусть задача ЛП имеет ограничения типа неравенств $Nz \geq b$, где N — матрица $m \times \nu$, $z \in \mathbb{R}^\nu_+$. Вводя дополнительные переменные $\xi \in \mathbb{R}^m_+$, представим вектор $x \in \mathbb{R}^n$ как объединение векторов z, ξ, т.е. $x^\top = [z^\top, \xi^\top]$. Теперь допустимое множество можно записать в том же виде, что и в задаче (P), где матрица $A = [N \mid -I_m]$. Тогда $K = [I_\nu \mid N^\top]$.

Ниже потребуются следующие обозначения, связанные с псевдообратной матрицей. Для произвольной прямоугольной $(m \times n)$ матрицы H полного ранга определим псевдообратную $(n \times m)$ матрицу H^+, $(m \times m)$ матрицу проектирования $H^\|$ на пространство столбцов матрицы H и матрицу проектирования $H^\perp$ на нуль-пространство матрицы $H^\top$. Если ранг H равен n, то $n \leq m$, столбцы матрицы H линейно независимы и

$$H^+ = (H^\top H)^{-1} H^\top, \qquad H^\| = HH^+, \qquad H^\perp = I_m - H^\|. \tag{6}$$

Если ранг H равен m, то $n \geq m$, строки H линейно независимы и

$$H^+ = H^\top (HH^\top)^{-1}, \qquad (H^\top)^\| = H^+ H, \qquad (H^\top)^\perp = I_n - (H^\top)^\|, \tag{7}$$

где $(H^\top)^\|$ есть $(n \times n)$ матрица проектирования на пространство столбцов матрицы $H^\top$, $(H^\top)^\perp$ — матрица проектирования на нуль-пространство матрицы H.

Определим вектор $d \in \mathbb{R}^\nu$ и вектор дополнительных переменных $v \in \mathbb{R}^n$ соотношениями $d = Kc$ и

$$v = c - A^\top u. \tag{8}$$

Введем в рассмотрение два аффинных множества:

$$\bar{X} = \{x \in \mathbb{R}^n : Ax = b\}, \qquad \bar{V} = \{v \in \mathbb{R}^n : Kv = d\}.$$

Всюду ниже через $\bar{x}$ и $\bar{v}$ обозначим произвольные фиксированные n-мерные векторы, удовлетворяющие соответственно условиям $\bar{x} \in \bar{X}$, $\bar{v} \in \bar{V}$. Вектор $\bar{v} \in \bar{V}$ находится согласно формуле (8) с помощью произвольного фиксированного вектора $\bar{u} \in \mathbb{R}^m$

$$\bar{v} = c - A^\top \bar{u}. \tag{9}$$

В простейшем варианте можно взять $\bar{u} = 0$, тогда $\bar{v} = c$. Заметим, что векторы $\bar{x}$, $\bar{v}$ или некоторые их компоненты могут быть отрицательными. В силу того что матрица A имеет ранг m и $n \geq m$, всегда $\bar{X} \neq \emptyset$ и $\bar{V} \neq \emptyset$. Отметим, что формула (8) с помощью матрицы $A^\top$ выражает общее решение неоднородной системы $Kv = d$. А общее решение неоднородной системы $Ax = b$ выражается с помощью матрицы K:

$$x = \bar{x} - K^\top y, \tag{10}$$

где $\bar{x}$ — частное решение системы, а $K^\top y$ — общее решение однородной системы $Ax = 0_m$ и $y \in \mathbb{R}^\nu$.

Далее вместо прямой задачи (P) будем в основном иметь дело со следующей задачей:

$$f_*^1(\bar{v}) = \min_{x \in X} \bar{v}^\top x, \qquad X = \{x \in \mathbb{R}^n : Ax = b,\ x \geq 0_n\}. \tag{P_x}$$

Множества решений задач (P) и (P_x) совпадают, а оптимальные значения целевых функций отличаются на константу, зависящую от выбора $\bar{v}$ и равную нулю, если $\bar{v} = c$. Более того, для любого $x \in \bar{X}$ имеем

$$c^\top x = \bar{v}^\top x + b^\top \bar{u}, \tag{11}$$

где $\bar{v}$ определен через $\bar{u}$ согласно (9).

Для задачи (P_x) приведем двойственную задачу

$$\max_{w \in W} b^\top w, \qquad W = \{w \in \mathbb{R}^m : \bar{v} - A^\top w \geq 0_n\}. \tag{C_w}$$

Между допустимыми множествами U и W задач (D) и (C_w) существует взаимно однозначное соответствие $U = \bar{u} + W$, где $\bar{u}$ определяет по формуле (9) вектор $\bar{v} \in \bar{V}$. Решение задачи (C_w) будем обозначать через w_*, а множество решений задачи — через W_*. Аналогичная связь $U_* = \bar{u} + W_*$ имеет место между множествами решений задач (D) и (C_w).

Используя соотношение (10), задачу (P_x) можно записать в виде стандартной задачи ЛП с ограничениями только типа неравенств

$$\max_{y \in Y} d^\top y, \qquad Y = \{y \in \mathbb{R}^\nu : \bar{x} - K^\top y \geq 0_n\}. \tag{P_y}$$

Решение задачи (P_y) обозначим через y_*. Между допустимыми множествами X и Y, множествами решений X_* и Y_* задач (P_x) и (P_y) существует взаимно однозначное соответствие. Для множеств решений X_*, Y_* это можно представить в виде

$$X_* = \bar{x} - K^\top Y_*, \qquad Y_* = (K^\top)^+(\bar{x} - X_*) = (KK^\top)^{-1}K(\bar{x} - X_*). \tag{12}$$

С помощью матрицы K и переменных v двойственную задачу (D) можно записать в форме, симметричной (P_x):

$$f_*^2(\bar{x}) = \min_{v \in V} \bar{x}^\top v, \qquad V = \{v \in \mathbb{R}^n : Kv = d,\ v \geq 0_n\}. \tag{C_v}$$

Решение задачи (C_v) обозначим через v_*. Существует взаимно однозначное соответствие между допустимыми множествами U и V, между множествами решений U_* и V_* соответственно задач (D) и (C_v). Задачи (P_y) и (C_v) являются взаимно двойственными, Укажем также на связь между множествами решений V_* и W_* задач (C_v) и (C_w):

$$V_* = \bar{v} - A^\top W_*, \qquad W_* = (A^\top)^+(\bar{v} - V_*). \tag{13}$$

Необходимые и достаточные условия оптимальности задач (P_x) и (C_v) могут быть представлены в виде

$$Ax = b, \qquad Kv = d, \qquad \bar{v}^\top x + \bar{x}^\top v = \bar{x}^\top \bar{v}, \qquad x \geq 0_n, \qquad v \geq 0_n. \tag{14}$$

Будем искать нормальное решение системы (14), т.е. рассмотрим следующую задачу квадратичного программирования:

$$\min_{x\in\mathbb{R}^n_+} \min_{v\in\mathbb{R}^n_+} \left[\frac{1}{2}(\|x\|^2 + \|v\|^2) : Ax = b,\ Kv = d,\ \bar{v}^\top x + \bar{x}^\top v = \bar{x}^\top \bar{v}\right]. \tag{15}$$

Для этой задачи выпишем функцию Лагранжа

$$L(x,v,p,q,\alpha) = \frac{1}{2}(\|x\|^2 + \|v\|^2) + p^\top(b - Ax) + q^\top(d - Kv) + \alpha(\bar{v}^\top x + \bar{x}^\top v - \bar{x}^\top \bar{v}), \tag{16}$$

где введены множители Лагранжа $p \in \mathbb{R}^m$, $q \in \mathbb{R}^\nu$, $\alpha \in \mathbb{R}^1$.

Необходимые и достаточные условия минимума функции Лагранжа по $x \in \mathbb{R}^n_+$ и $v \in \mathbb{R}^n_+$ имеют вид

$$\begin{array}{rcl} L_x(x,v,p,q,\alpha) & = & x - A^\top p + \alpha\bar{v} \geq 0_n, \qquad x \geq 0_n, \qquad D(x)L_x = 0_n, \\ L_v(x,v,p,q,\alpha) & = & v - K^\top q + \alpha\bar{x} \geq 0_n, \qquad v \geq 0_n, \qquad D(v)L_v = 0_n. \end{array}$$

Разрешая эти соотношения относительно x и v, получим

$$x = (A^\top p - \alpha\bar{v})_+, \qquad v = (K^\top q - \alpha\bar{x})_+. \tag{17}$$

Здесь и ниже через $(a)_+$ обозначается вектор из $\mathbb{R}^n$ с компонентами $a^i_+ = \max\{a^i, 0\}$, $i = 1,\dots,n$, где a^i есть i-я компонента вектора a.

