В. Г. Жадан, Прямо-двойственный метод Ньютона для задач линейного программирования, *Ж. вычисл. матем. и матем. физ.*, 1999, том 39, номер 1, 17–32

ЖУРНАЛ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ ФИЗИКИ, 1999, том 39, № 1, с. 17–32

УДК 519.852.6

ПРЯМО-ДВОЙСТВЕННЫЙ МЕТОД 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.

Для задачи линейного программирования рассматривается прямо-двойственный метод Ньютона, в котором текущие точки могут принадлежать границам допустимых множеств. В тех случаях, когда ньютоновская система для нахождения направлений перемещения является недоопределенной, для выбора ньютоновских направлений предлагается решать вспомогательную линейную задачу дополнительности. Исследуются основные свойства итеративного процесса.

1. ВВЕДЕНИЕ

В последнее время методам решения задач линейного программирования, особенно относящимся к классу методов внутренней точки, уделяется много внимания. Был предложен ряд алгоритмов, из которых наиболее эффективными оказались прямо-двойственные методы, аппроксимирующие так называемый "центральный путь" [1]–[13]. В [14], [15] рассматривался прямо-двойственный алгоритм, в котором для определения направлений перемещений методом Ньютона решалась система равенств, описывающих условия оптимальности для поставленной задачи. Шаги как в прямом, так и в двойственном пространствах выбирались на основе наискорейшего спуска, однако выход на границы допустимых множеств не разрешался. В [16] данный метод был обобщен на случай общей задачи линейного программирования, в которой присутствуют двусторонние ограничения.

В настоящей работе рассматривается предельный вариант метода [15], когда возможен выход на границы допустимых множеств. При этом для простоты предполагается, что шаги в прямом и двойственном пространствах совпадают и что все точки итеративного процесса являются допустимыми. В качестве функции, на основе минимизации которой определяется шаг, используется разность значений целевых функций в прямой и двойственной задачах. Так как выход на границы допустимых множеств может привести к недоопределенности системы уравнений для нахождения ньютоновских направлений, то для выбора конкретных направлений предлагается решать вспомогательную линейную задачу дополнительности.

Постановка задачи и основные обозначения, используемые в работе, даются в разд. 2. В разд. 3 описывается общая ньютоновская итерация. При этом рассматривается как регулярный случай, когда матрица ньютоновской системы является невырожденной, так и нерегулярный – когда она вырождена. В последнем случае формулируется специальная линейная задача дополнительности, на основе решения которой определяются единственные направления. Общие свойства ньютоновских направлений исследуются в разд. 4. Наконец, в разд. 5 описывается численная схема метода. Показывается, каким образом процесс ведет себя при попадании в вершины допустимых множеств. Рассматривается частный случай процесса, когда начальная пара точек совпадает с неоптимальной парой вершин.

2. ПОСТАНОВКА ЗАДАЧИ

Пусть требуется решить задачу линейного программирования

$$\min c^{\mathrm{T}}x, \quad Ax = b, \quad x \geq 0_n, \tag{1}$$

где c и x – векторы-столбцы из n-мерного евклидова пространства $\mathbb{R}^n$, A есть $(m \times n)$-матрица пол-

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

ного ранга, в которой $m < n$, 0_n – нулевой n-мерный вектор. Двойственной к (1) является задача

$$\max b^{\mathrm{T}} u, \quad c - A^{\mathrm{T}} u \geq 0_n. \tag{2}$$

Обозначим через X и U множества допустимых точек в задачах (1) и (2):

$$X = \{x \in \mathbb{R}^n_+ : Ax = b\}, \quad U = \{u \in \mathbb{R}^m : v(u) \geq 0_n\},$$

где $\mathbb{R}^n_+$ – неотрицательный ортант $\mathbb{R}^n$ и $v(u) = c - A^{\mathrm{T}}u$. Пара точек $[x, u] \in X \times U$ называется *допустимой*. Если, помимо того, $x > 0_n$, $v(u) > 0_n$, то такая допустимая пара называется *внутренней*.

Всюду ниже предполагается, что множества X и U не пусты. Считается также и нигде специальным образом не оговаривается, что обе задачи (1) и (2) не вырождены, т.е. не вырождены все вершины допустимых множеств X и U. Из сделанных предположений вытекает, что задачи (1) и (2) имеют единственные оптимальные решения x_* и u_*, являющиеся вершинами, соответственно, множеств X и U. Пара $[x_*, u_*]$ называется *оптимальной*.

Введем дополнительный набор обозначений, которым будем пользоваться в дальнейшем. Любым допустимым точкам $x \in X$, $u \in U$ поставим в соответствие индексные множества

$$\begin{aligned} J^B(x) &= \{i: x^i > 0\}, \quad J^N(x) = \{i: x^i = 0\}, \\ J_B(u) &= \{i: v^i(u) = 0\}, \quad J_N(u) = \{i: v^i(u) > 0\}. \end{aligned} \tag{3}$$

Кроме того, рассмотрим их всевозможные пересечения:

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

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

Пусть a_i есть i-й столбец матрицы A. В соответствии с разбиением множества всех индексов $[1 : n]$ на подмножества (3) разобьем также матрицу A на подматрицы A^B, A^N, A_B и A_N, включив в них те столбцы a_i, индексы которых принадлежат, соответственно, множествам $J^B(x)$, $J^N(x)$ и $J_B(u)$, $J_N(u)$. Символами A^B_B, A^N_N, A^B_N и A^N_B будем обозначать подматрицы матрицы A, составленные из столбцов с индексами, соответственно, из множеств $J^B_B(x, u)$, $J^N_N(x, u)$, $J^B_N(x, u)$ и $J^N_B(x, u)$. Аналогичные разбиения будем использовать и для n-мерных векторов. Например, x^B_B – часть вектора x, содержащая компоненты x^i с индексами i из множества $J^B_B(x, u)$; v^B_N – часть вектора $v = v(u)$, составленная из компонент с индексами из $J^B_N(x, u)$.

