В. Г. Жадан, О сходимости прямо-двойственного метода Ньютона для задач линейного программирования, *Ж. вычисл. матем. и матем. физ.*, 1999, том 39, номер 3, 431–445

ЖУРНАЛ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ ФИЗИКИ, 1999, том 39, № 3, с. 431–445

УДК 519.9:519.852

О СХОДИМОСТИ ПРЯМО-ДВОЙСТВЕННОГО МЕТОДА НЬЮТОНА ДЛЯ ЗАДАЧ ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ[1)]

© 1999 г. В. Г. Жадан

(117967 Москва, ул. Вавилова, 40, ВЦ РАН)

Поступила в редакцию 31.03.98 г.

Исследуется сходимость прямо-двойственного метода Ньютона для решения задач линейного программирования, в котором шаг перемещения выбирается из условия наискорейшего спуска. Показывается, что если в начальной паре одна из точек совпадает с оптимальной вершиной, то метод позволяет находить решение за одну итерацию. Доказывается также локальная сходимость метода к оптимальному решению за число итераций, не превышающее числа переменных в задаче.

1. ОПИСАНИЕ МЕТОДА

Рассмотрим задачу линейного программирования и двойственную к ней задачу

$$\begin{gathered}\min c^{\mathrm{T}}x, \quad Ax = b, \quad x \ge 0_n, \\ \max b^{\mathrm{T}}u, \quad c - A^{\mathrm{T}}u \ge 0_n,\end{gathered} \tag{1}$$

где c и x – вектор-столбцы из n-мерного евклидова пространства $\mathbb{R}^n$, A есть $(m \times n)$-матрица полного ранга, в которой $m < n$, 0_n – нулевой n-мерный вектор. Допустимое множество в прямой задаче обозначим через X, в двойственной – через U. Предполагается, что множества X и U не пусты и что обе задачи (1) не вырождены. Пара точек $[x, u] \in X \times U$ называется допустимой.

В [1] для решения задач (1) был предложен прямо-двойственный метод, в котором с помощью метода Ньютона отыскиваются точки, удовлетворяющие условиям оптимальности для обеих задач (1). Сходимость некоторых его вариантов была исследована в [1] и [2], однако среди них не было версии метода с полным наискорейшим спуском, когда разрешается выход на границы допустимых множеств. В настоящей работе изучается сходимость именно такого варианта метода, который был предложен в [3] и рассмотрен также в [4].

Пусть $[x_0, u_0]$ – начальная допустимая пара. Итерации в методе описываются следующими рекуррентными соотношениями:

$$x_{k+1} = x_k + \alpha_k \Delta x_k, \quad u_{k+1} = u_k + \alpha_k \Delta u_k. \tag{2}$$

Направления перемещения Δx и Δu в (2) находятся из решения системы линейных уравнений

$$\begin{gathered}D(v(u))\Delta x - D(x)A^{\mathrm{T}}\Delta u = -D(x)v(u), \\ A\Delta x = b - Ax,\end{gathered} \tag{3}$$

где $v(u) = c - A^{\mathrm{T}}u$ и $D(z)$ – диагональная матрица с вектором z на диагонали.

Разобьем в каждой допустимой паре $[x, u]$ все множество индексов от единицы до n на подмножества:

$$J_B^B(x, u) = J^B(x) \cap J_B(u), \quad J_N^N(x, u) = J^N(x) \cap J_N(u), \tag{4}$$

$$J_N^B(x, u) = J^B(x) \cap J_N(u), \quad J_B^N(x, u) = J^N(x) \cap J_B(u). \tag{5}$$

---

1) Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (коды проектов 96-01-01047 и 96-15-96124).

Здесь

$$J^B(x) = \{i: x^i > 0\}, \quad J^N(x) = \{i: x^i = 0\},$$
$$J_B(u) = \{i: v^i(u) = 0\}, \quad J_N(u) = \{i: v^i(u) > 0\}. \tag{6}$$

В соответствии с разбиениями (4)–(6) делятся на подвекторы все n-мерные векторы, связанные с данной парой точек, например x и v, а также матрица A делится на подматрицы, состоящие из ее столбцов.

Если в допустимой паре $[x, u]$ множество индексов $J_B^N(x, u)$ пусто, то такая пара называется регулярной, иначе – нерегулярной. Как показано в [4], в любой регулярной паре решение системы уравнений (3) единственно. В нерегулярной паре система (3) недоопределена. В качестве ее решения берется следующее:

$$\Delta u = Qb_1,$$
$$\Delta x_N^B = (G_N^B)^2 \hat{A}_N^B Q b_1 - x_N^B, \quad \Delta x_N^N = 0, \tag{7}$$
$$\Delta x_B^B = (\hat{A}_B^B A_B^B)^{-1} \hat{A}_B^B (I - \Gamma Q) b_1 - x_B^B.$$

Здесь и ниже I – единичная матрица, крышечка над матрицей обозначает ее транспонирование,

$$G_N^B = D^{1/2}(x_N^B) D^{-1/2}(v_N^B), \quad \Gamma = A_N^B (G_N^B)^2 \hat{A}_N^B, \quad Q = H(\hat{H}\Gamma H)^{-1}\hat{H}, \tag{8}$$
$$b_1 = b - A_B^N \Delta x_B^N.$$

Матрица полного ранга H в (8) берется произвольной, но такой, чтобы она задавала базис в ортогональном дополнении пространства столбцов матрицы A_B^B. Считается также, что Q в (8) является нулевой матрицей, когда число индексов в множестве $J_B^B(x, u)$ равно m, и совпадает с Γ^{-1}, когда оно равно нулю. Если $J_B^N(x, u) \neq \varnothing$, то вектор Δx_B^N находится из решения линейной задачи дополнительности

$$\Omega \Delta x_B^N - p \geq 0,$$
$$\Delta x_B^N \geq 0, \tag{9}$$
$$\langle \Omega \Delta x_B^N - p, \Delta x_B^N \rangle = 0,$$

где $\Omega = \hat{A}_B^N Q A_B^N$, $p = \hat{A}_B^N Q b$, угловые скобки обозначают евклидово скалярное произведение. Изменение $\Delta v = -\hat{A}\,\Delta u$ вектора $v(u)$ оказывается таким, что $\Delta v_B^B = 0$.

Шаг α_k в (2) полагается равным

$$\alpha_k = \max\{\alpha \geq 0: x_k + \alpha \Delta x_k \in X, u_k + \alpha \Delta u_k \in U\}. \tag{10}$$

Данный выбор шага соответствует наибольшему убыванию разности значений целевых функций в задачах (1) при условии сохранения допустимости новых точек x_{k+1} и u_{k+1}.

Если ввести индикаторные векторы y^B и y_N, определив их по формулам

$$y^B = e + D^{-1}(x^B)\Delta x^B, \quad y_N = -D^{-1}(v_N)\Delta v_N, \tag{11}$$

то из (7) получим

$$y_B^B = D^{-1}(x_B^B)(\hat{A}_B^B A_B^B)^{-1} \hat{A}_B^B (I - \Gamma Q) b_1,$$
$$y_N^B = D^{-1}(v_N^B) \hat{A}_N^B Q b_1, \tag{12}$$
$$y_N^N = D^{-1}(v_N^N) \hat{A}_N^N Q b_1.$$

Тогда процесс (2) в исходном пространстве и в пространстве дополнительных двойственных переменных можно переписать в виде

$$(x^B)_{k+1} = D((x^B)_k)[e + \alpha_k((y^B)_k - e)], \quad (v_N)_{k+1} = D((v_N)_k)[e - \alpha_k (y_N)_k]. \tag{13}$$

Здесь и ниже e – вектор со всеми компонентами, равными единице.