Подставляя решения (17) в выражение для функции Лагранжа (16), получаем двойственную функцию

$$\widetilde{L}(p,q,\alpha) = b^\top p + d^\top q - \alpha\bar{x}^\top\bar{v} - \frac{1}{2}\left[\|(A^\top p - \alpha\bar{v})_+\|^2 + \|(K^\top q - \alpha\bar{x})_+\|^2\right].$$

Двойственная к (15) задача является задачей безусловной максимизации и состоит в отыскании

$$\max_{p\in\mathbb{R}^m} \max_{q\in\mathbb{R}^\nu} \max_{\alpha\in\mathbb{R}^1} \widetilde{L}(p,q,\alpha). \tag{18}$$

Из теоремы двойственности, примененной к задачам (15), (18), и условий оптимальности (14) получаем следующий результат.

Теорема 1. *Задачи (P_x), (C_v) разрешимы тогда и только тогда, когда разрешима задача безусловной максимизации (18). Для каждого решения $[p^*, q^*, \alpha^*]$ задачи (18) векторы $\tilde{x}_*$ и $\tilde{v}_*$, определяемые по формулам*

$$\tilde{x}_* = (A^\top p^* - \alpha^*\bar{v})_+, \tag{19}$$

$$\tilde{v}_* = (K^\top q^* - \alpha^*\bar{x})_+, \tag{20}$$

являются решениями с минимальной евклидовой нормой соответственно задач (P) и (C_v).

Итак, решение исходной задачи ЛП свелось к решению задачи безусловной максимизации кусочно-квадратичной функции $n+1$ переменных. В отличие от классического метода внешнего квадратичного штрафа, примененного к задаче (P_x), в задаче (18), отсутствует стремящийся к бесконечности коэффициент штрафа.

Далее рассмотрим следующие вопросы:

1) Возможность сведения задачи (18) к двум независимым задачам безусловной максимизации меньшей размерности соответственно в пространствах $\mathbb{R}^{m+1}$ и $\mathbb{R}^{\nu+1}$.

2) Связь задачи (18) с регуляризованной задачей ЛП, с методом внешних квадратичных штрафных функций. Оценки оптимального значения α и пороговых значений параметра регуляризации, коэффициента штрафа.

3. Нормальные решения задач линейного программирования

Условия оптимальности (14) можно расщепить, вводя дополнительно в рассмотрение оптимальные значения целевых функций. Допустим, что известно оптимальное значение целевой функции исходной задачи (P). Тогда согласно (11) оптимальное значение целевой функции задачи (P_x) вычисляется следующим образом: $f_*^1(\bar{v}) = f_* - b^\top \bar{u}$, а оптимальное значение целевой функции задачи (C_v) находится по формуле $f_*^2(\bar{x}) = \bar{x}^\top \bar{v} - f_*^1(\bar{v})$. При этом необходимые и достаточные условия оптимальности (14) распадаются на две независимые группы условий, которые описывают множества решений соответственно задач (P_x) и $C_v)$:

$$\begin{aligned} X_* &= \{x \in \mathbb{R}^n : Ax = b,\ \bar{v}^\top x = f_*^1(\bar{v}),\ x \geq 0_n\}, \\ V_* &= \{v \in \mathbb{R}^n : Kv = d,\ \bar{x}^\top v = f_*^2(\bar{x}),\ v \geq 0_n\}. \end{aligned}$$

Обозначим через $\tilde{x}$, $\tilde{x}'$, $\tilde{x}_*$, $\tilde{v}$, $\tilde{v}'$, $\tilde{v}_*$ проекции начала координат на множества $\bar{X}$, X, X_*, $\bar{V}$, V, V_* соответственно. Векторы $\tilde{x}$ и $\tilde{v}$ являются нормальными решениями соответственно систем $Ax = b$ и $Kv = d$. Проекции $\tilde{x}_*$, $\tilde{v}_*$ называются нормальными решениями задач (P_x) и (C_v) соответственно. Их свойствам и методам нахождения будет уделено основное внимание ниже. Так как $X_* \subseteq X \subseteq \bar{X}$, $V_* \subseteq V \subseteq \bar{V}$, то имеем неравенства

$$\|\tilde{x}\| \leq \|\tilde{x}'\| \leq \|\tilde{x}_*\|, \qquad \|\tilde{v}\| \leq \|\tilde{v}'\| \leq \|\tilde{v}_*\|.$$

Наиболее просто вычисляются векторы $\tilde{x}$, $\tilde{v}$:

$$\begin{aligned} \tilde{x} &= A^+ b = A^\top (AA^\top)^{-1} b, \\ \tilde{v} &= K^+ d = K^\top (KK^\top)^{-1} d. \end{aligned} \tag{21}$$

Наиболее сложно — векторы $\tilde{x}_*$, $\tilde{v}_*$. Трудоемкость вычисления $\tilde{x}'$, $\tilde{v}'$ занимает промежуточное место. Интересны случаи, когда легко вычисляемые векторы совпадают с более трудно вычисляемыми. Например, если окажется, что $\tilde{x} \geq 0_n$, $\tilde{v} \geq 0_n$, то $\tilde{x} = x_*$, $\tilde{v} = v_*$. Если $\tilde{x} \in X$, то $\tilde{x} = \tilde{x}'$. Если $\tilde{x} \in X_*$, то $\tilde{x} = \tilde{x}' = \tilde{x}_*$.

Все задачи отыскания проекций ставятся как задачи квадратичного программирования, в которых ищется минимум функций $\|x\|^2/2$ и $\|v\|^2/2$ на соответствующих множествах. Таким образом, приходим к следующим шести задачам отыскания решений с минимальной евклидовой нормой (нормальные решения):

$$\frac{\|\tilde{x}_*\|^2}{2} = \min_{x \in X_*} \frac{\|x\|^2}{2}, \tag{P_{x_*}}$$

$$\frac{\|\tilde{v}_*\|^2}{2} = \min_{v \in V_*} \frac{\|v\|^2}{2}, \tag{D_{v_*}}$$

$$\frac{\|\tilde{x}'\|^2}{2} = \min_{x \in X} \frac{\|x\|^2}{2}, \tag{$P_{\tilde{x}'}$}$$

$$\frac{\|\tilde{v}'\|^2}{2} = \min_{v \in V} \frac{\|v\|^2}{2}, \qquad (D_{\tilde{v}'})$$

$$\frac{\|\tilde{x}\|^2}{2} = \min_{x \in \bar{X}} \frac{\|x\|^2}{2}, \qquad (P_{\tilde{x}})$$

$$\frac{\|\tilde{v}\|^2}{2} = \min_{v \in \bar{V}} \frac{\|v\|^2}{2}. \qquad (D_{\tilde{v}})$$

Все эти задачи имеют единственные решения, причем двойственные к ним суть задачи безусловной максимизации дифференцируемых функций в пространствах меньших размерностей: $m+1$, $\nu+1$, m, ν, m, ν, соответственно.

Выведем двойственные задачи для (P_{x_*}) и (D_{v_*}). Запишем функции Лагранжа для задач (P_{x_*}) и (D_{v_*}):

$$L(x, p, \beta) = \frac{\|x\|^2}{2} + p^\top(b - Ax) + \beta\left(\bar{v}^\top x - f_*^1(\bar{v})\right), \qquad (22)$$

$$L(v, q, \gamma) = \frac{\|v\|^2}{2} + q^\top(d - Kv) + \gamma\left(\bar{x}^\top v - f_*^2(\bar{x})\right). \qquad (23)$$

Здесь $p \in \mathbb{R}^m$, $q \in \mathbb{R}^\nu$, $\beta \in \mathbb{R}^1$, $\gamma \in \mathbb{R}^1$ — множители Лагранжа.