Обозначим через $|J|$ число элементов в индексном множестве J. Положим также $d = n - m$. Из предположения о невырожденности обеих задач (1) и (2) вытекают неравенства

$$|J^B(x)| \geq m, \quad |J^N(x)| \leq d, \quad |J_B(u)| \leq m, \quad |J_N(u)| \geq d, \tag{6}$$

имеющие место в произвольной допустимой паре точек $[x, u]$. Поэтому

$$|J^B_B(x, u)| \leq m, \quad |J^N_N(x, u)| \leq d, \quad |J^N_B(x, u)| \leq \min\{m, d\}. \tag{7}$$

Кроме того, если $[x, u]$ – внутренняя пара точек, то

$$|J^B_N(x, u)| = n, \quad |J^B_B(x, u)| = |J^N_N(x, u)| = |J^N_B(x, u)| = 0. \tag{8}$$

При выполнении строгого равенства $|J^B_B(x, u)| = m$ точка u является вершиной множества U. Аналогично, при выполнении строгого равенства $|J^N_N(x, u)| = d$ точка x является вершиной множества X.

Лемма 1. *Пусть $[x, u]$ – допустимая пара точек. Тогда имеют место следующие утверждения*:

1) $[x, u]$ *является оптимальной парой в том и только том случае, когда*

$$|J^B_B(x, u)| = m, \quad |J^N_N(x, u)| = d, \quad |J^B_N(x, u)| = |J^N_B(x, u)| = 0;$$

2) $[x, u]$ *отлична от оптимальной пары в том и только том случае, когда* $|J^B_N(x, u)| > 0$;

3) $[x, u]$ *является неоптимальной парой, состоящей из вершин, соответственно, множеств* X *и* U, *в том и только том случае, когда* $|J_B^N(x, u)| > 0$.

Доказательство. Докажем только последнее утверждение. Предположим, что вершина x множества X и вершина u множества U образуют неоптимальную пару точек $[x, u]$. В этом случае $|J^B(x)| = m$ и $|J_B(u)| = m$. Если допустить, что множество $J_B^N(x, u)$ пусто, то $J_B(u) = J_B^B(x, u)$. Но поскольку x – невырожденная вершина множества X, то $|J^B(x)| = m$. Поэтому $J^B(x) = J_B^B(x, u) = J_B(u)$. Так как при этом обязательно $J_N^B(x, u) = \varnothing$, то на основании утверждения 1) леммы приходим к выводу, что $[x, u]$ – оптимальная пара. Данное утверждение противоречит сделанному допущению. Лемма доказана.

Предположение о невырожденности задач (1) и (2) приводит также к тому, что в любой допустимой паре точек $[x, u]$ матрицы A^B и A_B имеют полные ранги, причем ранг матрицы A^B равен m. Отсюда следует, что подматрицы A_B^B и A_B^N также имеют полные ранги.

3. НЬЮТОНОВСКАЯ ИТЕРАЦИЯ

Для того чтобы пара точек $[x, u] \in \mathbb{R}_+^n \times U$ была решением задач (1) и (2), необходимо и достаточно, чтобы

$$D(x)v(u) = 0_n, \tag{9}$$

$$Ax = b, \tag{10}$$

где через $D(x)$ обозначена диагональная матрица с вектором x на диагонали.

Если для решения системы уравнений (9), (10) применить метод Ньютона с переменным шагом, то получим следующие формулы перехода из текущей точки $[x, u]$ в новую точку $[\bar{x}, \bar{u}]$:

$$\bar{x} = x + \alpha\Delta x, \quad \bar{u} = u + \alpha\Delta u, \tag{11}$$

где $\alpha > 0$. Векторы приращений Δx и Δu удовлетворяют следующей системе линейных уравнений:

$$W(x, u)\begin{bmatrix}\Delta x\\ \Delta u\end{bmatrix} = -\begin{bmatrix}D(x)v(u)\\ b - Ax\end{bmatrix}. \tag{12}$$

Матрица $W(x, u)$ в (12) имеет вид

$$W(x, u) = \begin{bmatrix}D(v(u)) & -D(x)\hat{A}\\ -A & 0_{mm}\end{bmatrix}. \tag{13}$$

Здесь и ниже 0_{ks} – нулевая матрица размера $k \times s$, крышечка над матрицей используется для обозначения ее транспонирования.

Обозначим через $W_1(x, u)$ квадратную подматрицу матрицы $W(x, u)$, получающуюся вычеркиванием из нее строк и столбцов, индексы которых принадлежат множеству $J_B^N(x, u)$. Если множество $J_B^N(x, u)$ пусто, то считаем, что $W_1(x, u)$ совпадает с $W(x, u)$.

Лемма 2. *Для любых допустимых пар* $[x, u]$ *матрица* $W_1(x, u)$ *является неособой.*

Доказательство. Предположим для определенности, что $J_N^N(x, u) = \{1, 2, ..., r\}$, $J_N^B = \{r + 1, ...$ $..., r + s\}$ и $J_B^B(x, u) = \{r + s + 1, ..., r + s + t\}$, где $r + s + t \le n$. Тогда матрица $W_1(x, u)$ имеет вид

$$W_1(x, u) = \begin{bmatrix}D(v_N^N) & 0_{r(s+t+m)}\\ w_1 & W_2(x, u)\end{bmatrix},$$

где

$$W_2(x,u) = \begin{bmatrix} D(v_N^B) & 0_{st} & -D(x_N^B)\hat{A}_N^B \\ 0_{ts} & 0_{tt} & -D(x_B^B)\hat{A}_B^B \\ -A_N^B & (-A_B^B) & 0_{mm} \end{bmatrix}, \quad w_1 = \begin{bmatrix} 0_{sr} \\ 0_{tr} \\ -A_N^N \end{bmatrix}.$$

Отсюда следует, что

$$\det W_1(x,u) = \det W_2(x,u) \prod_{i \in J_N^N(x,u)} v^i(u). \tag{14}$$

После умножения первой строки матрицы $W_2(x, u)$ слева на $A_N^B D^{-1}(v_N^B)$ и сложения ее с последней строкой приходим к матрице

$$\overline{\overline{W}}_2(x,u) = \begin{bmatrix} D(v_N^B) & w_2 \\ 0_{(t+m)s} & W_3(x,u) \end{bmatrix}, \tag{15}$$

где

$$W_3(x,u) = \begin{bmatrix} 0_{tt} & -D(x_B^B)\hat{A}_B^B \\ -A_B^B & -\Gamma \end{bmatrix}, \quad \hat{w}_2 = \begin{bmatrix} 0_{ts} \\ -A_N^B D(x_N^B) \end{bmatrix}, \quad \Gamma = A_N^B D(x_N^B) D^{-1}(v_N^B)\hat{A}_N^B.$$

Так как $\det \overline{\overline{W}}_2(x, u) = \det W_2(x, u)$, то из (14) и (15) получаем, что

$$\det W_1(x,u) = \det W_3(x,u) \prod_{i \in J_N(u)} v^i(u).$$

Поэтому $W_1(x, u)$ будет неособой матрицей в том и только том случае, когда неособой является матрица $W_3(x, u)$.