Обозначим через y_*^B минимальную компоненту вектора y^B, через y_N^* – максимальную компоненту вектора y_N. Тогда для шага α_k наряду с (10) имеет место формула

$$\alpha_k = \min\left\{\frac{1}{[1-(y_*^B)_k]_+}, \frac{1}{[(y_N^*)_k]_+}\right\}, \qquad (14)$$

где $[a]_+ = \max[0, a]$.

В данной работе исследуется сходимость метода (2) в предположении, что по крайней мере одна из начальных точек (x_0 или u_0) достаточно близка к оптимальному решению. В разд. 2 рассматривается частный случай, когда имеет место совпадение одной из точек x_0 или u_0 с оптимальной вершиной. Показано, что для таких начальных пар метод позволяет находить решение за одну итерацию. Далее, в разд. 3 исследуется случай начальных пар $[x_0, u_0]$, когда x_0 и u_0 принадлежат граням допустимых множеств, прилегающим к оптимальным вершинам, причем одна из начальных точек лежит на ребре. Приводятся условия, при которых имеет место сходимость за две итерации. Наконец, в разд. 4 доказывается локальная сходимость метода (2) за число итераций, не превышающее число переменных в прямой задаче.

Для упрощения записи точки x_{k+1} и u_{k+1}, вычисляемые по формуле (2) при $x_k = x$, $u_k = u$, обозначаются также через $\bar{x}$ и $\bar{u}$ соответственно.

2. СХОДИМОСТЬ ЗА ОДНУ ИТЕРАЦИЮ

Введем обозначения

$$J_{B_*} = J^B(x_*) = J_B(u_*), \quad J_{N_*} = J^N(x_*) = J_N(u_*),$$

где $[x_*, u_*]$ – пара оптимальных решений задач (1). Считаем для определенности, что базис точки x_* состоит из первых m столбцов матрицы A, т.е. $J_{B_*} = \{1, 2, ..., m\}$, $J_{N_*} = \{m+1, ..., n\}$.

Определение 1. Индекс i в допустимой паре точек $[x, u]$ называется правильным, если

$$i \in J_B^B(x, u) \cap J_{B_*} \quad \text{или} \quad i \in J_N^N(x, u) \cap J_{N_*}.$$

В противном случае он называется неправильным.

Обозначим через $\sigma(x, u)$ множество всех неправильных индексов в допустимой паре точек $[x, u]$, через $\sigma_{B_*}(x, u)$ и $\sigma_{N_*}(x, u)$ – те подмножества множества $\sigma(x, u)$, которые содержатся, соответственно, в J_{B_*} и J_{N_*}. Обозначим также через Υ_k множество тех допустимых пар точек $[x, u]$, из которых метод (2) позволяет решать задачи линейного программирования (1) не более чем за k итераций.

Теорема 1. *Пусть в допустимой паре* $[x, u]$ *точка* u *совпадает с оптимальной вершиной* u_*. *Тогда* $[x, u] \in \Upsilon_1$.

Доказательство. Предположим сначала, что $\sigma_{B_*}(x, u) = \varnothing$. Тогда

$$J_B^B(x, u) = J_{B_*}, \quad \sigma(x, u) = \sigma_{N_*}(x, u) = J_N^B(x, u) \subseteq J_{N_*}.$$

Так как в этом случае число индексов в множестве $J_B^B(x, u)$ равно m, то матрица Q в (7) является нулевой, поэтому из (7) получаем

$$\Delta u = 0, \quad \Delta x_N^N = 0, \quad \Delta x_N^B = -x_N^B.$$

Кроме того, поскольку множество $J_B^N(x, u)$ пусто, то $b_1 = b = A_B^B x_B^B + A_N^B x_N^B$. Следовательно,

$$\Delta x_B^B = (A_B^B)^{-1} A_N^B x_N^B.$$

Тогда, согласно (11) и (12),

$$y_N = 0, \quad y_B^B = D^{-1}(x_B^B)(A_B^B)^{-1} b = D^{-1}(x_B^B)(x_*)_B^B > 0. \qquad (15)$$

Поэтому $\alpha = 1$ и, в соответствии с (13), получаем

$$\bar{u} = u = u_*, \quad \bar{x}_N^B = 0, \quad \bar{x}_B^B = (A_B^B)^{-1} b = (x_*)_B^B,$$

следовательно, $\bar{x} = x_*$.

Если $\sigma_{B_*}(x, u) \neq \varnothing$, то обязательно $\sigma_{N_*}(x, u) \neq \varnothing$, причем, в силу предположения о невырожденности исходной задачи (1), $|\sigma_{N_*}(x, u)| \geq |\sigma_{B_*}(x, u)|$ и $\sigma_{N_*}(x, u) = J_N^B(x, u)$, $\sigma_{B_*}(x, u) = J_B^N(x, u)$. Покажем, что уравнение

$$\Omega \Delta x_B^N - p = 0 \tag{16}$$

имеет решение $\Delta x_B^N > 0$ и, следовательно, именно оно является решением линейной задачи дополнительности (9). Так как матрицы H и A_B^N имеют одинаковое число столбцов, то

$$\Omega = \hat{Z}(\hat{H}\Gamma H)^{-1} Z, \quad p = \hat{Z}(\hat{H}\Gamma H)^{-1} \hat{H} b,$$

где $Z = \hat{H} A_B^N$ – квадратная неособая матрица. Но из равенства

$$\hat{H} b = \hat{H}[A_B^B (x_*)_B^B + A_B^N (x_*)_B^N] = \hat{H} A_B^N (x_*)_B^N$$

вытекает, что $p = \hat{Z}(\hat{H}\Gamma H)^{-1} Z (x_*)_B^N$. Поэтому уравнение (16) преобразуется к виду

$$\hat{Z}(\hat{H}\Gamma H)^{-1} Z(\Delta x_B^N - (x_*)_B^N) = 0.$$

В силу невырожденности матрицы $\hat{Z}(\hat{H}\Gamma H)^{-1} Z$, оно имеет единственное решение $\Delta x_B^N = (x_*)_B^N$, подставляя которое в (7), получаем, с учетом того, что $b - A_B^N \Delta x_B^N = A_B^B (x_*)_B^B$ и $Q(b - A_B^N \Delta x_B^N) = 0$,

$$\Delta u = 0, \quad \Delta x_N^B = -x_N^B,$$

$$\Delta x_B^B = (\hat{A}_B^B A_B^B)^{-1} \hat{A}_B^B A_B^B (x_*)_B^B - x_B^B = (x_* - x)_B^B.$$

Таким образом, для компонент векторов y_B^B и y_N имеют место формулы (15) и, следовательно, $\bar{x} = x_*$, $\bar{u} = u_*$.

Теорема 2. *Пусть в допустимой паре* $[x, u]$ *точка* x *совпадает с оптимальной вершиной* x_*. *Тогда* $[x, u] \in \Upsilon_1$.

Доказательство. Предположим, что $\sigma_{N_*}(x, u) = \varnothing$. Тогда

$$J_N^N(x, u) = J_{N_*}, \quad \sigma(x, u) = \sigma_{B_*}(x, u) = J_N^B(x, u) \subseteq J_{B_*}.$$

Поскольку $J_B^N(x, u) = \varnothing$, то решение системы (3) единственно. Беря $\Delta x = 0$, $\Delta u = u_* - u$, получаем, что пара $[\Delta x, \Delta u]$ удовлетворяет данной системе. Тогда имеем

$$\Delta v_B^B = 0, \quad \Delta v_N^B = -v_N^B, \quad \Delta v_N^N = (v_* - v)_N^N.$$

Поэтому

$$y_N^B = e, \quad y_N^N = e - D^{-1}(v_N^N)(v_*)_N^N < e,$$

откуда вытекает, что $\alpha = 1$ и, стало быть, $\bar{x} = x_*$, $\bar{u} = u_*$.