Будем искать седловые точки функций Лагранжа $L(x, p, \beta)$ и $L(v, q, \gamma)$, решая, соответственно, следующие задачи:

$$\max_{p \in \mathbb{R}^m} \max_{\beta \in \mathbb{R}^1} \min_{x \in \mathbb{R}^n_+} L(x, p, \beta), \qquad (24)$$

$$\max_{q \in \mathbb{R}^\nu} \max_{\gamma \in \mathbb{R}^1} \min_{v \in \mathbb{R}^n_+} L(v, q, \gamma). \qquad (25)$$

Внутренние задачи минимизации решаются аналитически, внешние задачи сводятся к решению задач безусловной максимизации вогнутой дифференцированной кусочно-квадратичной функции. Необходимые и достаточные условия минимума функций Лагранжа по $x \in \mathbb{R}^n_+$ и $v \in \mathbb{R}^n_+$ имеют вид

$$L_x(x, p, \beta) = x - A^\top p + \beta\bar{v} \geq 0_n, \quad x \geq 0_n, \quad D(x)L_x(x, p, \beta) = 0_n, \qquad (26)$$

$$L_v(v, q, \gamma) = v - K^\top q + \gamma\bar{x} \geq 0_n, \quad v \geq 0_n, \quad D(v)L_v(v, q, \gamma) = 0_n. \qquad (27)$$

Разрешая соотношения (26) относительно x и (27) относительно v, получаем

$$x = (A^\top p - \beta\bar{v})_+, \qquad (28)$$

$$v = (K^\top q - \gamma\bar{x})_+. \qquad (29)$$

Подставляя решения (28), (29) соответственно в выражения для функций Лагранжа (22), (23), получаем двойственные функции

$$\widetilde{L}(p, \beta) = b^\top p - \beta f_*^1(\bar{v}) - \frac{\|(A^\top p - \beta\bar{v})_+\|^2}{2}, \qquad (30)$$

$$\widetilde{L}(q, \gamma) = d^\top q - \gamma f_*^2(\bar{x}) - \frac{\|(K^\top q - \gamma\bar{x})_+\|^2}{2}. \qquad (31)$$

Эти функции дифференцируемы, вогнуты. Задача, двойственная к (P_{x_*}), является задачей безусловной максимизации в $(m+1)$-мерном пространстве и состоит в отыскании

$$\max_{p \in \mathbb{R}^m} \max_{\beta \in \mathbb{R}^1} \widetilde{L}(p, \beta). \qquad (D_p)$$

Необходимые и достаточные условия максимума задачи (D_p) имеют вид

$$\begin{array}{rcl} \widetilde{L}_p(p,\beta) & = & b - A(A^\top p - \beta\bar{v})_+ = 0_m, \\ \widetilde{L}_\beta(p,\beta) & = & -f_*^1(\bar{v}) + \bar{v}^\top(A^\top p - \beta\bar{v})_+ = 0. \end{array} \tag{32}$$

Если тройка $[x,\ p,\ \beta]$ — решение задачи (24), то x — решение задачи (P_x), а пара $[p,\ \beta]$ — решение задачи (D_p). Для рассматриваемой задачи квадратичного программирования (P_{x_*}) справедливо и обратное утверждение: если пара $[p,\beta]$ — решение задачи (D_p), то, подставляя эту пару в формулу (28), получим решение x задачи (P_{x_*}) и тройку $[x,\ p,\ \beta]$, являющуюся решением задачи (24). Аналогично, двойственной к (D_{v_*}) является задача безусловной максимизации в $(\nu+1)$-мерном пространстве

$$\max_{q\in\mathbb{R}^\nu}\max_{\gamma\in\mathbb{R}^1}\widetilde{L}(q,\gamma). \tag{D_q}$$

Необходимые и достаточные условия максимума задачи (D_q) имеют вид

$$\begin{array}{rcl} \widetilde{L}_q(q,\gamma) & = & q - K(K^\top q - \gamma\bar{x})_+ = 0_\nu, \\ \widetilde{L}_\gamma(q,\gamma) & = & -f_*^2(\bar{x}) + \bar{x}^\top(K^\top q - \gamma\bar{x})_+ = 0. \end{array} \tag{33}$$

Отметим две особенности задач квадратичного программирования (P_{x_*}), (D_{v_*}) и двойственных к ним задач (D_p), (D_q) соответственно.

1. Переменные прямых задач не входят в постановку двойственных.

2. Двойственные задачи являются задачами безусловной максимизации вогнутых кусочно-квадратичных функций в пространствах размерности $m+1$ и $\nu+1$.

Поэтому вместо решения задач на условный экстремум (P_{x_*}) или (D_{v_*}) целесообразно решать задачи на безусловной экстремум (D_p) или (D_q), имеющие меньшие размерности. Тогда решения исходных задач находятся по простым формулам (28) или (29) соответственно.

Без ограничения общности можно считать, что первые ℓ компонент нормального решения $\tilde{x}_*$ задачи (P_x) и последние r компонент нормального решения $\tilde{v}_*$ задачи (C_v) строго больше нуля. В соответствии с этим предположением представим векторы $\tilde{x}_*$, $\bar{x}$, $\tilde{v}_*$, $\bar{v}$ и матрицы A, K в виде

$$\begin{array}{rclcrcl} \tilde{x}_*^\top & = & \left[[\tilde{x}_*^\ell]^\top,[\tilde{x}_*^s]^\top,[\tilde{x}_*^r]^\top\right], & \qquad & \bar{x}^\top & = & \left[[\bar{x}^\ell]^\top,[\bar{x}^s]^\top,[\bar{x}^r]^\top\right], \\ \tilde{v}_*^\top & = & \left[[\tilde{v}_*^\ell]^\top,[\tilde{v}_*^s]^\top,[\tilde{v}_*^r]^\top\right], & & \bar{v}^\top & = & \left[[\bar{v}^\ell]^\top,[\bar{v}^s]^\top,[\bar{v}^r]^\top\right], \\ A & = & [A_\ell \mid A_s \mid A_r], & & K & = & [K_\ell \mid K_s \mid K_r], \end{array} \tag{34}$$

где $\tilde{x}_*^\ell > 0_\ell$, $\tilde{x}_*^s = 0_s$, $\tilde{x}_*^r = 0_r$, $\tilde{v}_*^\ell = 0_\ell$, $\tilde{v}_*^s = 0_s$, $\tilde{v}_*^r > 0_r$, $\ell, s, m \ge 0$ и $\ell+s+r=n$.

С учетом эквивалентности (26) и (28) и разбиения вектора $\tilde{x}_*$ необходимые и достаточные условия оптимальности (условия Куна–Таккера) задачи (P_{x_*}) можно переписать в развернутом виде:

$$\tilde{x}_*^\ell = A_\ell^\top p - \beta\bar{v}^\ell > 0_\ell, \tag{35}$$
$$\tilde{x}_*^s = 0_s, \qquad A_s^\top p - \beta\bar{v}^s \le 0_s, \tag{36}$$
$$\tilde{x}_*^r = 0_r, \qquad A_r^\top p - \beta\bar{v}^r \le 0_r, \tag{37}$$
$$A_\ell\tilde{x}_*^\ell = b, \quad (\bar{v}^\ell)^\top\tilde{x}_*^\ell = f_*^1(\bar{v}). \tag{38}$$

Вектор $\tilde{v}_*$ является нормальным решением задачи (C_v) и, согласно формуле (13), однозначно определяет решение w_* задачи (C_w). С учетом введенных разбиений векторов $\tilde{x}_*$ и $\tilde{v}_*$ запишем развернутые условия Куна–Таккера для (P_x) и (C_w) в точке $[\tilde{x}_*, \tilde{v}_*, w_*]$:

$$\tilde{x}_*^\ell > 0_\ell, \qquad \tilde{v}_*^\ell = \bar{v}^\ell - A_\ell^\top w_* = 0_\ell, \tag{39}$$
$$\tilde{x}_*^s = 0_s, \qquad \tilde{v}_*^s = \bar{v}^s - A_s^\top w_* = 0_s, \tag{40}$$
$$\tilde{x}_*^r = 0_r, \qquad \tilde{v}_*^r = \bar{v}^r - A_r^\top w_* > 0_r, \tag{41}$$
$$A_\ell \tilde{x}_*^\ell = b. \tag{42}$$

Аналогично записываются развернутые условия Куна–Таккера для задачи (D_{v_*}):

$$\tilde{v}_*^\ell = 0_\ell, \qquad K_\ell^\top q - \gamma \bar{x}^\ell \leq 0_\ell, \tag{43}$$
$$\tilde{v}_*^s = 0_s, \qquad K_s^\top q - \gamma \bar{x}^s \leq 0_s, \tag{44}$$
$$\tilde{v}_*^r = K_r^\top q - \gamma \bar{x}^r > 0_r, \tag{45}$$
$$K_r \tilde{v}_*^r = d, \quad (\bar{x}^r)^\top \tilde{v}_*^r = f_*^2(\bar{x}), \tag{46}$$

а также развернутые условия Куна–Таккера для (C_v) и (P_y):

$$\tilde{v}_*^\ell = 0_\ell, \qquad \tilde{x}_*^\ell = \bar{x}^\ell - K_\ell^\top y_* > 0_\ell, \tag{47}$$
$$\tilde{v}_*^s = 0_s, \qquad \tilde{x}_*^s = \bar{x}^s - K_s^\top y_* = 0_s, \tag{48}$$
$$\tilde{v}_*^r > 0_r, \qquad \tilde{x}_*^r = \bar{x}^r - K_r^\top y_* = 0_r, \tag{49}$$
$$K_r \tilde{v}_*^r = d, \tag{50}$$

вычисленные в точке $[\tilde{v}_*, \tilde{x}_*, y_*]$, где y_* определяется по $\tilde{x}_*$ согласно формуле (12).