Рассмотрим сначала случай, когда $J_B^B(x, u) = \varnothing$. Тогда $W_3(x, u) = \Gamma$. Так как при этом предположении обязательно $J^B(x) = J_N^B(x, u)$, то матрица A_N^B имеет полный ранг, равный m, поэтому матрица Γ неособая.

Перейдем теперь к случаю, когда $J_B^B(x, u) \neq \varnothing$. Для доказательства невырожденности матрицы $W_3(x, u)$ убедимся, что линейная система $W_3(x, u)\bar{z} = 0$ имеет только нулевое решение. Используя для вектора $\bar{z}$ разбиение $\bar{z} = [\bar{x}_B^B, \bar{u}]$, записываем эту систему в более подробном виде:

$$D(x_B^B)\hat{A}_B^B\bar{u} = 0, \quad A_B^B\bar{x}_B^B + \Gamma\bar{u} = 0. \tag{16}$$

Пусть s – ранг матрицы A_B^B. Если $s = m$, то, согласно первому равенству (16), $\bar{u} = 0$. Второе равенство (16) тогда переходит в $A_B^B\bar{x}_B^B = 0$, следовательно, $\bar{x}_B^B = 0$.

Предположим теперь, что $s < m$. Так как A_B^B – матрица полного ранга, то число столбцов в ней равно s. Кроме того, на основании (16) имеем $\bar{u} = Hq$, где H – $m \times (m - s)$ – матрица, составленная из векторов произвольного базиса нуль-пространства матрицы $\hat{A}_B^B$ (ортогонального дополнения пространства столбцов матрицы A_B^B), q – некоторый вектор из $\mathbb{R}^{m-s}$. Поэтому второе равенство (16) можно переписать в виде $\Gamma Hq + A_B^B\bar{x}_B^B = 0$ или (после его умножения слева на матрицу $\hat{H}$)

$$\hat{H}\Gamma Hq = 0. \tag{17}$$

Число столбцов матрицы $\hat{A}_N^B H$, равное $m - s$, не больше числа строк. Действительно, в противном случае имели бы $m - |J_B^B(x, u)| > |J_N^B(x, u)|$. Поэтому

$$\left|J_B^B(x, u)\right| + \left|J_N^B(x, u)\right| = \left|J^B(x)\right| < m,$$

что противоречит (6). Таким образом, ранг матрицы $\hat{A}_N^B H$ не превышает $m - s$. Убедимся, что он равен $m - s$. Пусть это не так. Тогда столбцы матрицы $\hat{A}_N^B H$ линейно зависимы, т.е. $\hat{A}_N^B Hh = 0$ для некоторого ненулевого вектора $h \in \mathbb{R}^{m-s}$. Поэтому ненулевой вектор Hh, принадлежащий нуль-пространству матрицы $\hat{A}_B^B$, принадлежит одновременно нуль-пространству матрицы $\hat{A}_N^B$, т.е. он принадлежит нуль-пространству матрицы $\hat{A}^B$, что невозможно из-за того, что ранг $\hat{A}^B$ равен m.

Так как столбцы матрицы $\hat{A}_N^B H$ линейно независимы, то матрица Грама $\hat{H}\Gamma H$ не вырождена и из (17) получаем, что $q = 0$. Тогда на основании (16) имеем $\bar{x}_B^B = 0$. Таким образом, $\bar{z} = 0$ и, следовательно, $W_3(x, u)$ – неособая матрица. Лемма доказана.

Определение 1. Допустимая пара точек $[x, u]$ называется *регулярной*, если $J_B^N(x, u) = \varnothing$. В противном случае пара $[x, u]$ называется *нерегулярной*.

Используя утверждение леммы 2, приходим к следующему результату.

Теорема 1. *Матрица $W(x, u)$ является неособой в том и только том случае, когда пара $[x, u]$ регулярна.*

Из (8) следует, что все внутренние пары точек являются регулярными. Кроме того, согласно утверждению 1) леммы 1, оптимальная пара также регулярна. Если пара точек $[x, u]$ нерегулярна, то система (12) является недоопределенной, т.е. число неизвестных больше числа уравнений.

Рассмотрим произвольную допустимую пару точек $[x, u]$ и обозначим

$$n_B^B = \left|J_B^B(x, u)\right|, \quad n_N^N = \left|J_N^N(x, u)\right|,$$
$$n_N^B = \left|J_N^B(x, u)\right|, \quad n_B^N = \left|J_B^N(x, u)\right|.$$

Система (12) в этой паре точек запишется в виде

$$D(v_N^B)\Delta x_N^B - D(x_N^B)\hat{A}_N^B\Delta u = -D(x_N^B)v_N^B, \tag{18}$$

$$D(v_N^N)\Delta x_N^N = 0_{n_N^N}, \tag{19}$$

$$D(x_B^B)\hat{A}_B^B\Delta u = 0_{n_B^B}, \tag{20}$$

$$A_B^B\Delta x_B^B + A_N^B\Delta x_N^B + A_N^N\Delta x_N^N + A_B^N\Delta x_B^N = 0_m. \tag{21}$$

Эта система состоит из $n + m - n_B^N$ уравнений. Рассмотрим возможные случаи ее решения в зависимости от числа индексов в множестве $J_B^B(x, u)$.

Случай 1: $n_B^B = m$. Тогда обязательно $[x, u]$ является регулярной парой точек и, в силу полноты ранга матрицы A_B^B, из (18)–(21) получаем

$$\Delta u = 0_m, \quad \Delta x_N^N = 0_{n_N^N}, \quad \Delta x_N^B = -x_N^B, \quad \Delta x_B^B = (A_B^B)^{-1}A_N^B x_N^B. \tag{22}$$

Заметим, что точка u является вершиной допустимого множества U.