Допустим теперь, что $\sigma_{N_*}(x, u) \neq \varnothing$. Тогда $\sigma_{B_*}(x, u) \neq \varnothing$, причем из-за невырожденности двойственной задачи (1) будет $|\sigma_{B_*}(x, u)| \geq |\sigma_{N_*}(x, u)|$ и $\sigma_{N_*}(x, u) = J_B^N(x, u)$, $\sigma_{B_*}(x, u) = J_N^B(x, u)$.

Покажем, что система неравенств

$$\Omega \Delta x_B^N - p \geq 0$$

имеет тривиальное решение $\Delta x_B^N = 0$. Для этого достаточно убедиться, что у вектора p все компоненты строго отрицательны.

Так как $\hat{A}^B_B(u_* - u) = 0$, то вектор $u_* - u$ принадлежит нуль-пространству матрицы $\hat{A}^B_B$ и, следовательно, $u_* - u = H\xi$ для некоторого вектора $\xi \in \mathbb{R}^s$, где $s = m - |J^B_B(x, u)|$. Умножая это равенство слева на матрицу $\hat{A}^B_N$ и учитывая, что $s = |J^B_N(x, u)|$, получаем

$$\Delta v^B_N = (v - v_*)^B_N = \hat{A}^B_N H\xi = Z_1\xi,$$

где $Z_1 = \hat{A}^B_N H$ – квадратная матрица порядка s. Она не вырождена, так как иначе столбцы матрицы H принадлежали бы нуль-пространству всей матрицы A^B, что невозможно.

Имеем $p = \hat{A}^N_B H(\hat{Z}_1(G^B_N)^2 Z_1)^{-1}\hat{H}b$. Отсюда, используя равенство $\hat{H}b = \hat{H}A^B_N(x_*)^B_N$, вытекающее из $b = A^B_B(x_*)^B_B + A^B_N(x_*)^B_N$, приходим к следующему:

$$p = \hat{A}^N_B H Z_1^{-1}(G^B_N)^{-2}(x_*)^B_N = \hat{A}^N_B H Z_1^{-1} v^B_N = \hat{A}^N_B H\xi = \hat{A}^N_B(u_* - u) = (v - v_*)^N_B = (-v_*)^N_B < 0.$$

Поскольку $\Delta x^N_B = 0$, то $\Delta x = 0$, $\Delta u = u_* - u$ являются решением системы (3), и, аналогично сказанному выше, опять убеждаемся, что $\bar{x} = x_*$, $\bar{u} = u_*$.

Лемма 1. *Пусть в допустимой паре* $[x, u]$ *все неправильные индексы принадлежат либо множеству* J_{B_*}, *либо множеству* J_{N_*}. *Тогда* $[x, u] \in \Upsilon_1$.

Доказательство следует из утверждений двух предыдущих теорем, так как при сделанном предположении одна из точек (x или u) является оптимальной.

Таким образом, если в допустимой паре $[x, u]$ только один неправильный индекс, то обязательно $[x, u] \in \Upsilon_1$.

3. СХОДИМОСТЬ ЗА ДВЕ ИТЕРАЦИИ

Рассмотрим теперь вопрос о конечной сходимости метода (2), предполагая, что в начальной паре $[x, u]$ либо точка x, либо u принадлежит относительной внутренности ребра допустимого множества, имеющего в качестве одного из своих концов оптимальную вершину. Считаем также, что

$$\sigma(x, u) = J^B_N(x, u). \tag{17}$$

Это предположение гарантирует, что пара $[x, u]$ является регулярной. Из него также следует, что обе точки x и u принадлежат граням допустимых множеств X и U, прилегающим, соответственно, к оптимальным вершинам x_* и u_*.

Ниже добавление символа B_* или символа N_* (независимо от того, вверху или внизу) обозначает подмножество множества индексов, которые одновременно входят, соответственно, в множество J_{B_*} или в множество J_{N_*}. Например, x^{B, B_*}_N – совокупность таких компонент x^i вектора x^B_N, что $i \in J^B_N(x, u) \cap J_{B_*}$; A^B_{N, N_*} – подматрица матрицы A^B_N, составленная из таких столбцов a_i, номера которых принадлежат множеству $J^B_N(x, u) \cap J_{N_*}$.

Пусть S^B_B – подпространство, порожденное столбцами матрицы A^B_B, и $(S^B_B)^\perp$ – его ортогональное дополнение. Возьмем оптимальную вершину $u_* \in U$. Из $v^i = v^i_* = 0$ для $i \in J^B_B(x, u)$ следует, что $u_* - u \in (S^B_B)^\perp$. Поэтому

$$u_* - u = H\xi_* \tag{18}$$

для некоторого вектора $\xi_* \in \mathbb{R}^r$, где $r = m - |J^B_B(x, u)|$. Отсюда получаем $v^{B, B_*}_N = \hat{A}^{B, B_*}_N H\xi_*$, и если ввести вектор

$$\psi = D^{1/2}(x)D^{1/2}(v)e,$$

то

$$\psi_N^{B,B_*} = G_N^{B,B_*}\hat{A}_N^{B,B_*}H\xi_*.$$

Рассмотрим вектор $p_N^B = Z_N^B\xi_*$, где матрица Z_N^B имеет вид $G_N^B\hat{A}_N^B H$. Согласно (18),

$$p_N^B = G_N^B\hat{A}_N^B H\xi_* = G_N^B(v_N^B - (v_*)_N^B) = D(\eta_N^B)\psi_N^B, \tag{19}$$

где

$$\eta_N = e - D^{-1}(v_N)(v_*)_N \le e.$$

Из $u \in U$, $u \ne u_*$ следует также, что

$$\langle\psi_N^B, p_N^B\rangle = \sum_{i \in J_N^B(x,u)} x^i v^i \eta^i = \sum_{i \in J_N^B(x,u)} x^i(v^i - v_*^i) = \sum_{i \in J^B(x)} x^i(v^i - v_*^i) = $$
$$= \sum_{i \in J^B(x)} x^i\langle a_i, u_* - u\rangle = \langle b, u_* - u\rangle > 0. \tag{20}$$

Так как $J_B^N(x, u) = \varnothing$, то для вектора y_N^B наряду с (12) имеет место представление

$$y_N^B = D^{-1}(\psi_N^B)PD(\psi_N^B)e = D^{-1}(\psi_N^B)\bar{\psi}_N^B,$$

где $\bar{\psi}_N^B = P\psi_N^B$ и $P = Z_N^B(\hat{Z}_N^B Z_N^B)^{-1}\hat{Z}_N^B$ – матрица ортогонального проектирования.