Нормальное решение $\tilde{x}_*$ прямой задачи (P_x) назовем *невырожденным*, если оно содержит не менее чем m ненулевых компонент. Аналогично нормальное решение $\tilde{v}_*$ двойственной задачи (C_v) назовем *невырожденным*, если оно содержит не менее чем ν ненулевых компонент.

В принятых выше обозначениях нормальные решения $\tilde{x}_*$ и $\tilde{v}_*$ не вырождены, если $\ell \geq m$ и $r \geq \nu$. Заметим, что случай одновременного выполнения неравенств $\ell > m$ и $r > \nu$ невозможен. Так, если $m < \ell$, то из неравенства $\ell + r \leq n$ имеем $m + r < n$ и, следовательно, $r < \nu$. Если $\nu < r$, то $\ell + r \leq n < r + m$ и, следовательно, $\ell < m$.

Далее будем использовать формулы для псевдообратных матриц и матриц проектирования вида (6)–(7) с заменой матриц A и K на "урезанные" $(m \times \ell)$ матрицу $A_\ell A_\ell$ и $(\nu \times r)$ матрицу K_r.

Теорема 2. *Пусть множества X_* и V_* не пусты. Тогда решения задач (P_{x_*}), (D_{v_*}) существуют и единственны, и если столбцы матрицы A, соответствующие ненулевым компонентам вектора $\tilde{x}_*$, образуют матрицу A_ℓ максимального ранга, то компоненты решения $\tilde{x}_*$ представимы в виде*

$$\tilde{x}_*^\ell = A_\ell^+ b, \qquad \tilde{x}_*^s = 0_s, \qquad \tilde{x}_*^r = 0_r. \tag{51}$$

Если столбцы матрицы K, соответствующие ненулевым компонентам вектора $\tilde{v}_*$, образуют матрицу K_r максимального ранга, то компоненты решения $\tilde{v}_*$ представимы в виде

$$\tilde{v}_*^\ell = 0_\ell, \qquad \tilde{v}_*^s = 0_s, \qquad \tilde{v}_*^r = K_r^+ d. \tag{52}$$

Доказательство. В задачах квадратичного программирования (P_{x_*}) и (D_{v_*}) целевые функции строго выпуклы, множества X_* и V_* не пусты. Поэтому обе задачи имеют единственные решения $\tilde{x}_*$ и $\tilde{v}_*$ и существуют такие множители Лагранжа p, β и q, γ, что выполнены условия оптимальности (26), (32) и (27), (33), а также соответствующие условия Куна–Таккера (35)–(38) и (43)–(46).

С учетом разбиения (34) условия $\tilde{x}_* \in X$, $\tilde{v}_* \in V$ можно записать в виде

$$A_\ell \tilde{x}_*^\ell = b, \qquad K_r \tilde{v}_*^r = d. \tag{53}$$

Если $\ell \leq m$, то по условию теоремы ранг матрицы A_ℓ равен ℓ. Тогда первое уравнение в (53) имеет единственное решение $\tilde{x}_*^\ell$, задаваемое формулой в (51), где $A_\ell^+ = (A_\ell^\top A_\ell)^{-1} A_\ell^\top$ согласно (6). В этом случае вектор b принадлежит пространству столбцов матрицы A_ℓ, и поэтому его проекция на это пространство есть $A_\ell^\| b = A_\ell A_\ell^+ b = b$, вектор $\tilde{x}_*$ является вершиной многогранного допустимого множества. Аналогично, если $\nu \geq r$, то единственное решение $\tilde{v}_*^r$ второго уравнения в (53) задается формулой в (52), где $K_r^+ = (K_r^\top K_r)^{-1} K_r^\top$.

Рассмотрим теперь случай невырожденного решения $\tilde{x}_*$, когда $\ell \geq m$. По условию теоремы ранг A_ℓ равен m. Тогда системы уравнений (35) и (39) имеют единственные решения, которые с учетом (38) представим в виде

$$p = (A_\ell A_\ell^\top)^{-1} b + \beta (A_\ell^\top)^+ \bar{v}^\ell, \qquad w_* = (A_\ell^\top)^+ \bar{v}^\ell, \tag{54}$$

где $(A_\ell^\top)^+ = (A_\ell A_\ell^\top)^{-1} A_\ell$. Подставляя в (35) найденное решение для вектора p, получим

$$\begin{aligned} \tilde{x}_*^\ell &= A_\ell^\top (A_\ell A_\ell^\top)^{-1} b + \beta (A_\ell^\top)^\| \bar{v}^\ell - \beta \bar{v}^\ell = \\ &= A_\ell^\top (A_\ell A_\ell^\top)^{-1} b - \beta (A_\ell^\top)^\perp \bar{v}^\ell = A_\ell^\top (A_\ell A_\ell^\top)^{-1} b. \end{aligned} \tag{55}$$

Здесь использовано свойство $(A_\ell^\top)^\perp \bar{v}^\ell = 0_m$, которое следует из формулы (39), показывающей, что вектор $\bar{v}^\ell$ лежит в пространстве столбцов матрицы $A_\ell^\top$. Таким образом, приходим к первой формуле (51), в которой $A_\ell^+ = A_\ell^\top (A_\ell A_\ell^\top)^{-1} b$.

Аналогично доказывается последняя формула в (52) в случае, когда $r \geq \nu$. При этом используются развернутые условия Куна–Таккера (43)–(46) для задачи (D_{v_*}) и развернутые условия Куна–Таккера (47)–(50) для (C_v) и (P_y). Аналогом формул (54) и (55) являются следующие:

$$\begin{aligned} q &= (K_r K_r^\top)^{-1} d + \gamma y_*, \qquad y_* = (K_r^\top)^+ \bar{x}^r, \\ \tilde{v}_*^r &= K_r^\top (K_r K_r^\top)^{-1} d + \gamma K_r^\top (K_r^\top)^+ \bar{x}^r - \gamma \bar{x}^r = K_r^+ d. \end{aligned}$$

Здесь $(K_r^\top)^+ = (K_r K_r^\top)^{-1} K_r$ и $K_r^+ = K_r^\top (K_r K_r^\top)^{-1}$. Теорема доказана. $\square$

Доказанная теорема утверждает, что при достаточно естественных предположениях ненулевые компоненты решений задач (P_{x_*}), (D_{v_*}) совпадают с нормальным решением соответствующих систем уравнений (53).

Заметим, что, в отличие от задачи (P_{x_*}), двойственная к ней задача (D_p) имеет неединственное решение. Естественно возникает вопрос о нахождении среди всех решений задачи (D_p) минимального значения β_* множителя Лагранжа β. Тогда в двойственной задаче (D_p) можно зафиксировать $\beta \geq \beta_*$ и решать задачу максимизации двойственной функции $\widetilde{L}(p, \beta)$ только по переменным p. При этом пара $[p, \beta]$ является решением задачи (D_p), тройка $[\tilde{x}_*, p, \beta]$ — решением задачи (24), и нормальное решение $\tilde{x}_*$ задачи (P_x) определится по формуле (28).

Для нахождения минимального β обратимся к условиям Куна–Таккера для задачи (P_{x_*}). Среди решений системы (35)–(38) найдем такие множители Лагранжа $[p, \beta]$, что β является минимальным, т.е. приходим к задаче линейного программирования

$$\beta_* = \inf_{\beta \in \mathbb{R}^1} \inf_{p \in \mathbb{R}^m} \{\beta : A_\ell^\top p - \beta \bar{v}^\ell = \tilde{x}_*^\ell,\ A_s^\top p - \beta \bar{v}^s \leq 0_s,\ A_r^\top p - \beta \bar{v}^r \leq 0_r\}. \tag{56}$$

Целевая функция в этой задаче может быть не ограничена снизу. В этом случае $\beta_* = -\infty$. Это означает, что для любого фиксированного значения β существует решение p задачи $\max\limits_{p \in \mathbb{R}^m} \widetilde{L}(p, \beta)$ и тройка $[\tilde{x}_*, p, \beta]$ есть решение задачи (24).