Случай 2: $0 < n_B^B < m$. Предполагая заданным Δx_B^N, разрешаем систему (18)–(21) относительно Δx_B^B, Δx_N^B, Δx_N^N и Δu. Пусть $(S_B^B)^\perp$ – ортогональное дополнение подпространства S_B^B, порожденного столбцами матрицы A_B^B, и пусть H – матрица, составленная из векторов произвольного базиса в $(S_B^B)^\perp$. Так как ранг матрицы A_B^B равен n_B^B, то из (19), (20) получаем, что $\Delta x_N^N = 0$ и $\Delta u = Hq$, где $q \in \mathbb{R}^s$, $s = m - n_B^B$. Из (18), (21) с учетом того, что $v_N^B > 0$, приходим к более простой системе из $m + n_N^B$ уравнений:

$$\Delta x_N^B - (G_N^B)^2 \hat{A}_N^B H q = -x_N^B, \quad A_B^B \Delta x_B^B + A_N^B \Delta x_N^B + A_B^N \Delta x_B^N = 0_m,$$

где

$$G_N^B = D^{1/2}(x_N^B) D^{-1/2}(v_N^B).$$

Разрешая ее относительно $m + n_N^B$ переменных Δx_B^B, Δx_N^B и q, получаем

$$q = (\hat{H}\Gamma H)^{-1} \hat{H}(A_N^B x_N^B - A_B^N \Delta x_B^N), \tag{23}$$

$$\Delta x_N^B = (G_N^B)^2 \hat{A}_N^B Q (A_N^B x_N^B - A_B^N \Delta x_B^N) - x_N^B, \tag{24}$$

$$\Delta x_B^B = (\hat{A}_B^B A_B^B)^{-1} \hat{A}_B^B (I - \Gamma Q)(A_N^B x_N^B - A_B^N \Delta x_B^N). \tag{25}$$

Здесь и далее I – единичная матрица,

$$\Gamma = A_N^B (G_N^B)^2 \hat{A}_N^B, \quad Q = H(\hat{H}\Gamma H)^{-1} \hat{H}, \tag{26}$$

причем из доказательства леммы 2 следует, что $\hat{H}\Gamma H$ – неособая матрица.

Обозначим $b_1 = b - A_B^N \Delta x_B^N$. В произвольной точке $x \in X$ имеет место равенство

$$A_B^B x_B^B + A_N^B x_N^B = b. \tag{27}$$

Тогда, поскольку $\hat{H} A_B^B = 0$, из (23)–(27) получаем

$$\Delta u = Q b_1, \tag{28}$$

$$\Delta x_N^B = (G_N^B)^2 \hat{A}_N^B Q b_1 - x_N^B, \quad \Delta x_N^N = 0_{n_N^N}, \tag{29}$$

$$\Delta x_B^B = (\hat{A}_B^B A_B^B)^{-1} \hat{A}_B^B (I - \Gamma Q) b_1 - x_B^B. \tag{30}$$

Случай 3: $n_B^B = 0$. Тогда $J^B(x) = J_N^B(x, u)$, поэтому матрица $\hat{A}_N^B$ состоит из m линейно независимых столбцов. Решение системы (18)–(21) имеет вид

$$\Delta u = Q_1 (A_N^B x_N^B - A_B^N \Delta x_B^N), \tag{31}$$

$$\Delta x_N^B = (G_N^B)^2 \hat{A}_N^B Q_1 (A_N^B x_N^B - A_B^N \Delta x_B^N) - x_N^B, \quad \Delta x_N^N = 0_{n_N^N}. \tag{32}$$

Здесь $Q_1 = (A_N^B (G_N^B)^2 \hat{A}_N^B)^{-1}$. Матрица Q_1 совпадает с Q в том случае, когда H – матрица базиса во всем пространстве $\mathbb{R}^m$ и, следовательно, является невырожденной квадратной матрицей порядка m. Из (31) и (32) с учетом равенства $A_N^B x_N^B = b$ получаем

$$\Delta u = Q_1 b_1, \quad \Delta x_N^B = (G_N^B)^2 \hat{A}_N^B Q_1 b_1 - x_N^B, \quad \Delta x_N^N = 0_{n_N^N}.$$

В дальнейшем решение системы (18)–(21) для всех трех случаев будем записывать в виде (28)–(30), считая, что матрица Q является нулевой, если $n_B^B = m$, и совпадает с Q_1, если $n_B^B = 0$. В последнем случае отпадает надобность в пересчете вектора x_B^B.

Изменение компонент вектора $v(u)$, соответствующее изменению (28) вектора u, равно $\Delta v = -\hat{A}\,\Delta u$, или (в более подробной записи)

$$\Delta v_B^B = -\hat{A}_B^B Q b_1 = 0, \quad \Delta v_N^B = -\hat{A}_N^B Q b_1, \tag{33}$$

$$\Delta v_B^N = -\hat{A}_B^N Q b_1, \quad \Delta v_N^N = -\hat{A}_N^N Q b_1. \tag{34}$$

Так как в текущей точке $x_N^N = 0$, $v_B^B = 0$, то первое равенство совместно с равенством $\Delta x_N^N = 0$ означает, что компоненты векторов x_N^N и v_B^B остаются равными нулю.

Согласно утверждению теоремы 1, если допустимая пара точек $[x, u]$ является регулярной, то система (18)–(21) имеет единственное решение. Если нет – то она имеет множество решений и выбор конкретного решения из этого множества зависит от выбора вектора Δx_B^N. Распорядимся этим вектором таким образом, чтобы приращение Δv_B^N вектора v_B^N было неотрицательным. Согласно (34) имеем $\Delta v_B^N = \Omega \Delta x_B^N - p$, где $\Omega = \hat{A}_B^N Q A_B^N$, $p = \hat{A}_B^N Qb$. Отсюда приходим к неравенству

$$\Delta v_B^N = \Omega \Delta x_B^N - p \geq 0_{n_B^N}. \tag{35}$$

Так как матрицы Γ или $\hat{H}\Gamma H$, входящие в матрицу Ω, являются матрицами Грама, составленными для линейно независимых столбцов, то симметричная матрица Ω неотрицательно определена. На самом деле имеет место более сильное утверждение.

Лемма 3. *В любой допустимой нерегулярной паре точек* $[x, u]$ *матрица* Ω *положительно определена.*

Доказательство. Прежде всего отметим, что если множество $J_B^N(x, u)$ не пусто, то обязательно $n_B^B < m$.