Лемма 2. *Пусть допустимая пара точек* $[x, u]$ *такова, что выполнено* (17) *и*

$$\left|J_N^{B,B_*}(x,u)\right| = 1, \quad \left|J_{N,N_*}^B(x,u)\right| \ge 1.$$

Тогда

$$\max_{i \in J_{N,N_*}(u)} y^i < \max_{i \in J_N(u)} y^i = \max_{i \in J_N^B(x,u)} y^i = \max_{i \in J_N^{B,B_*}(x,u)} y^i. \tag{21}$$

Доказательство. Считаем для определенности, что неправильными являются индексы $\{m, m+1, \ldots, m+s-1\}$, где $s = |J_N^B(x, u)|$. Тогда при сделанных предположениях

$$J_B^B(x,u) = J_B^{B,B_*}(x,u) = \{1, 2, \ldots, m-1\},$$
$$J_N^N(x,u) = J_{N,N_*}^N(x,u) = \{m+s, \ldots, n\}.$$

Таким образом, подпространство $(S_B^B)^\perp$ является одномерным и вектор $h = u_* - u$ принадлежит этому подпространству. Взяв его в качестве вектора базиса в $(S_B^B)^\perp$, получим, что $H = \{h\}$ и

$$Z_N^B = G_N^B\hat{A}_N^B(u_* - u) = D(\eta_N^B)\psi_N^B = D(\psi_N^B)\eta_N^B = p_N^B.$$

Вычислим вектор y_N. С учетом равенства

$$\hat{H}b = \hat{H}A_N^B x_N^B = \hat{Z}_N^B\psi_N^B = \langle p_N^B, \psi_N^B\rangle$$

из определений (11), (12) и (19) получаем, что

$$y_N = c_1\eta_N, \quad c_1 = \langle\psi_N^B, p_N^B\rangle/\|p_N^B\|^2, \tag{22}$$

причем, согласно (20), $c_1 > 0$. Так как $\eta^m = 1$, $\eta^i < 1$, $i = m+1, \ldots, n$, то через (22) приходим к неравенству (21).

Лемма 3. *Пусть допустимая пара точек* $[x, u]$ *такова, что выполнено* (17) *и*

$$\left|J_{N,N_*}^B(x,u)\right| = 1, \quad \left|J_N^{B,B_*}(x,u)\right| \ge 1.$$

Тогда

$$\min_{i \in J^{B, B_*}(x)} y^i > \min_{i \in J^B(x)} y^i = \min_{i \in J^B_N(x, u)} y^i = \min_{i \in J^B_{N, N_*}(x, u)} y^i. \tag{23}$$

Доказательство. Считаем для определенности, что $J_N^B(x, u) = \{m - s + 2, ..., m + 1\}$, где $s = |J_N^B(x, u)|$. Тогда

$$J^B_B(x, u) = J^{B, B_*}_B(x, u) = \{1, 2, \ldots, m - s + 1\},$$

$$J^N_N(x, u) = J^N_{N, N_*}(x, u) = \{m + 2, \ldots, n\}.$$

Представим матрицу Z_N^B в виде

$$Z_N^B = \begin{bmatrix} Z_N^{B, B_*} \\ Z^B_{N, N_*} \end{bmatrix},$$

где

$$Z_N^{B, B_*} = G_N^{B, B_*}\hat{A}_N^{B, B_*}H, \quad Z^B_{N, N_*} = G^B_{N, N_*}\hat{A}^B_{N, N_*}H.$$

Так как, на основании (19), $\psi_N^{B, B_*} = p_N^{B, B_*}$, то

$$\bar{\psi}^B_N = P[p^B_N + (\psi^B_N - p^B_N)] = p^B_N + Z^B_N(\hat{Z}^B_N Z^B_N)^{-1}\hat{Z}^B_{N, N_*}(\psi^B_{N, N_*} - p^B_{N, N_*}),$$

причем, согласно (19),

$$\psi^B_{N, N_*} - p^B_{N, N_*} > 0.$$

Обозначим $Y = \hat{Z}_N^{B, B_*} Z_N^{B, B_*}$, $\Phi = (\hat{Z}_N^{B, B_*})^{-1}\hat{Z}^B_{N, N_*}$. На основании формулы Моррисона–Шермана–Вудбери получаем

$$(\hat{Z}^B_N Z^B_N)^{-1} = (Y + \hat{Z}^B_{N, N_*}Z^B_{N, N_*})^{-1} = Y^{-1} - Y^{-1}\hat{Z}^B_{N, N_*}(I + Z^B_{N, N_*}Y^{-1}\hat{Z}^B_{N, N_*})^{-1}Z^B_{N, N_*}Y^{-1}.$$

Поэтому

$$\bar{\psi}_N^{B, B_*} = \psi_N^{B, B_*} + \Phi(I + \hat{\Phi}\Phi)^{-1}(\psi^B_{N, N_*} - p^B_{N, N_*}), \tag{24}$$

$$\bar{\psi}^B_{N, N_*} = p^B_{N, N_*} + [I - (I + \hat{\Phi}\Phi)^{-1}](\psi^B_{N, N_*} - p^B_{N, N_*}) = \psi^B_{N, N_*} - (I + \hat{\Phi}\Phi)^{-1}(\psi^B_{N, N_*} - p^B_{N, N_*}). \tag{25}$$

Вычислим матрицу Φ. Из

$$A^B_B(x_*)^B_B + A_N^{B, B_*}(x_*)_N^{B, B_*} = b \tag{26}$$

следует, что

$$\hat{H}b = \hat{H}A_N^{B, B_*}(x_*)_N^{B, B_*} = \hat{Z}_N^{B, B_*}(G_N^{B, B_*})^{-1}(x_*)_N^{B, B_*}. \tag{27}$$

Кроме того, из

$$A^B_B x^B_B + A_N^{B, B_*}x_N^{B, B_*} + A^B_{N, N_*}x^B_{N, N_*} = b \tag{28}$$

получаем

$$\hat{H}b = \hat{Z}_N^{B, B_*}(G_N^{B, B_*})^{-1}x_N^{B, B_*} + \hat{Z}^B_{N, N_*}(G^B_{N, N_*})^{-1}x^B_{N, N_*}. \tag{29}$$

Умножая левую и правую части равенств (27) и (29) на матрицу $(\hat{Z}_N^{B, B_*})^{-1}$, получаем

$$(G_N^{B, B_*})^{-1}(x_*)_N^{B, B_*} = (\hat{Z}_N^{B, B_*})^{-1}\hat{H}b, \tag{30}$$

$$(G_N^{B, B_*})^{-1}x_N^{B, B_*} + \Phi(G^B_{N, N_*})^{-1}x^B_{N, N_*} = (\hat{Z}_N^{B, B_*})^{-1}\hat{H}b. \tag{31}$$

Из (30) и (31) следует, что

$$\Phi(G^B_{N,N_*})^{-1}x^B_{N,N_*} = (G^{B,B_*}_N)^{-1}[(x_*)^{B,B_*}_N - x^{B,B_*}_N].$$

Так как $|J^B_{N,N_*}(x, u)| = 1$, то вектор $(G^B_{N,N_*})^{-1}x^B_{N,N_*}$ является скаляром и

$$(G^B_{N,N_*})^{-1}x^B_{N,N_*} = \psi^B_{N,N_*} = \psi^{m+1}.$$

Поэтому

$$\Phi = -\frac{1}{\psi^B_{N,N_*}}D(\mu^{B,B_*}_N)\psi^{B,B_*}_N, \tag{32}$$

где $\mu^B = e - D^{-1}(x^B)(x_*)^B \le e$. Кроме того,

$$\psi^B_{N,N_*} - p^B_{N,N_*} = D(e - \eta^B_{N,N_*})\psi^B_{N,N_*} = \psi^B_{N,N_*}\frac{(v_*)^B_{N,N_*}}{v^B_{N,N_*}}. \tag{33}$$