Аналогичная ситуация и в задаче (D_q), которая, в отличие от задачи (D_{v_*}), имеет неединственное решение. Так же возникает вопрос о нахождении минимального значения γ_* множителя Лагранжа γ. Тогда, зафиксировав $\gamma \geq \gamma_*$, можно решать задачу максимизации двойственной функции $\widetilde{L}(q, \gamma)$ только по переменным q. Минимальный множитель Лагранжа γ находится из условий Куна–Таккера (43)–(46) для задачи (D_{v_*}). То есть требуется найти такую пару $[q, \gamma]$, что множитель γ является минимальным:

$$\gamma_* = \inf_{\gamma \in \mathbb{R}^1} \inf_{q \in \mathbb{R}^\nu} \{\gamma : K_\ell^\top q - \gamma \bar{x}^\ell \leq 0_\ell,\ K_s^\top q - \gamma \bar{x}^s \leq 0_s,\ K_r^\top q - \gamma \bar{x}^r = \tilde{v}_*^r\}. \tag{57}$$

Для этой задачи так же возможен случай $\gamma_* = -\infty$. Тогда в задаче (D_q) можно зафиксировать любое значение γ.

Из приведенных результатов и теоремы 1 из [6] вытекает следующая

Теорема 3. *Пусть задача (P_x) разрешима. Тогда существуют такие β_* и γ_*, определяемые в* (56) *и* (57), *что имеют место следующие утверждения:*
1) *при любом фиксированном $\beta \geq \beta_*$ пара $[p, \beta]$, где p — решение задачи*

$$\max_{p \in \mathbb{R}^m} \widetilde{L}(p, \beta), \tag{58}$$

определяет нормальное решение $\tilde{x}_*$ задачи (P_x) по формуле

$$\tilde{x}_* = (A^\top p - \beta \bar{v})_+;$$

2) *при любом фиксированном $\gamma \geq \gamma_*$ пара $[q, \gamma]$, где q — решение задачи*

$$\max_{q \in \mathbb{R}^\nu} \widetilde{L}(q, \gamma), \tag{59}$$

определяет нормальное решение $\tilde{v}_*$ задачи (C_v) по формуле

$$\tilde{v}_* = (K^\top q - \gamma \bar{x})_+.$$

Определим значение β_* по формуле

$$\beta_* = \begin{cases} \max\limits_{\ell+s+1 \leq i \leq n} \dfrac{\left[A_r^\top (A_\ell^\top)^+ \tilde{x}_*^\ell\right]^i}{[\tilde{v}_*^r]^i}; & \\ -\infty, & \text{если} \quad \|\tilde{v}_*\| = 0. \end{cases} \tag{60}$$

Теорема 4. *Пусть нормальное решение $\tilde{x}_*$ задачи (P_x) не вырождено и ранг матрицы A_ℓ равен m. Тогда тройка $[x, p, \beta]$, где*

$$x = \tilde{x}_*, \qquad p = (A_\ell A_\ell^\top)^{-1} b + \beta w_*, \qquad \beta \geq \beta_*, \tag{61}$$

является седловой точкой функции Лагранжа $L(x, p, \beta)$ для задачи (P_{x_*}), при этом пара $[p, \beta]$ является решением двойственной задачи (D_p) и выполнены неравенства

$$A_s^\top (A_\ell^\top)^+ \tilde{x}_*^\ell \le 0_s, \tag{62}$$

$$A_r^\top (A_\ell^\top)^+ \tilde{x}_*^\ell \le \beta \tilde{v}_*^r. \tag{63}$$

Доказательство. Из (39)–(41) определим $\bar{v}^\ell$, $\bar{v}^s$, $\bar{v}^r$, подставим соответственно в соотношения (35)–(37). В результате получим

$$\tilde{x}_*^\ell = A_\ell^\top (p - \beta w_*), \tag{64}$$

$$A_s^\top (p - \beta w_*) \le 0_s, \tag{65}$$

$$A_r^\top (p - \beta w_*) \le 0_r. \tag{66}$$

Из (64) однозначно определяется вектор p, который с учетом условия $A_\ell \tilde{x}_*^\ell = b$ представим в виде

$$p = (A_\ell A_\ell^\top)^{-1} b + \beta w_*. \tag{67}$$

Подставляя это выражение для p в (65) и (66), приходим к неравенствам (62) и (63). Итак, условия Куна–Таккера (35)–(38) для задачи (P_{x_*}) сводятся к (62), (63), (38), где $\tilde{x}_*^\ell > 0_\ell$, $\tilde{x}_*^s = 0_s$, $\tilde{x}_*^r = 0_r$, а множители Лагранжа p выражаются формулой (67) и линейно зависят от β. Из (63) следует, что при любом $\beta \ge \beta_*$, где β_* задается формулой (60), тройка $[x, p, \beta]$, определяемая в (61), удовлетворяет условиям Куна–Таккера для задачи (P_{x_*}) и, следовательно, является седловой точкой функции Лагранжа для задачи (P_{x_*}). Тогда пара $[p, \beta]$ является решением двойственной задачи (D_p) к задаче (P_{x_*}). Теорема доказана. $\square$

Из доказательства теоремы 4 следует, что задача (56) нахождения β_* может быть разрешена явно, если $\ell \ge m$ и ранг матрицы A_ℓ равен m. При выполнении условий теоремы 4 оптимальное значение целевой функции выражается формулой (60), а оптимальный вектор p определяется формулой из (61).

Значение γ_* определим следующей формулой:

$$\gamma_* = \begin{cases} \max\limits_{1 \le i \le \ell} \dfrac{\left[K_\ell^\top (K_r^\top)^+ \tilde{v}_*^r\right]^i}{[\tilde{x}_*^\ell]^i}; & \\ -\infty, & \text{если} \quad \|\tilde{x}_*\| = 0. \end{cases} \tag{68}$$

Аналогом теоремы 4 в случае задачи (D_q) будет следующая

Теорема 5. *Пусть нормальное решение $\tilde{v}_*$ задачи (C_v) не вырождено и ранг матрицы K_r равен ν. Тогда тройка $[v, q, \gamma]$, где*

$$v = \tilde{v}_*, \qquad q = (K_r K_r^\top)^{-1} d + \gamma y_*, \qquad \gamma \ge \gamma_*, \tag{69}$$

является седловой точкой функции Лагранжа $L(v, q, \gamma)$ для задачи (D_{v_*}), при этом пара $[q, \gamma]$ является решением двойственной задачи (D_q) и выполнены неравенства

$$K_\ell^\top (K_r^\top)^+ \tilde{v}_*^r \le \gamma \tilde{x}_*^\ell, \qquad K_s^\top (K_r^\top)^+ \tilde{v}_*^r \le 0_s.$$

Доказательство совершенно аналогично доказательству теоремы 4 и основано на использовании развернутых условий Куна–Таккера (43)–(46) для задачи (D_{v_*}) и условий (47)–(50) для задач (C_v) и (P_y).

При выполнении условий теоремы 5 задача (57) нахождения γ_* может быть явно разрешена. При этом γ_* определяется формулой (68), а оптимальный вектор q находится из (69).

В простейшем случае, когда A_ℓ и K_r — квадратные невырожденные матрицы, $\ell = m$, $s = 0$, $r = \nu$, нормальные решения $\tilde{x}_*$ и $\tilde{v}_*$ являются невырожденными базисными решениями. Формулы (60) и 68) при этом упрощаются:

$$\begin{aligned} \beta_* &= \max_{m+1\le i\le n} \frac{\left[A_r^\top (A_\ell^\top)^{-1}\tilde{x}_*^\ell\right]^i}{[\tilde{v}_*^r]^i} \\ \gamma_* &= \max_{1\le i\le m} \frac{\left[K_\ell^\top (K_r^\top)^{-1}\tilde{v}_*^r\right]^i}{[\tilde{x}_*^\ell]^i}. \end{aligned} \tag{70}$$

4. Регуляризованные задачи линейного программирования

Для задач линейного программирования (P_x) и (C_v) введем регуляризованные задачи

$$\min_{x\in X}\left[\beta\bar{v}^\top x + \frac{\|x\|^2}{2}\right], \qquad X = \{x \in \mathbb{R}^n : Ax = b,\ x \ge 0_n\}, \tag{$P_x(\beta)$}$$

$$\min_{v\in V}\left[\gamma\bar{x}^\top v + \frac{\|v\|^2}{2}\right], \qquad V = \{v \in \mathbb{R}^n : Kv = d,\ v \ge 0_n\}. \tag{$D_v(\gamma)$}$$

Здесь $\beta, \gamma \in \mathbb{R}^1$ — некоторые параметры. В отличие от регуляризации по А.Н. Тихонову (см., например, [2, 4, 5]) здесь параметры регуляризации β и γ умножаются на целевые функции, и в ряде случаев они могут принимать нулевые или отрицательные значения.