Пусть $0 < n_B^B < m$, и пусть $t = n_B^N$. Возьмем произвольный ненулевой вектор $z \in \mathbb{R}^t$ и обозначим $w = \hat{H} A_B^N z$. Вектор $w \neq 0$, ибо иначе ненулевой вектор $A_B^N z$ был бы ортогонален столбцам матрицы H. Так как эти столбцы образуют базис в нуль-пространстве матрицы $\hat{A}_B^B$, то это означает, что вектор $A_B^N z$ принадлежит пространству столбцов матрицы A_B^B, являющемуся ортогональным дополнением нуль-пространства $\hat{A}_B^B$. Поэтому столбцы матрицы A_B линейно зависимы, что невозможно, поскольку, в силу невырожденности двойственной задачи, матрица A_B имеет полный ранг. Отсюда с учетом положительной определенности матрицы $\hat{H}\Gamma H$ получаем, что $\langle z, \Omega z\rangle = \langle w, (\hat{H}\Gamma H)^{-1} w\rangle > 0$. Таким образом, матрица Ω также положительно определена.

В случае, когда $n_B^B = 0$, матрица Ω имеет вид $\Omega = \hat{A}_B^N Q_1 A_B^N$ и неравенство $\langle z, \Omega z\rangle = \langle A_B^N z, Q_1 A_B^N z\rangle >$ > 0 вытекает из полноты ранга матрицы A_B^N. Лемма доказана.

В случае, когда множество $J_B^N(x, u)$ состоит из одного индекса, матрица Ω является положительным числом. Если $p \leq 0$, то из (35) видно, что при $\Delta x_B^N = 0$ выполняется $\Delta v_B^N \geq 0$. Если $p > 0$, то, полагая $\Delta x_B^N = p/\Omega$, будем иметь $\Delta v_B^N = 0$. Таким образом, всегда существует неотрицательное решение неравенства (35), удовлетворяющее условию $\Delta x_B^N \Delta v_B^N = 0$. Покажем, что такое свойство имеет место и при большем количестве индексов в множестве $J_B^N(x, u)$. Ниже через $\langle \cdot, \cdot \rangle$ обозначено евклидово скалярное произведение.

Теорема 2. *Пусть выполнены предположения леммы* 3. *Тогда существует единственное неотрицательное решение* Δx_B^N *системы линейных неравенств* (35), *удовлетворяющее условию*

$$\langle \Delta x_B^N, \Delta v_B^N \rangle = 0. \tag{36}$$

Доказательство. Составим функцию

$$f(\Delta x_B^N) = \frac{1}{2}\langle \Delta x_B^N, \Omega \Delta x_B^N\rangle - \langle p, \Delta x_B^N\rangle.$$

Так как Ω – положительно-определенная матрица, то эта квадратичная функция сильно выпукла. Поэтому существует единственное решение задачи

$$\min_{\Delta x_B^N \geq 0} f(\Delta x_B^N). \tag{37}$$

В силу необходимых условий минимума, в этом решении выполняются соотношения

$$\nabla f(\Delta x_B^N) = \Omega \Delta x_B^N - p \geq 0, \quad \langle \nabla f(\Delta x_B^N), \Delta x_B^N \rangle = 0.$$

Отсюда приходим к утверждению теоремы.

Объединяя (35) и (36), получаем линейную задачу дополнительности

$$\Omega \Delta x_B^N - p \geq 0, \quad \Delta x_B^N \geq 0, \quad \langle \Omega \Delta x_B^N - p, \Delta x_B^N \rangle = 0. \tag{38}$$

Из неравенства (6) следует, что

$$\left|J_B^B(x, u)\right| + \left|J_B^N(x, u)\right| \leq m, \quad \left|J_B^B(x, u)\right| + \left|J_N^B(x, u)\right| \geq m. \tag{39}$$

Поэтому

$$\left|J_B^N(x, u)\right| \leq \left|J_N^B(x, u)\right|. \tag{40}$$

Таким образом, размерность линейной задачи дополнительности (38) не превосходит числа индексов в множестве $J_N^B(x, u)$ и, согласно неравенству (7), не больше максимального из двух чисел m и d. Из (39) следует также, что она не превышает размерность нуль-пространства матрицы A_B^B. Отметим также, что, в силу единственности решения (37), задача (38) также имеет единственное решение, поэтому независимо от того, является пара $[x, u]$ регулярной или нет, направления Δx и Δu определяются единственным образом.

4. СВОЙСТВА НЬЮТОНОВСКИХ НАПРАВЛЕНИЙ

Приведем простейшие свойства ньютоновских направлений Δx и Δu, которые вытекают непосредственно из их определений (28)–(30).

Лемма 4. *Пусть в регулярной паре* $[x, u]$ *точка* u *является вершиной множества* U. *Тогда* $\Delta u = 0$.

Доказательство. Так как u – вершина U, то имеет место равенство $|J_B(u)| = m$. Но $J_B(u) = J_B^B(x, u) \cup J_B^N(x, u)$, и поскольку $J_B^N(x, u) = \varnothing$, то $J_B(u) = J_B^B(x, u)$. Следовательно, $|J_B^B(x, u)| = m$. Поэтому на основании (22) получаем, что $\Delta u = 0$.

Определение 2. Задача (1) сильно невырожденна, если вектор b не принадлежит никакому подпространству пространства $\mathbb{R}^m$, порожденному менее чем m столбцами матрицы A.

Лемма 5. *Пусть задача* (1) *сильно невырожденна. Тогда в любой допустимой паре* $[x, u]$ *если* $\Delta u = 0$, *то* u – *вершина множества* U.

Доказательство. Согласно (28) имеем $\Delta u = H(\hat{H}\Gamma H)^{-1}\hat{H} b_1 = 0$. Так как столбцы матрицы H линейно независимы, то $(\hat{H}\Gamma H)^{-1}\hat{H} b_1 = 0$ и, следовательно, $\hat{H} b_1 = 0$, вектор b_1 принадлежит нуль-пространству матрицы $\hat{H}$, совпадающему с пространством столбцов матрицы A_B^B. Но тогда для некоторого $z \in \mathbb{R}^{n_B^B}$ выполнено соотношение

$$A_B^B z + A_B^N \Delta x_B^N = b. \tag{41}$$

Матрица A_B состоит из не более чем m столбцов. Если их меньше m, то равенство (41) противоречит предположению о сильной невырожденности задачи (1). Если же их число равно m, то это означает, что m линейно независимых столбцов A_B являются базисом вершины u допустимого множества U.