Обозначим

$$c_2 = \frac{(v_*)^B_{N,N_*}}{v^B_{N,N_*}(1+\hat{\Phi}\Phi)}. \tag{34}$$

Так как $\mu^B_{N,N_*} = 1$, то

$$1+\hat{\Phi}\Phi = \frac{1}{x^B_{N,N_*}v^B_{N,N_*}}\|q^B_N\|^2, \quad q^B_N = D(\mu^B_N)\psi^B_N.$$

Поэтому

$$c_2 = (v_*)^B_{N,N_*}x^B_{N,N_*}/\|q^B_N\|^2 > 0.$$

Согласно (24) и (32)–(34) имеем

$$y^{B,B_*}_N = e - c_2\mu^{B,B_*}_N. \tag{35}$$

Из (25), (33) и (34) получаем также

$$y^B_{N,N_*} = 1 - c_2\mu^B_{N,N_*}. \tag{36}$$

Вычислим y^B_B. Согласно (12) и равенству $b = A^B_Bx^B_B + A^B_Nx^B_N$,

$$y^B_B = D^{-1}(x^B_B)(\hat{A}^B_BA^B_B)^{-1}\hat{A}^B_B[I - \Gamma Q]b = e + D^{-1}(x^B_B)(\hat{A}^B_BA^B_B)^{-1}\hat{A}^B_B[A^B_Nx^B_N - \Gamma Qb]. \tag{37}$$

Имеем также

$$\Gamma Qb = A^B_NG^B_NP\psi^B_N = A^B_NG^B_N\bar{\psi}^B_N. \tag{38}$$

Подставляя в (38) выражение для $\bar{\psi}^B_N = D(y^B_N)\psi^B_N$, получаем

$$A^B_NG^B_N\bar{\psi}^B_N = A^B_Nx^B_N - c_2A^B_N[x^B_N - (x_*)^B_N]. \tag{39}$$

Из (26) и (28) следует, что

$$A^B_N[x^B_N - (x_*)^B_N] = A^B_B[(x_*)^B_B - x^B_B].$$

Поэтому, в силу (37) и (39), $y^B_B = e - c_2\mu^B_B$. Отсюда и из (35), (36) получаем окончательно $y^B = e - c_2\mu^B$. Так как $\mu^{m+1} = 1$, $\mu^i < 1$, $i = 1, 2, \ldots, m$, то имеет место неравенство (23).

Теорема 3. *Пусть в допустимой паре точек* $[x, u]$ *выполнены предположения леммы* 2 *и*

$$y^B_* + y^*_N > 1. \tag{40}$$

Тогда $[x, u] \in \Upsilon_2$.

Доказательство. Из неравенства (40) с учетом формулы (14) получаем, что шаг α в точке $[x, u]$ равняется $1/y_N^*$. Но, в силу утверждения леммы 2, максимальной компонентой вектора y_N является та, индекс которой содержится в множестве $J_N^{B, B_*}(x, u)$. Так как это множество состоит только из одного индекса, то в новой точке $\bar{u} = u + \alpha\Delta u$ соответствующая компонента $v^i(\bar{u})$ станет равной нулю, что означает совпадение точки $\bar{u}$ с оптимальной вершиной u_*. На основании утверждения теоремы 1, из такой пары точек $[\bar{x}, \bar{u}]$, в которой $\bar{u} = u_*$, при любом $\bar{x} = x + \alpha\Delta x$ метод (2) позволяет находить решения задач (1) за одну итерацию.

Теорема 4. *Пусть в допустимой паре точек* $[x, u]$ *выполнены предположения леммы* 3 *и*

$$y_*^B + y_N^* < 1.$$

Тогда $[x, u] \in \Upsilon_2$.

Доказательство аналогично доказательству предыдущей теоремы.

Теорема 5. *Пусть в допустимой паре точек* $[x, u]$ *два неправильных индекса и имеет место равенство* (17). *Тогда* $[x, u] \in \Upsilon_2$.

Доказательство. Если оба неправильных индекса принадлежат множеству J_{B_*} или J_{N_*}, то, по лемме 1, $[x, u] \in \Upsilon_1$. Если неправильные индексы находятся в разных подмножествах, то выполнены условия как леммы 2, так и леммы 3. Поэтому $[x, u] \in \Upsilon_2$.

В [3] показано, что результат теоремы 5 сохраняется и без предположения о выполнении равенства (17). Наличие только двух неправильных индексов означает, что обе начальные точки x и u лежат на ребрах допустимых множеств, у которых один из концов является оптимальной вершиной.

4. КОНЕЧНАЯ ЛОКАЛЬНАЯ СХОДИМОСТЬ

Пусть $j_i \in J_{N_*}$ $(i = 1, 2, ..., r)$. Обозначим через $\gamma_{j_1, \dots, j_r}(X)$ множество

$$\gamma_{j_1, \dots, j_r}(X) = \{x \in X\colon x^j = 0, j \in J_1;\ x^j > 0, j \in J_2\},$$

где

$$J_1 = J_{N_*} \backslash \{j_1, \dots, j_r\}, \quad J_2 = J_{B_*} \cup \{j_1, \dots, j_r\}.$$

Аналогичным образом для $i_j \in J_{B_*}$ $(j = 1, 2, ..., s)$ определим множество

$$\gamma_{i_1, \dots, i_s}(U) = \{u \in U\colon v^i(u) = 0, i \in J_3;\ v^i(u) > 0, i \in J_4\}.$$

Здесь

$$J_3 = J_{B_*} \backslash \{i_1, \dots, i_s\}, \quad J_4 = J_{N_*} \cup \{i_1, \dots, i_s\}.$$

Замыкание $\bar{\gamma}_{j_1, \dots, j_r}(X)$ полиэдрального множества $\gamma_{j_1, \dots, j_r}(X)$ совпадает с r-мерной гранью X, содержащей оптимальную вершину x_*. Замыкание $\bar{\gamma}_{i_1, \dots, i_s}(U)$ множества $\gamma_{i_1, \dots, i_s}(U)$ также является s-мерной гранью множества U, прилегающей к вершине u_*. Множество $\bar{\gamma}_i(U)$, где $i \in J_{B_*}$, является ребром (или крайним лучом) множества U, имеющего в качестве одного из своих концов вершину u_*.

Непосредственно из определений множеств $\gamma_{j_1, \dots, j_r}(X)$ и $\gamma_{i_1, \dots, i_s}(U)$ следует, что для любых $x \in \gamma_{j_1, \dots, j_r}(X)$ и $u \in \gamma_{i_1, \dots, i_s}(U)$ множество $J_B^N(x, u)$ пусто и множества $J_B^B(x, u)$, $J_N^N(x, u)$ и $J_N^B(x, u)$ не зависят от конкретного выбора точек x и u. Кроме того,

$$\{i_1, \dots, i_s\} = J_N^{B, B_*}(x, u), \quad \{j_1, \dots, j_r\} = J_{N, N_*}^B(x, u), \tag{41}$$

$$J_{B_*} = J_B^B(x, u) \cup J_N^{B, B_*}(x, u), \quad J_{N_*} = J_N^N(x, u) \cup J_{N, N_*}^B(x, u). \tag{42}$$

Пусть w_i – вектор единичной длины, выходящий из оптимальной вершины u_* и имеющий направление i-го ребра $\bar{\gamma}_i(U)$ многогранника U. Составим из векторов w_i $(1 \le i \le m)$ матрицу

$$W = [w_1, \dots, w_m]$$

и положим $\Lambda = -\hat{A}\, W$. Элемент λ_{ij} матрицы Λ есть производная функции $v^i(u)$ по направлению w_j. В силу линейности функции $v^i(u)$, данная производная не зависит от выбора точки u. Поэтому у матрицы Λ верхняя квадратная подматрица порядка m является диагональной матрицей с положительными диагональными элементами $\lambda_{11} > 0, \dots, \lambda_{mm} > 0$.