Задачи $(P_x(\beta))$, $(D_v(\gamma))$ можно представить, соответственно, в виде

$$\min_{x\in X}\frac{1}{2}\ \left[\|x+\beta\bar{v}\|^2 - \beta^2\|\bar{v}\|^2\right], \qquad \min_{v\in V}\frac{1}{2}\ \left[\|v+\gamma\bar{x}\|^2 - \gamma^2\|\bar{x}\|^2\right],$$

т.е. в виде задач проектирования векторов $-\beta\bar{v}$ и $-\gamma\bar{x}$ на множества соответственно X и V.

Выведем двойственную задачу к $(P_x(\beta))$. Для этого выпишем функцию Лагранжа задачи $(P_x(\beta))$

$$L(x,p) = \beta\bar{v}^\top x + \frac{\|x\|^2}{2} + p^\top(b - Ax), \tag{71}$$

где $p \in \mathbb{R}^m$ — множители Лагранжа.

Отметим, что функция Лагранжа $L(x,p)$ отличается от определенной в (22) для задачи (P_{x_*}) функции Лагранжа $L(x,p,\beta)$ на величину $-\beta f_*^1(\bar{v})$, являющуюся константой при фиксированном β. Необходимые и достаточные условия минимума по $x \in \mathbb{R}^n_+$ функции Лагранжа $L(x,p)$ имеют вид, полностью совпадающий с (26). Следовательно, и формула для определения точки минимума совпадает с (28). Подставляя ее в функцию Лагранжа (71), получим двойственную функцию

$$\widetilde{L}(p) = b^\top p - \frac{\|(A^\top p - \beta\bar{v})_+\|^2}{2}. \tag{72}$$

Эта функция дифференцируема, вогнута по p. Двойственная к $(P_x(\beta))$ задача является задачей безусловной максимизации и состоит в отыскании

$$\max_{p\in\mathbb{R}^m}\widetilde{L}(p). \tag{$D_p(\beta)$}$$

Необходимые и достаточные условия максимума в этой задаче имеют вид

$$\widetilde{L}_p(p) = b - A(A^\top p - \beta\bar{v})_+ = 0_m.$$

Аналогично выводится двойственная к $(D_v(\gamma))$ задача, являющаяся задачей безусловной максимизации в ν-мерном пространстве:

$$\max_{q\in\mathbb{R}^\nu} \widetilde{L}(q), \qquad (D_q(\gamma))$$

где двойственная функция $\widetilde{L}(q)$ содержит параметр γ и имеет вид

$$\widetilde{L}(q) = d^\top q - \frac{\|(K^\top q - \gamma\bar{x})_+\|^2}{2}.$$

Сравнивая $\widetilde{L}(p)$ в (72) с $\widetilde{L}(p,\beta)$ в (30), замечаем, что задачи $(D_p(\beta))$ и (D_p) отличаются на величину $-\beta f_*^1(\bar{v})$. Следовательно, при фиксированном β решения задач $(D_p(\beta))$ и (58) совпадают. Аналогичный факт справедлив и для задач $(D_q(\gamma))$ и (59): их решения при фиксированных γ совпадают. Поэтому следующая важная теорема является простой переформулировкой теоремы 3.

Теорема 6. *Пусть задача (P_x) разрешима. Тогда существуют такие β_* и γ_*, определяемые в* (56) *и* (57), *что имеют место следующие утверждения:*
1) *при любом фиксированном $\beta \geq \beta_*$ решение $p(\beta)$ задачи $(D_p(\beta))$ определяет нормальное решение $\tilde{x}_*$ задачи (P_x) по формуле*

$$\tilde{x}_* = [A^\top p(\beta) - \beta\bar{v}]_+; \qquad (73)$$

2) *при любом фиксированном $\gamma \geq \gamma_*$ решение $q(\gamma)$ задачи $(D_q(\gamma))$ определяет нормальное решение $\tilde{v}_*$ задачи (C_v) по формуле*

$$\tilde{v}_* = [K^\top q(\gamma) - \gamma\bar{x}]_+. \qquad (74)$$

Если нормальное решение $\tilde{x}_*$ не вырождено и ранг матрицы A_ℓ равен m, или если нормальное решение $\tilde{v}_*$ не вырождено и ранг матрицы K_r равен ν, то формулировку теоремы 6 можно уточнить аналогично тому, как это было сделано в теоремах 4 и 5. Простым следствием теоремы 4 является следующая

Теорема 7. *Пусть выполнены условия теоремы* 4. *Тогда при любом фиксированном $\beta \geq \beta_*$, где β_* находится по формуле* (60), *решение $p(\beta)$ задачи $(D_p(\beta))$ определяет нормальные решения задач (P_x) и (C_v) соответственно по формулам* (73) *и $\tilde{v}_* = \bar{v} - A^\top w_*$, где*

$$w_* = \frac{p(\beta) - (A_\ell A_\ell^\top)^{-1} b}{\beta}.$$

Следствием теоремы 5 является

Теорема 8. *Пусть выполнены условия теоремы* 5. *Тогда при любом фиксированном $\gamma \geq \gamma_*$, где γ_* находится по формуле* (68), *решение $q(\gamma)$ задачи $(D_q(\gamma))$ определяет нормальные решения задач (C_v) и (P_x) соответственно по формулам* (74) *и $\tilde{x}_* = \bar{x} - K^\top y_*$, где*

$$y_* = \frac{q(\gamma) - (K_r K_r^\top)^{-1} d}{\gamma}.$$

Обозначим через $\mathrm{pr}\,(z, Z)$ проекцию вектора z на множество Z. Из теорем 6–8 следует, что для любых $\beta \geq \beta_*$, $\gamma \geq \gamma_*$ справедливы равенства

$$\begin{aligned} \tilde{x}_* &= \mathrm{pr}\,(-\beta\bar{v}, X) = [A^\top p(\beta) - \beta\bar{v}]_+ = \mathrm{pr}\,(0_n, X_*), \\ \tilde{v}_* &= \mathrm{pr}\,(-\gamma\bar{x}, V) = [K^\top q(\gamma) - \gamma\bar{x}]_+ = \mathrm{pr}\,(0_n, V_*). \end{aligned}$$

Из этих формул сделаем следующее заключение: при $\beta \geq \beta_*$, $\gamma \geq \gamma_*$ нормальные решения задач ЛП (P_x) и (C_v) совпадают с решениями регуляризованных задач $(P_x(\beta))$ и $(D_v(\gamma))$ и совпадают с проекциями начала координат на множества решений X_*, V_* и проекциями $-\beta\bar{v}$ и $-\gamma\bar{x}$ на X и V соответственно. Все эти задачи суть задачи квадратичного программирования, двойственные к задачам безусловной максимизации в различных пространствах. Между множествами решений этих задач существует связь, выражаемая простыми формулами. Отметим, что двойственная к (P_{x_*}) задача безусловной максимизации (D_p) имеет на единицу бóльшую размерность, чем двойственная задача $(D_p(\beta))$. Неизвестный параметр β в $(D_p(\beta))$ является переменной (множителем Лагранжа) в задаче (D_p).

Рассмотрим важный частный случай задачи (P_x), для которого $\beta_* \leq 0$. Согласно формуле (60) этот случай имеет место, если выполнены условия теоремы 4, т.е. нормальное решение $\tilde{x}_*$ задачи (P_x) не вырождено и ранг матрицы A_ℓ равен m, и если $\|\tilde{v}_*\| = 0$ или $A_r^\top(A_\ell^\top)^+\tilde{x}_*^\ell \leq 0_r$. Положим $\beta = 0$; тогда регуляризованная задача $(P_x(\beta))$ превращается в задачу нахождения наименьшего по норме вектора $\tilde{x}_*$ из допустимого множества X. Согласно теореме 7 этот вектор одновременно является решением задачи $(P_x(\beta))$ при любом $\beta \geq \beta_*$ и нормальным решением задач (P_x), (P), а также решением задачи (P_{x_*}).