Лемма 6. *Пусть в регулярной паре* $[x, u]$ *точка* x *является вершиной множества* X. *Тогда* $\Delta x = 0$.

Доказательство. Так как x – вершина, то столбцы матрицы $A^B = [A_B^B, A_N^B]$ линейно независимы. Тогда на основании равенства

$$A_B^B \Delta x_B^B + A_N^B \Delta x_N^B = 0_m, \tag{42}$$

вытекающего из (21) и (29), получаем, что $\Delta x_B^B = 0$, $\Delta x_N^B = 0$. Таким образом, $\Delta x = 0$.

Определение 3. Задача (2) сильно невырожденна, если не существует $u \in \mathbb{R}^m$, для которого более чем m компонент вектора $v(u)$ равны нулю.

Лемма 7. *Пусть задача* (2) *сильно невырожденна. Тогда в любой допустимой паре* $[x, u]$ *если* $\Delta x = 0$, *то* x – *вершина множества* X.

Доказательство. Так как, по лемме 2, матрица $W_1(x, u)$ является неособой, то система (18)–(21) при $\Delta x_B^N = 0$ имеет единственное решение относительно остальных переменных. Но в этом случае она переходит в

$$\hat{A}_B^B \Delta u = 0, \quad \hat{A}_N^B \Delta u = v_N^B. \tag{43}$$

Покажем, что число столбцов в матрице $A^B = [A_B^B, A_N^B]$ равно m. Действительно, если бы их было меньше m, то система (43) имела бы неединственное решение и, следовательно, неединственное решение имела бы система (18)–(21). Если число уравнений в (43) больше m, то существование ее решения равносильно существованию такого вектора $\tilde{u} \in \mathbb{R}^m$, что $\hat{A}_B^B(\tilde{u} - u) = 0$ и $\hat{A}_N^B(\tilde{u} - u) = v_N^B$. Но $\hat{A}_B^B(\tilde{u} - u) = -v_B^B(\tilde{u})$, $\hat{A}_N^B(\tilde{u} - u) = v_N^B - v_N^B(\tilde{u})$. Таким образом, в точке $\tilde{u} \in \mathbb{R}^m$ более чем m компонент вектора $v(\tilde{u})$ равны нулю, что противоречит предположению о сильной невырожденности двойственной задачи (2). Лемма доказана.

Согласно (18) имеем $D^{-1}(x_N^B)\Delta x_N^B + D^{-1}(v_N^B)\Delta v_N^B = -e$, где e – вектор со всеми компонентами, равными единице. Из формул (29), (33) следует также, что

$$\langle \Delta x_N^B, \Delta v_N^B \rangle = -\langle Q\Gamma Q b_1 - QA_N^B x_N^B, b_1 \rangle. \tag{44}$$

Из (34) и (44), учитывая равенства $Q\Gamma Q = Q$, $QA_N^B x_N^B = Qb$ и (36), получаем

$$\langle \Delta x_N^B, \Delta v_N^B \rangle = \langle QA_B^N \Delta x_B^N, b_1 \rangle = -\langle \Delta x_B^N, \Delta v_B^N \rangle = 0. \tag{45}$$

Таким образом, $\langle \Delta x, \Delta v \rangle = 0$, т.е. сохраняется ортогональность направлений Δx и Δv, имеющая место для внутренних пар.

Введем в рассмотрение вектор-функцию $\psi(x, u)$ и функцию $V(x, u)$, положив

$$\psi(x, u) = D^{1/2}(x)D^{1/2}(v(u))e, \quad V(x, u) = \|\psi(x, u)\|^2, \tag{46}$$

где $\|\cdot\|$ – евклидова норма в $\mathbb{R}^n$. Как известно (см., например, [7]), существует взаимно однозначное соответствие между компонентами вектора ψ и точками множества $X \times U$. На множестве $X \times U$ функция $V(x, u)$ неотрицательна и

$$V(x, u) = \langle c, x \rangle - \langle b, u \rangle. \tag{47}$$

Условие (9) в точке $[x_*, u_*] \in X \times U$ эквивалентно тому, что $V(x_*, u_*) = 0$.

Вычислим изменение функции $V(x, u)$ при переходе из пары точек $[x, u]$ в новую пару $[\bar{x}, \bar{u}]$, определяемую формулой (11). Имеем

$$V(\bar{x}, \bar{u}) = \sum_{i \in J_N^B(x, u)} \bar{x}^i v^i(\bar{u}) + \sum_{i \in J_B^N(x, u)} \bar{x}^i v^i(\bar{u}) = c_1 + c_2\alpha + c_3\alpha^2,$$

где

$$c_1 = \langle x_N^B, v_N^B \rangle = V(x, u),$$

$$c_2 = \langle v_N^B, \Delta x_N^B \rangle + \langle x_N^B, \Delta v_N^B \rangle = -\langle x_N^B, v_N^B \rangle = -V(x, u),$$

$$c_3 = \langle \Delta x_N^B, \Delta v_N^B \rangle + \langle \Delta x_B^N, \Delta v_B^N \rangle.$$

Из (45) вытекает, что $c_3 = 0$. Следовательно, изменение функции $V(x, u)$ не зависит от выбора вектора Δx_B^N и

$$V(\bar{x}, \bar{u}) = (1-\alpha)V(x, u). \tag{48}$$

Поэтому для наибольшего уменьшения значения функции $V(x, u)$ шаг в (11) следует брать максимально возможным при условии, что новая пара точек $[\bar{x}, \bar{u}]$ остается допустимой.

Теорема 3. *Пусть задача* (1) *сильно невырожденна. Тогда в любой допустимой паре точек* $[x, u]$ *имеет место неравенство*

$$\langle b, \Delta u\rangle \geq 0, \tag{49}$$

причем равенство возможно лишь в том случае, когда u – вершина множества U.