Лемма 4. *Пусть* $u \in \gamma_{i_1, \dots, i_s}(U)$. *Тогда для любого фиксированного* $x \in \gamma_{j_1, \dots, j_r}(X)$ *имеют место представления*

$$y^i(x, u) = x^i_*/x^i + \rho^i_1(x, u), \quad i \in J_{B*}, \tag{43}$$

$$y^i(x, u) = \rho^i_1(x, u), \quad i \in J_{N*}, \tag{44}$$

где $|\rho^i_1(x, u)| = O(\|u - u_*\|)$ $(1 \le i \le n)$.

Доказательство. Считаем для определенности, что $\{i_1, \dots, i_s\} = \{1, 2, \dots, s\}$. Тогда, согласно (41) и (42),

$$J_N^{B, B*}(x, u) = \{1, 2, \dots, s\}, \quad J_B^B(x, u) = \{s+1, \dots, m\}.$$

Пусть K – конус, порожденный векторами $w_1, \dots, w_s$, $\operatorname{int} K$ – его внутренность. Так как $u \in \gamma_{i_1, \dots, i_s}(U)$, то $u - u_* \in \operatorname{int} K$ и, следовательно,

$$u - u_* = \sum_{i=1}^{s} \beta^i w_i, \quad \beta^i > 0, \quad i = 1, 2, \dots, s.$$

Линейно независимые векторы w_i $(1 \le i \le s)$ принадлежат подпространству $(S_B^B)^\perp$, поэтому они могут быть использованы в качестве базиса в $(S_B^B)^\perp$. Положим

$$H = -W_s D(\beta), \tag{45}$$

где W_s – подматрица матрицы W, состоящая из первых s столбцов W и $\beta^{\mathrm{T}} = [\beta^1, \dots, \beta^s]$. Тогда $\hat{A}\, H = \Lambda_s D(\beta)$, где Λ_s – подматрица матрицы Λ, состоящая из ее первых s столбцов.

Для любого $i = 1, 2, \dots, n$ выполняется равенство

$$v^i(u) = v^i_* + \sum_{j=1}^{s} \beta^j \lambda_{ij}.$$

Отсюда, в частности, получаем

$$v^i(u) = \beta^i \lambda_{ii} > 0, \quad i = 1, 2, \dots, s. \tag{46}$$

Рассмотрим матрицу $M_N(\beta) = D^{-1}(v_N)\hat{A}_N H$ с элементами $\mu_{ij}(\beta)$ $(i = 1, 2, \dots, s, m+1, \dots, n, j = 1, 2, \dots, s)$, и выясним ее поведение при $\|\beta\| \longrightarrow 0$. Имеет место формула

$$\mu_{ij}(\beta) = \frac{\beta^j \lambda_{ij}}{v^i_* + \Sigma_{l=1}^{s} \beta^l \lambda_{il}},$$

откуда следует, что

$$\mu_{ii}(\beta) \equiv 1, \quad i = 1, 2, \dots, s, \tag{47}$$

$$\mu_{ij}(\beta) \equiv 0, \quad i, j = 1, 2, \dots, s, \quad i \ne j, \tag{48}$$

и

$$\lim_{\|\beta\| \to 0} \mu_{ij}(\beta) = 0, \quad i = m+1, \dots, n, \quad j = 1, 2, \dots, s. \tag{49}$$

Поэтому предельная матрица $M_N(0)$ имеет вид

$$M_N(0) = \begin{bmatrix} D(e) \\ 0_{ds} \end{bmatrix}.$$

Здесь $d = n - m$.

Введем теперь в рассмотрение вектор $c(\beta) = \tilde{\Gamma}^{-1}(\beta)g$, где

$$\tilde{\Gamma}(\beta) = D^{-1}(\beta)\hat{H}\Gamma H, \quad g = D^{-1}(\beta)\hat{H}A_N^B x_N^B = -\hat{W}_s A_N^B x_N^B,$$

и найдем его предел при $\|\beta\| \longrightarrow 0$. Вектор $c(\beta)$ является решением системы линейных уравнений

$$\tilde{\Gamma}(\beta)c(\beta) = g. \tag{50}$$

Вычислим вектор g и матрицу $\tilde{\Gamma}(\beta)$. С учетом (46) получаем

$$g^i = \sum_{k \in J_N^B(x, u)} x^k \lambda_{ki} = x^i \lambda_{ii} + \sum_{k \in J_{N, N_*}^B(x, u)} x^k \lambda_{ki}, \quad i = 1, 2, \ldots, s,$$

$$\tilde{\Gamma}_{ij}(\beta) = \beta^j \sum_{k \in J_N^B(x, u)} \frac{x^k \lambda_{ki} \lambda_{kj}}{v_*^k + \Sigma_{l=1}^s \beta^l \lambda_{kl}}, \quad i, j = 1, 2, \ldots, s. \tag{51}$$

Из (51), в частности, следует

$$\tilde{\Gamma}_{ii}(\beta) = x^i \lambda_{ii} + \beta^i \sum_{k \in J_{N, N_*}^B(x, u)} \frac{x^k \lambda_{ki}^2}{v_*^k + \Sigma_{l=1}^s \beta^l \lambda_{kl}},$$

$$\tilde{\Gamma}_{ij}(\beta) = \beta^j \sum_{k \in J_{N, N_*}^B(x, u)} \frac{x^k \lambda_{ki} \lambda_{kj}}{v_*^k + \Sigma_{l=1}^s \beta^l \lambda_{kl}}, \quad i \neq j.$$

Переходя к пределу при $\|\beta\| \longrightarrow 0$, получаем, что

$$\tilde{\Gamma}(0) = \mathrm{diag}(x^1 \lambda_{11}, \ldots, x^s \lambda_{ss}),$$

причем все диагональные элементы матрицы $\tilde{\Gamma}(0)$ строго положительны.

Умножая равенство

$$\sum_{k=1}^m x_*^k a_k = \sum_{k=1}^m x^k a_k + \sum_{k \in J_{N, N_*}^B(x, u)} x^k a_k = b$$

слева на w_i^{T} $(1 \le i \le s)$, получаем, что

$$x_*^i \lambda_{ii} = x^i \lambda_{ii} + \sum_{k \in J_{N, N_*}^B(x, u)} x^k \lambda_{ki}, \quad i = 1, 2, \ldots, s.$$

Поэтому $g^{\mathrm{T}} = [x_*^1 \lambda_{11}, \ldots, x_*^s \lambda_{ss}]$ и, следовательно, вектор $c(0)$ удовлетворяет системе уравнений

$$x^i c^i(0) = x_*^i, \quad 1 \le i \le s,$$

решением которой является вектор

$$\hat{c}(0) = [x_*^1/x^1, \ldots, x_*^s/x^s]. \tag{52}$$

Согласно [5], решение системы (50) для достаточно малых по норме β представимо в виде

$$c(\beta) = c(0) + O(\|\beta\|). \tag{53}$$

Так как, в силу (12), $y_N(x, u) = M_N(\beta)c(\beta)$, то из (47)–(49) и (52), (53), с учетом того, что $\|u - u_*\| = O(\|\beta\|)$, приходим к разложениям (43) при $i \in J_N^{B, B_*}(x, u)$ и к разложениям (44) при $i \in B_{N, N_*}(x, u)$.