Рассмотрим случай, когда $\beta \geq \beta_*$ и $\beta > 0$. В задаче $(D_p(\beta))$ сделаем замену переменных, положив $w = p/\beta$. Тогда

$$\widetilde{L}(p) = \beta \left[b^\top w - \frac{\beta}{2} \|(A^\top w - \bar{v})_+\|^2 \right],$$

и задача $(D_p(\beta))$ заменяется эквивалентной задачей

$$\max_{w \in \mathbb{R}^m} \left[b^\top w - \frac{\beta}{2} \|A^\top w - \bar{v})_+\|^2 \right], \tag{75}$$

т.е. приходим к методу внешнего квадратичного штрафа, примененному к задаче (C_w). В этом случае из теоремы 6 следует, что вектор $w(\beta)$, полученный в результате максимизации дифференцируемой внешней штрафной функции, определяет по формуле $\tilde{x}_* = \beta[A^\top w(\beta) - \bar{v}]_+$ нормальное решение задачи (P_x) при $\beta > \beta_*$. Из хорошо известных свойств метода внешнего квадратичного штрафа [8] имеем $w(\beta) = p(\beta)/\beta \to w_*$ при $\beta \to +\infty$.

Таким образом, полученные выше соотношения для β_* дают оценки коэффициента штрафа для классического метода внешнего квадратичного штрафа, примененного к задаче (C_w), т.е. к задаче ЛП с ограничениями типа неравенств. Если $\beta_* \leq 0$, то при любом положительном β, и если $\beta_* > 0$, то при любом $\beta \geq \beta_*$ с помощью решения $w(\beta)$ задачи (75) получаем нормальное решение $\tilde{x}_*$ задачи (P_x).

Итак, задачу $(D_p(\beta))$ можно рассматривать как нестандартную реализацию метода внешнего квадратичного штрафа для решения задачи линейного программирования (C_w). Заметим, что задача (75) является двойственной к задаче регуляризации

$$\min_{x \in X} \left[\bar{v}^\top x + \frac{1}{2\beta} \|x\|^2 \right].$$

Полагая $\varepsilon_* = 1/\beta_*$, если $\beta_* > 0$, и $\varepsilon_* = +\infty$, если $\beta_* \leq 0$, где β_* определяется приведенными выше соотношениями, получаем оценку для параметра регуляризации в классической задаче регуляризации

$$\min_{x \in X} \left[\bar{v}^\top x + \frac{\varepsilon}{2} \|x\|^2 \right]. \tag{76}$$

При $0 < \varepsilon \leq \varepsilon_*$ решения задач (76) и (P_x) совпадают.

Еще раз вернемся к задаче безусловной минимизации (18) для одновременного нахождения нормальных решений задач ЛП (P_x) и (C_v). Если в функции $\widetilde{L}(p, q, \alpha)$ зафиксировать скаляр α и вместо задачи (18) рассмотреть упрощенную задачу

$$\max_{p \in \mathbb{R}^m} \max_{q \in \mathbb{R}^\nu} \widetilde{L}(p, q, \alpha), \tag{77}$$

то последняя распадается на две независимые задачи $(D_p(\beta))$ и $(D_q(\gamma))$, в которых $\beta = \gamma = \alpha$. Тогда из теоремы 6 следует

Теорема 9. *Пусть задача (P_x) разрешима. Тогда существует такой скаляр $\alpha^* = \max\{\beta_*, \gamma_*\}$, где β_* и γ_* определяются из (56) и (57), что при любом фиксированном $\alpha \geq \alpha^*$ решение $[p(\alpha), q(\alpha)]$ задачи (77) определяет нормальные решения задач (P_x) и (C_v) соответственно по формулам*

$$\tilde{x}_* = [A^\top p(\alpha) - \alpha \bar{v}]_+, \qquad \tilde{v}_* = [K^\top q(\alpha) - \alpha \bar{x}]_+.$$

5. Нормальные решения систем линейных уравнений и регуляризация

Рассмотрим совместную систему линейных уравнений

$$Ax = b. \tag{78}$$

Здесь A — $(m \times n)$ ненулевая матрица, вектор $b \in \mathbb{R}^m$, $\|b\| \neq 0$. В методе регуляризации рассматривается последовательность задач безусловной минимизации

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} \left(\|b - Ax\|^2 + \varepsilon \|x\|^2 \right) \tag{79}$$

с положительным параметром ε, стремящимся к нулю. Единственное решение $x(\varepsilon)$ задачи (79) при фиксированном ε выражается явно следующей формулой:

$$x(\varepsilon) = (\varepsilon I_n + A^\top A)^{-1} A^\top b. \tag{80}$$

Здесь и ниже через I_k обозначена единичная матрица порядка k. При $\varepsilon \to 0$ решение $x(\varepsilon)$ сходится к нормальному решению системы (78) [13]. В (80) обратная матрица существует при любом ранге матрицы A и любом $\varepsilon > 0$.

Выражение (80) для вычисления $x(\varepsilon)$ можно представить в другом виде, используя формулу Шермана–Моррисона–Вудбери [14]:

$$x(\varepsilon) = \frac{1}{\varepsilon} [I_n - A^\top (\varepsilon I_m + AA^\top)^{-1} A] A^\top b. \tag{81}$$

Отметим, что в этой формуле требуется обращать матрицу порядка m в отличие от формулы (80), в которой обращается матрица порядка n. Ниже будет выведена еще одна формула вычисления $x(\varepsilon)$, в которой требуется обращать матрицу порядка m, но с меньшим числом перемножений матриц, чем в формуле (80).

При фиксированном параметре ε задачу (79) можно рассматривать как метод наименьших квадратов (метод минимизации невязок), примененный к следующей несовместной системе:

$$Ax = b, \qquad -\sqrt{\varepsilon}x = 0_n. \tag{82}$$

Вектор $x(\varepsilon)$ — решение задачи безусловной минимизации (79) — является псевдорешением системы (82). Через $z_1(\varepsilon) = b - Ax(\varepsilon)$ и $z_2(\varepsilon) = \sqrt{\varepsilon}x(\varepsilon)$ обозначим составляющие вектора минимальных невязок

$$z^\top(\varepsilon) = [z_1^\top(\varepsilon), z_2^\top(\varepsilon)]$$

системы (82).

Согласно теореме об альтернативах (см., например [15, 16]) для несовместной системы (82) можно построить совместную альтернативную систему вида

$$A^\top u_1 - \sqrt{\varepsilon}u_2 = 0_n, \qquad b^\top u_1 = \rho > 0. \tag{83}$$

Здесь ρ — фиксированная положительная константа, векторы неизвестных $u_1 \in \mathbb{R}^m$, $u_2 \in \mathbb{R}^n$. Согласно [16] нормальный вектор $\tilde{u}(\varepsilon)^\top = [\tilde{u}_1^\top(\varepsilon), \tilde{u}_2^\top(\varepsilon)]$ альтернативной системы (83) выражается через вектор минимальных невязок $z(\varepsilon)$ по формулам

$$\tilde{u}_1(\varepsilon) = \frac{\rho z_1(\varepsilon)}{\|z(\varepsilon)\|^2}, \qquad \tilde{u}_2(\varepsilon) = \frac{\rho z_2(\varepsilon)}{\|z(\varepsilon)\|^2}. \tag{84}$$

Формально задача безусловной минимизации (79) не имеет функцию Лагранжа, и, следовательно, для нее нельзя непосредственно построить двойственную задачу. Тем не менее, с помощью дополнительных переменных можно ввести искусственные ограничения и получить эквивалентную задачу нелинейного программирования, для которой уже стандартным образом определяется двойственная задача. Следуя [11] и [16], воспользуемся этим приемом для получения задачи, двойственной к (79). Пусть вектор переменных $z \in \mathbb{R}^{m+n}$ имеет разбиение $z^\top = [z_1^\top, z_2^\top]$, где $z_1 \in \mathbb{R}^m$, $z_2 \in \mathbb{R}^n$. Тогда двойственная к (79) задача имеет вид

$$\max_{z \in Z} \left(b^\top z_1 - \frac{1}{2}\|z\|^2 \right), \qquad Z = \{z \in \mathbb{R}^{m+n} : A^\top z_1 - \sqrt{\varepsilon}z_2 = 0_n\}. \tag{85}$$

В свою очередь для этой задачи двойственной является (79). Поэтому задачи (79) и (85) можно назвать взаимно двойственными.