Доказательство. Из (28) с учетом (34) получаем

$$\begin{aligned}\langle b, \Delta u\rangle &= \langle b, Qb_1\rangle = \langle b_1, Qb_1\rangle + \langle A_B^N\Delta x_B^N, Qb_1\rangle = \langle b_1, Qb_1\rangle + \langle \Delta x_B^N, \hat{A}_B^N Qb - \hat{A}_B^N QA_B^N\Delta x_B^N\rangle = \\ &= \langle b_1, Qb_1\rangle - \langle \Delta x_B^N, \Delta v_B^N\rangle = \langle b_1, Qb_1\rangle = \langle \hat{H}b_1, (\hat{H}\Gamma H)^{-1}\hat{H}b_1\rangle.\end{aligned} \tag{50}$$

Так как матрица $\hat{H}\Gamma H$ является положительно-определенной, то из (50) следует, что $\langle b, \Delta u\rangle \geq 0$. Это неравенство переходит в равенство, если $\hat{H}b_1 = 0$. Но тогда $\Delta u = 0$ и, по лемме 5, получаем, что u – вершина множества U.

Теорема 4. *Пусть задача* (2) *сильно невырожденна. Тогда в любой допустимой паре точек* $[x, u]$ *имеет место неравенство*

$$\langle c, \Delta x\rangle \leq 0, \tag{51}$$

причем равенство возможно лишь в том случае, когда x – вершина множества X.

Доказательство. На основании (29), (30) имеем $\langle c, \Delta x\rangle = \Delta_1 + \Delta_2$, где

$$\Delta_1 = -\langle c^B, x^B\rangle + \langle c_B^B, (\hat{A}_B^B A_B^B)^{-1}\hat{A}_B^B(I-\Gamma Q)b\rangle + \langle c_N^B, (G_N^B)^2\hat{A}_N^B Qb\rangle,$$

$$\Delta_2 = \langle c_B^N, \Delta x_B^N\rangle - \langle c_B^B(\hat{A}_B^B A_B^B)^{-1}\hat{A}_B^B(I-\Gamma Q)A_B^N\Delta x_B^N\rangle - \langle c_N^B, (G_N^B)^2\hat{A}_N^B QA_B^N\Delta x_B^N\rangle.$$

Воспользуемся равенствами

$$c_B = \hat{A}_B u, \quad c_N^B = v_N^B + \hat{A}_N u, \quad x_N^B = G_N^B\psi_N^B. \tag{52}$$

Из последних двух следует, что

$$\langle c_N^B, x_N^B\rangle = \|\psi_N^B\|^2 + \langle u, A_N^B G_N^B\psi_N^B\rangle. \tag{53}$$

Обозначим

$$Z_N^B = G_N^B\hat{A}_N^B H, \tag{54}$$

$$P = Z_N^B(\hat{Z}_N^B Z_N^B)^{-1}\hat{Z}_N^B, \quad P_1 = A_B^B(\hat{A}_B^B A_B^B)^{-1}\hat{A}_B^B. \tag{55}$$

Аналогично (53) получаем

$$\begin{aligned}\langle c_N^B, (G_N^B)^2\hat{A}_N^B QA_N^B x_N^B\rangle &= \langle c_N^B, G_N^B P\psi_N^B\rangle = \\ &= \langle G_N^B v_N^B, P\psi_N^B\rangle + \langle u, A_N^B G_N^B P\psi_N^B\rangle = \langle \psi_N^B, P\psi_N^B\rangle + \langle u, A_N^B G_N^B P\psi_N^B\rangle.\end{aligned}$$

Тогда

$$\Delta_1 = -\langle \psi_N^B, (I-P)\psi_N^B\rangle - \langle u, (I-P_1)A_N^B G_N^B(I-P)\psi_N^B\rangle.$$

Матрица $I - P_1$ есть матрица ортогонального проектирования на подпространство $(S_B^B)^\perp$. Отсюда следует, что

$$(I-P_1)A_N^B G_N^B(I-P)\psi_N^B = H(\hat{H}H)^{-1}\hat{Z}_N^B(I - Z_N^B(\hat{Z}_N^B Z_N^B)^{-1}\hat{Z}_N^B)\psi_N^B = 0_m.$$

Поэтому, в силу неотрицательной определенности матрицы ортогонального проектирования $I - P$, получаем

$$\Delta_1 = -\langle \psi_N^B, (I-P)\psi_N^B\rangle \leq 0. \tag{56}$$

На основании (52) имеем также

$$\Delta_2 = \langle u, A_B^N \Delta x_B^N \rangle - \langle u, P_1(I - \Gamma Q) A_B^N \Delta x_B^N \rangle - \langle v_N^B, (G_N^B)^2 \hat{A}_N^B Q A_B^N \Delta x_B^N \rangle - \langle u, \Gamma Q A_B^N \Delta x_B^N \rangle =$$

$$= \langle u, (I - P_1)(I - \Gamma Q) A_B^N \Delta x_B^N \rangle - \langle v_N^B, (G_N^B)^2 \hat{A}_N^B Q A_B^N \Delta x_B^N \rangle.$$

Но

$$(I - P_1)(I - \Gamma Q) A_B^N \Delta x_B^N = H(\hat{H}H)^{-1} \hat{H}(I - \Gamma Q) A_B^N \Delta x_B^N = 0_m,$$

следовательно,

$$\Delta_2 = -\langle A_N^B x_N^B, Q A_B^N \Delta x_B^N \rangle = -\langle b, Q A_B^N \Delta x_B^N \rangle = -\langle p, \Delta x_B^N \rangle.$$

Отсюда, с учетом (38), приходим к

$$\Delta_2 = -\langle \Delta x_B^N, \Omega \Delta x_B^N \rangle \le 0, \tag{57}$$

так как, по лемме 3, матрица Ω положительно определена.

Из (56) и (57) окончательно получаем

$$\langle c, \Delta x \rangle = -[\langle \psi_N^B, (I - P)\psi_N^B \rangle + \langle \Delta x_B^N, \Omega \Delta x_B^N \rangle] \le 0. \tag{58}$$

Равенство нулю в (58) возможно только в том случае, когда $\Delta x_B^N = 0$ и $P\psi_N^B = \psi_N^B$, т.е. когда вектор ψ_N^B принадлежит подпространству, порожденному столбцами матрицы Z_N^B. Но тогда из (29), (52) и (55) получаем

$$\Delta x_N^B = G_N^B P \psi_N^B - x_N^B = G_N^B \psi_N^B - x_N^B = 0,$$

поэтому равенство $A_B^B \Delta x_B^B + A_N^B \Delta x_N^B = 0$ переходит в $A_B^B \Delta x_B^B = 0$, из которого, в силу линейной независимости столбцов матрицы A_B^B, следует, что $\Delta x_B^B = 0$. Отсюда приходим к выводу, что равенство в (58) возможно только при $\Delta x = 0$. Из утверждения леммы 7 следует, что x есть вершина множества X. Лемма доказана.