Покажем теперь, что разложение (43) справедливо и при $i \in J_B^B(x, u)$. Имеем

$$b = A_B^B D(x_B^B) y_B^B(x, u) + A_N^B D(x_N^B) y_N^B(x, u) = A_B^B D(x_B^B) y_B^B(x, u) + A_N^{B, B_*}(x_*)_N^{B, B_*} + q(x, u), \tag{54}$$

где $\|q(x, u)\| = O(\|u - u_*\|)$. Отсюда следует, что

$$A_B^B D(x_B^B) y_B^B(x, u_*) = b - A_N^{B, B_*}(x_*)_N^{B, B_*}. \tag{55}$$

Но, согласно (26), $A_B^B (x_*)_B^B = b - A_N^{B, B_*}(x_*)_N^{B, B_*}$. Вычитая данное равенство из (55), получаем

$$A_B^B[D(x_B^B) y_B^B(x, u_*) - (x_*)_B^B] = 0.$$

Так как столбцы матрицы A_B^B линейно независимы, то $y_B^B(x, u_*) = D^{-1}(x_B^B)(x_*)_B^B$, и, на основании (54), приходим к выводу, что разложение (43) остается справедливым при $i \in J_B^B(x, u)$.

Лемма 5. *Пусть $x \in \gamma_{j_1, \dots, j_r}(X)$. Тогда для любого фиксированного $u \in \gamma_{i_1, \dots, i_s}(U)$ имеют место представления*

$$y^i(x, u) = 1 + \rho_2^i(x, u), \quad i \in J_{B_*}, \tag{56}$$

$$y^i(x, u) = 1 - \frac{v_*^i}{v^i} + \rho_2^i(x, u), \quad i \in J_{N_*}, \tag{57}$$

где $|\rho_2^i(x, u)| = O(\|x - x_*\|)$ $(1 \le i \le n)$.

Доказательство. Согласно (12),

$$y_N(x, u) = D^{-1}(v_N)\hat{A}_N H c_1(x), \tag{58}$$

где вектор $c_1(x)$ удовлетворяет следующей системе линейных уравнений:

$$\tilde{\Gamma}(x)c_1(x) = \hat{H}b, \quad \tilde{\Gamma}(x) = \hat{Z}_N^B Z_N^B. \tag{59}$$

Здесь $Z_N^B = G_N^B \hat{A}_N^B H$.

Возьмем, как и при доказательстве предыдущей леммы, в качестве базиса в $(S_B^B)^\perp$ набор векторов (45). Тогда

$$\hat{A}_N^{B, B_*} H = D(v_N^{B, B_*}), \tag{60}$$

$$\hat{A}_{N, N_*} H = (\Lambda_s)_{N, N_*} D(\beta). \tag{61}$$

Из (26) и (60) следует, что

$$\hat{H}b = \hat{H} A_N^{B, B_*}(x_*)_N^{B, B_*} = D(v_N^{B, B_*})(x_*)_N^{B, B_*}. \tag{62}$$

Кроме того, имеем, в силу (59)–(61),

$$\begin{aligned} \tilde{\Gamma}(x) &= \hat{Z}_N^{B, B_*} Z_N^{B, B_*} + \hat{Z}_{N, N_*}^B Z_{N, N_*}^B = G_N^{B, B_*} D^2(v_N^{B, B_*}) G_N^{B, B_*} + G_{N, N_*}^B (\hat{\Lambda}_s)_{N, N_*}^B D^2(\beta)(\Lambda_s)_{N, N_*}^B G_{N, N_*}^B = \\ &= D(x_N^{B, B_*}) D(v_N^{B, B_*}) + D^{1/2}(x_{N, N_*}^B)\tilde{\Omega} D^{1/2}(x_{N, N_*}^B), \end{aligned} \tag{63}$$

где матрица

$$\tilde{\Omega} = D^{-1/2}(v_{N, N_*}^B)(\hat{\Lambda}_s)_{N, N_*}^B D^2(\beta)(\Lambda_s)_{N, N_*}^B D^{-1/2}(v_{N, N_*}^B)$$

не зависит от x.

Из формулы (63) следует, что

$$\tilde{\Gamma}(x_*) = D((x_*)_N^{B, B_*}) D(v_N^{B, B_*}).$$

Отсюда и из (62) получаем, что решением системы уравнений (59) при $x = x_*$ является $c_1(x_*) = e$. Поэтому решение уравнения (59) представимо в виде

$$c_1(x) = e + q(x),$$

где $\|q(x)\| = O(\|x - x_*\|)$. Подставляя его в (58) и учитывая, что $He = u_* - u$, получаем

$$y_N(x, u) = e - D^{-1}(v_N)(v_*)_N + q_1(x),$$

где $\|q_1(x)\| = O(\|x - x_*\|)$. Отсюда следуют разложения (56) и (57) при $i \in J_N(x, u)$.

Вычислим теперь вектор $y_B^B(x, u)$. На основании (12) получим

$$y_B^B(x, u) = e + D^{-1}(x_B^B)(\hat{A}_B^B A_B^B)^{-1} A_B^B A_N^B[I - D(y_N^B(x, u))]x_N^B.$$

Но поскольку $y^i(x, u) = 1 + \rho_2^i(x, u)$, $i \in J_N^{B, B_*}(x, u)$, то

$$[I - D(y_N^B(x, u))]x_N^B = [I - D(y_N^B(x, u))^{B_*}]x_N^{B, B_*} + [I - D(y_N^B(x, u))_{N_*}]x_{N, N_*}^B = q_2(x, u),$$

где $\|q_2(x, u)\| = O(\|x - x_*\|)$. Таким образом, разложение (56) имеет место и при $i \in J_B^B(x, u)$.

Через $\Delta_\delta(\bar{z}) = \{z \in \mathbb{R}^l : \|z - \bar{z}\| \le \delta\}$ обозначим δ-окрестность точки z.

Пусть X_1 – подмножество множества X, содержащее оптимальную вершину x_*, и пусть U_1 – подмножество множества U, содержащее оптимальную вершину u_*.

Определение 2. Метод (2) сходится к решениям задач (1) на множестве $X_1 \times U_1$ локально по x за k итераций, если найдется такое $\delta > 0$, что $X_1^\delta \times U_1 \subseteq \Upsilon_k$, где $X_1^\delta = X_1 \cap \Delta_\delta(x_*)$.

Определение 3. Метод (2) сходится к решениям задач (1) на множестве $X_1 \times U_1$ локально по u за k итераций, если найдется такое $\delta > 0$, что $X_1 \times U_1^\delta \subseteq \Upsilon_k$, где $U_1^\delta = U_1 \cap \Delta_\delta(u_*)$.

Определение 4. Метод (2) локально сходится к решениям задач (1) на множестве $X_1 \times U_1$ за k итераций, если найдутся такие $\delta_1 > 0$ и $\delta_2 > 0$, что $X_1^{\delta_1} \times U_1^{\delta_2} \subseteq \Upsilon_k$.

Теорема 6. *Метод* (2) *сходится к решениям задач* (1) *на* $\gamma_{j_1, \ldots, j_r}(X) \times \gamma_{i_1, \ldots, i_s}(U)$ *локально по u за r + s итераций.*

Доказательство проведем, используя индукцию по r и s. При $r = s = 1$ множества $\bar{\gamma}_{j_1}(X)$ и $\bar{\gamma}_{i_1}(U)$ являются ребрами допустимых множеств X и U, прилегающими, соответственно, к оптимальным вершинам x_* и u_*. Поэтому, по теореме 5, метод (2) позволяет находить решения задач (1) не более чем за две итерации. Отсюда, в частности, следует его локальная сходимость на $\gamma_{j_1}(X) \times \gamma_{i_2}(U)$ по u за две итерации.