В задаче (85) можно исключить переменные z_2, выразив их через z_1 и подставив $z_2 = (1/\sqrt{\varepsilon})A^\top z_1$ в целевую функцию задачи (85). Тогда приходим к следующей эквивалентной задаче безусловной максимизации строго вогнутой квадратичной функции:

$$\max_{z_1 \in \mathbb{R}^m} \left(b^\top z_1 - \frac{1}{2}\|z_1\|^2 - \frac{1}{2\varepsilon}\|A^\top z_1\|^2 \right). \tag{86}$$

Таким образом, эту задачу можно назвать двойственной к задаче (79). Тогда для взаимно двойственных задач (79) и (86) по теореме слабой двойственности для любых $z_1 \in \mathbb{R}^m$, $x \in \mathbb{R}^n$, $\varepsilon > 0$ справедливо неравенство

$$b^\top z_1 - \frac{1}{2}\|z_1\|^2 - \frac{1}{2\varepsilon}\|A^\top z_1\|^2 \le \frac{1}{2}(\|b - Ax\|^2 + \varepsilon\|x\|^2),$$

а по теореме двойственности оптимальные значения целевых функций этих задач совпадают:

$$b^\top z_1(\varepsilon) - \frac{1}{2}\,\|z_1(\varepsilon)\|^2 - \frac{1}{2\varepsilon}\,\|A^\top z_1(\varepsilon)\|^2 = \frac{1}{2}\left(\|b - Ax(\varepsilon)\|^2 + \varepsilon\|x(\varepsilon)\|^2\right). \qquad (87)$$

Здесь $z_1(\varepsilon)$ — решение задачи (86). Задачи квадратичного программирования (79), (85), (86) имеют единственные решения, которые связаны между собой следующими соотношениями:

$$z_1(\varepsilon) = b - Ax(\varepsilon), \qquad z_2(\varepsilon) = \sqrt{\varepsilon}x(\varepsilon), \qquad (88)$$

$$x(\varepsilon) = \frac{1}{\varepsilon}\,A^\top z_1(\varepsilon). \qquad (89)$$

В этом легко убедиться, подставляя выражение (89) для $x(\varepsilon)$ в необходимые и достаточные условия оптимальности для задачи (79)

$$-A^\top(b - Ax(\varepsilon)) + \varepsilon x(\varepsilon) = 0_n \qquad (90)$$

и подставляя выражение (88) для $z_1(\varepsilon)$ в необходимые и достаточные условия оптимальности для задачи (86)

$$b - z_1(\varepsilon) - \frac{1}{\varepsilon}\,AA^\top z_1(\varepsilon) = 0_m. \qquad (91)$$

Равенство (87) оптимальных значений целевых функций взаимно двойственных задач (79) и (86) с учетом связей (88), (89) между решениями $x(\varepsilon)$ и $z(\varepsilon)$ можно представить в следующих двух видах, в которых присутствуют решения только одной из задач:

$$b^\top z_1(\varepsilon) = \|z(\varepsilon)\|^2, \qquad (92)$$

$$b^\top(b - Ax(\varepsilon)) = \|b - Ax(\varepsilon)\|^2 + \varepsilon\|x(\varepsilon)\|^2. \qquad (93)$$

Из (93) следует $\varepsilon\|x(\varepsilon)\|^2 = z_1(\varepsilon)^\top Ax(\varepsilon)$. Отсюда получаем известный результат применения метода наименьших квадратов: при $\varepsilon = 0$ имеет место $z_1(0) \perp Ax(0)$. Заметим, что если система (78) не совместна, то $z_1(0) \neq 0$.

Задача (86) при любом $\varepsilon > 0$ и любом ранге матрицы A может быть явно разрешена:

$$z_1(\varepsilon) = \varepsilon(\varepsilon I_m + AA^\top)^{-1}b. \qquad (94)$$

Подставляя это выражение в (89), получаем еще одну формулу для вычисления решения $x(\varepsilon)$ задачи (79):

$$x(\varepsilon) = A^\top(\varepsilon I_m + AA^\top)^{-1}b. \qquad (95)$$

В этой формуле требуется обращать матрицу порядка m в отличие от формулы (80), в которой требуется обращать матрицу порядка n. Поэтому если в задаче (78) $m < n$, то для вычисления $x(\varepsilon)$ целесообразно использовать формулу (95) или (89). В случае применения последней формулы требуется решать задачу безусловной оптимизации (86) с m неизвестными.

Выражение (95) для вычисления $x(\varepsilon)$ также можно представить в другом виде с использованием формулы Шермана–Моррисона–Вудбери

$$x(\varepsilon) = \frac{1}{\varepsilon}\,A^\top(I_m - A(\varepsilon I_n + A^\top A)^{-1}A^\top)b. \qquad (96)$$

В этой формуле требуется обращать матрицу порядка n в отличие от формулы (95), в которой обращается матрица порядка m.

В [17] показано, что для всякой матрицы A из системы (78) следующим образом определена псевдообратная матрица A^+:

$$A^+ = \lim_{\varepsilon\to 0}(\varepsilon I_n + A^\top A)^{-1}A^\top = \lim_{\varepsilon\to 0} A^\top(\varepsilon I_m + AA^\top)^{-1} \tag{97}$$

и при любом векторе b (в том числе и при b, делающем систему (78) несовместной) вектор $\tilde{x}_* = A^+b$ является вектором с минимальной нормой среди всех векторов, минимизирующих $\|b - Ax\|^2$.

В [16] предложен другой метод нахождения нормального решения системы (78), основанный на применении теорем об альтернативах. При этом для несовместной альтернативной к (78) системы

$$A^\top u = 0_n, \qquad b^\top u = \rho \neq 0 \tag{98}$$

решается следующая задача минимизации ее невязок:

$$\min_{u\in\mathbb{R}^m} \frac{1}{2}\left[\|A^\top u\|^2 + (\rho - b^\top u)^2\right]. \tag{99}$$

Здесь ρ — произвольная ненулевая константа.

Пусть u_* — решение задачи безусловной минимизации (99). Тогда нормальное решение исходной системы (78) выражается по формуле

$$\tilde{x}_* = \frac{A^\top u_*}{\rho - b^\top u_*}. \tag{100}$$

Если в исходной системе ранг матрицы A равен m, то решение задачи (99) находится аналитически

$$u_* = \rho(AA^\top + bb^\top)^{-1}b. \tag{101}$$

Подставляя эту формулу в (100), получаем еще одно выражение для нормального решения системы (78):

$$\tilde{x}_* = \frac{A^\top(AA^\top + bb^\top)^{-1}b}{1 - b^\top(AA^\top + bb^\top)^{-1}b}. \tag{102}$$

Литература

1. *Тихонов А.Н.* О некорректных задачах оптимального планирования // Журн. вычислит. матем. и матем. физ. 1966. Т. 6. № 1. С. 81–90.

2. *Тихонов А.Н., Арсенин В.Я.* Методы решения некорректных задач. М.: Наука, 1979.

3. *Поляк Б.Т.* Введение в оптимизацию. М.: Наука, 1983.

4. *Васильев Ф.П., Иваницкий А.Ю.* Линейное программирование. М.: Факториал, 1998.

5. *Еремин И.И.* Теория линейной оптимизации. Екатеринбург, УрО РАН, 1998.

6. *Mangasarian O.L., Meyer R.R.* Nonlinear perturbation of linear programs // SIAM J. Control and Optimizat. 1979. V. 17. № 6. P. 745–752.

7. *Mangasarian O.L.* Normal solutions of linear programs // Math. Program. Study. 1984. V. 22. P. 206–216.

8. *Фиакко А., Мак-Кормик Г.* Нелинейное программирование. Методы последовательной безусловной минимизации. М.: Мир, 1972.

9. *Удзава Х.* Итерационные методы вогнутого программирования / К. Дж. Эрроу, Л. Гурвиц, Х. Удзава. Исследования по линейному и нелинейному программированию. М.: Изд-во иностр. лит., 1962. С. 228–245.

10. 1*Еремин И.И., Астафьев Н.Н.* Введение в теорию линейного и выпуклого программирования. М.: Наука, 1976.

11. *Еремин И.И.* О квадратичных и полноквадратичных задачах выпуклого программирования // Известия ВУЗов. Математика. 1998. № 12. С. 22–28.

12. *Golikoy A.I., Evtushenko Yu.G.* Two Parametric Families of LP Problems and Their Applications // Proceedings of Steklov Institute of Mathematics, Suppl. 1. 2002. P. S52–S66.

13. *Тихонов А.Н.* О некорректных задачах линейной алгебры и устойчивом методе их решения // ДАН СССР. 1965. Т. 163. № 4. С. 591–594.

14. *Ортега Дж., Рейнболдт В.* Итерационные методы решения нелинейных систем уравнений со многими неизвестными. М.: Мир, 1975.

15. *Mangasarian O.L.* Nonlinear Prograrnming. New York: McCrow-Hill, 1969.

16. *Голиков А.И., Евтушенко Ю.Г.* Теоремы об альтернативах и их применение в численных методах // Журн. вычислит. матем. и матем. физики. 2003. Т. 43. № 3. С. 354–375.

17. *Альбер А.* Регрессия, псевдоинверсия и рекуррентное оценивание. М.: Наука, 1977.