Предположим теперь, что утверждение теоремы 6 справедливо при всех $1 \le p \le r$, $1 \le q \le s$, $p + q < r + s$, и докажем его для случая, когда $p = r$, $q = s$. Возьмем всевозможные подмножества множеств индексов $\{j_1, \ldots, j_r\}$ и $\{i_1, \ldots, i_s\}$, состоящие, соответственно, из p и q индексов, $p + q < r + s$. По предположению индукции, для каждой пары таких подмножеств индексов существует $\delta > 0$, при котором метод сходится к решениям задач (1) локально по u. Пусть $\bar{\delta}$ – минимальное из таких δ. Так как число подмножеств множеств $\{j_1, \ldots, j_r\}$ и $\{i_1, \ldots, i_s\}$ конечно, то $\bar{\delta} > 0$. Обозначим $U_1 = \gamma_{i_1, \ldots, i_s}(U)$ и

$$S(\xi) = \{u \in U_1 : \langle b, u \rangle \ge \xi\}, \quad \xi(\delta) = \min_{u \in U_1 \cap \Delta_\delta(u_*)} \langle b, u \rangle.$$

Рассмотрим отдельно три случая.

Случай 1:

$$\nu = \max_{i \in J_N^{B, B_*}(x, u)} (x_*^i / x^i) > 1. \tag{64}$$

Выберем $\delta_1 \le \bar{\delta}$ настолько малым, чтобы для всех $u \in U_1 \cap \Delta_{\delta_1}(u_*)$ выполнялось включение

$S(\xi(\delta_1)) \subseteq \Delta_{\bar\delta}(u_*)$. Возьмем, далее, $\delta_2 > 0$ такое, что

$$\max_{i \in J_{N, B_*}(x, u)} y^i(x, u) > \max_{i \in J_{N, N_*}(x, u)} y^i(x, u), \tag{65}$$

$$\min_{i \in J^B_B(x, u)} y^i(x, u) > 0, \tag{66}$$

$$\max_{i \in J^{B, B_*}_N(x, u)} y^i(x, u) + \min_{i \in J^B_{N, N_*}(x, u)} y^i(x, u) > 1 \tag{67}$$

для всех $u \in U_1 \cap \Delta_{\delta_2}(u_*)$. В силу утверждения леммы 4 и неравенства (64), такое $\delta_2 > 0$ всегда можно указать.

Положим $\delta = \min\{\delta_1, \delta_2\}$. Тогда если $u \in U_1 \cap \Delta_\delta(u_*)$, то, на основании формулы (14) и (65)–(67), шаг α равняется

$$\alpha = 1/\beta^*, \quad \beta^* = \max_{i \in J^{B, B_*}_N(x, u)} y^i(x, u)$$

и при переходе в новую точку $\bar u = u + \alpha\Delta u$ по крайней мере одна компонента $v^i(\bar u)$ вектора $v_N^{B, B_*}(\bar u)$ станет равной нулю. Компоненты $\bar x^i$, $i \in J^B(x)$, вектора $\bar x^B$ останутся положительными. Останутся положительными также компоненты $v^i(\bar u)$, $i \in J_{N, N_*}(u)$.

Случай 2:

$$\nu = \max_{i \in J^{B, B_*}_N(x, u)} (x^i_*/x^i) < 1.$$

Возьмем теперь $\delta_2 > 0$ таким, чтобы выполнялось (65), (66) и неравенство, обратное к (67):

$$\max_{i \in J^{B, B_*}_N(x, u)} y^i + \min_{i \in J^B_{N, N_*}(x, u)} y^i < 1$$

для всех $u \in U_1 \cap \Delta_{\delta_2}(u_*)$. Поскольку в данном случае $y^B_* \le 0$, то шаг α равняется

$$\alpha = (1 - \beta_*)^{-1}, \quad \beta_* = \min_{i \in J^B_{N, N_*}(x, u)} y^i(x, u).$$

В новой паре точек станет равной нулю по крайней мере одна из компонент $\bar x^i$ ($i \in J^B_{N, N_*}(x, u)$) вектора $\bar x^B$. Компоненты $\bar x^i$ ($i \in J^{B, B_*}(x)$) и компоненты $v^i(\bar u)$ ($i \in J_N(u)$) вектора $v_N(\bar u)$ останутся положительными.

Случай 3:

$$\nu = \max_{i \in J^{B, B_*}_N(x, u)} (x^i_*/x^i) = 1.$$

При этом предположении может обнуляться как компонента вектора $\bar x^B$, так и компонента вектора $v_N(\bar u)$. Но, если δ_2 взято достаточно малым, то неравенства (65) и (66) сохраняются. Поэтому у вектора $\bar x^B$ может стать равной нулю лишь компонента $\bar x^i$ ($i \in J^{B, N_*}(x)$), а у вектора $v_N(\bar u)$ – лишь компонента $v^i(\bar u)$ ($i \in J_{N, N_*}(u)$).

В любом случае новая пара точек $[\bar x, \bar u]$ оказывается такой, что $\bar x \in \gamma_{j_1, \ldots, j_p}(X)$, $\bar u \in \gamma_{i_1, \ldots, i_q}(U)$ и $p \le r$, $q \le s$, $p + q < r + s$. Кроме того, $\langle b, \bar u\rangle \ge \langle b, u\rangle$. Поэтому $\bar u \in \Delta_{\bar\delta}(u_*)$. По предположению индукции, метод (2) позволяет находить из пары точек $[\bar x, \bar u]$ решения обеих задач (1) за число итераций не большее, чем $p + q$. Следовательно, общее число итераций не превышает $r + s$.

Аналогично доказывается

Теорема 7. *Метод* (2) *сходится к решениям задач* (1) *на* $\gamma_{j_1, \ldots, j_r}(X) \times \gamma_{i_1, \ldots, i_s}(U)$ *локально по x за r + s итераций.*

Из утверждений предыдущих двух теорем приходим к выводу, что справедлива

Теорема 8. *Метод* (2) *локально сходится к решениям задач* (1) *на* $X \times U$ *за n итераций.*

Результат теоремы 8 показывает, что существуют такие окрестности оптимальных вершин x_* и u_*, что если текущие точки x_k, u_k оказываются в этих окрестностях и принадлежат граням соответствующих допустимых множеств, прилегающим к оптимальным вершинам, то все последующие точки также принадлежат этим граням, причем суммарная разность граней постепенно уменьшается. Общее число итераций, которое необходимо методу для решения задачи из произвольной допустимой пары точек, зависит от того, как скоро метод попадает в данную окрестность.

СПИСОК ЛИТЕРАТУРЫ

1. *Евтушенко Ю.Г., Жадан В.Г.* Барьерно-проективные и барьерно-ньютоновские численные методы оптимизации (случай линейного программирования) // Сообщ. по вычисл. матем. М.: ВЦ РАН, 1992.
2. *Евтушенко Ю.Г., Жадан В.Г., Черенков А.П.* Применение метода Ньютона к решению задач линейного программирования // Ж. вычисл. матем. и матем. физ. 1995. Т. 35. № 6. С. 850–866.
3. *Жадан В.Г.* Метод Ньютона с наискорейшим спуском для задач линейного программирования // Сообщ. по вычисл. матем. М.: ВЦ РАН, 1997.
4. *Жадан В.Г.* Прямо-двойственный метод Ньютона для задач линейного программирования // Ж. вычисл. матем. и матем. физ. 1999. Т. 39. № 1. С. 17–32.
5. *Воеводин В.В.* Вычислительные основы линейной алгебры. М.: Наука, 1977.