Заметим, что если в регулярной паре $[x, u]$ точка u является вершиной множества U, то обязательно имеет место равенство $\langle b, \Delta u \rangle = 0$. Точно так же если x является вершиной X, то $\Delta x = 0$.

5. ОПИСАНИЕ МЕТОДА

Предположим, что задана начальная допустимая пара точек $[x_0, u_0]$, и пусть на k-й итерации получена допустимая пара точек $[x_k, u_k]$. На следующей, $(k+1)$-й итерации определяем новую пару точек:

$$x_{k+1} = x_k + \alpha_k \Delta x_k, \quad u_{k+1} = u_k + \alpha_k \Delta u_k. \tag{59}$$

Здесь приращения Δx_k и Δu_k вычисляются по формулам (28)–(30) при $x = x_k$, $u = u_k$, причем если $J_B^N(x_k, u_k) \ne \varnothing$, то неотрицательное приращение Δx_B^N выбирается таким образом, чтобы оно было решением линейной задачи дополнительности (38). Шаг α_k полагается равным

$$\alpha_k = \max\{\alpha \ge 0\colon x_k + \alpha \Delta x_k \in X,\ u_k + \alpha \Delta u_k \in U\}. \tag{60}$$

Итерации ведутся до тех пор, пока не будет выполнено условие $V(x_k, u_k) = 0$.

Введем индикаторные векторы y^B и y_N, определив их как

$$y^B = e + D^{-1}(x^B)\Delta x^B, \quad y_N = -D^{-1}(v_N)\Delta v_N. \tag{61}$$

По смыслу они совпадают с индикаторным вектором Тапиа, используемом в прямо-двойственных методах внутренней точки [17]. Согласно (29), (30), (33) и (34) имеют место представления

$$y_B^B = D^{-1}(x_B^B)(\hat{A}_B^B A_B^B)^{-1} \hat{A}_B^B (I - \Gamma Q) b_1, \tag{62}$$

$$y_N^B = D^{-1}(v_N^B) \hat{A}_N^B Q b_1, \tag{63}$$

$$y_N^N = D^{-1}(v_N^N) \hat{A}_N^N Q b_1. \tag{64}$$

Используя векторы y^B и y_N, формулы (11) пересчета компонент векторов x^B и v_N можно переписать в виде

$$\bar{x}^B = D(x^B)[e + \alpha(y^B - e)], \quad \bar{v}_N = D(v_N)[e - \alpha y_N], \tag{65}$$

где $\bar{v}_N = v_N(\bar{u})$.

Пусть $i \in J_N^B(x, u)$. Из (65) следует, что если $0 < y^i < 1$, то компоненты x^i и v^i векторов x и v убывают с ростом шага α. Если $y^i > 1$, то убывает только компонента v^i, а компонента x^i возрастает. Если $y^i < 0$, то, наоборот, x^i убывает, а v^i возрастает. Обозначим через y_*^B минимальную компоненту вектора y^B, через y_N^* – максимальную компоненту вектора y_N. Тогда

$$\alpha_k = \min\{[1 - y_*^B]_+^{-1}, [y_N^*]_+^{-1}\}. \tag{66}$$

Здесь $[a]_+ = \max[0, a]$, векторы y^B и y_N вычислены в паре точек $[x_k, u_k]$.

Определение 4. Степенью неоптимальности допустимой пары точек $[x, u]$ называется число индексов в множестве $J_N^B(x, u)$.

Если степень неоптимальности пары точек $[x, u]$ равна нулю, то, по лемме 1, $[x, u]$ – оптимальная пара. Степень неоптимальности любой другой допустимой пары строго положительна. Степень неоптимальности внутренней пары точек наибольшая и равняется n. Согласно определениям множеств индексов (4), (5), суммарная размерность граней, которым принадлежат точки x и u, равна разности между степенью неоптимальности пары $[x, u]$ и числом индексов в множестве $J_B^N(x, u)$. Поэтому данная размерность максимальна, когда пара $[x, u]$ регулярна.

Нерегулярные пары $[x, u]$ степени неоптимальности единица состоят из вершин допустимых множеств X и U. Регулярные пары той же степени состоят из таких точек, одна из которых является вершиной допустимого множества, а другая принадлежит ребру.

Из формул для нахождения ньютоновских направлений Δx_k, Δu_k из правила выбора шага α_k следует, что

$$J_N^B(x_k, u_k) \subseteq J_N^B(x_{k+1}, u_{k+1}) \cup J_B^B(x_{k+1}, u_{k+1}) \cup J_N^N(x_{k+1}, u_{k+1}),$$

$$J_B^B(x_k, u_k) \subseteq J_B^B(x_{k+1}, u_{k+1}) \cup J_B^N(x_{k+1}, u_{k+1}),$$

$$J_N^N(x_k, u_k) \subseteq J_N^N(x_{k+1}, u_{k+1}) \cup J_B^N(x_{k+1}, u_{k+1}),$$

$$J_B^N(x_k, u_k) \subseteq J_B^N(x_{k+1}, u_{k+1}) \cup J_B^B(x_{k+1}, u_{k+1}) \cup J_N^N(x_{k+1}, u_{k+1}).$$

Отсюда приходим к выводу, что на каждой итерации имеет место включение

$$J_N^B(x_{k+1}, u_{k+1}) \subseteq J_N^B(x_k, u_k), \quad k = 1, 2, \ldots. \tag{67}$$

Таким образом, если на некоторой k-й итерации индекс i выводится из множества $J_N^B(x_k, u_k)$, т.е. $i \notin J_N^B(x_{k+1}, u_{k+1})$, то обратно попасть в множество J_N^B на какой-нибудь последующей итерации он не может. Другими словами, степень неоптимальности не возрастает в ходе итеративного процесса. В дальнейшем те итерации, на которых происходит уменьшение степени неоптимальности пары точек $[x_k, u_k]$, будем называть *активными*. Имеет место следующий результат, вытекающий непосредственно из включения (67).

Теорема 5. *Для любых допустимых начальных пар* $[x_0, u_0]$ *метод* (59) *находит решения обеих задач линейного программирования* (1) *и* (2) *не более чем за* n *активных итераций.*

Отметим, что если итерация не является активной, то следующая пара точек обязательно будет нерегулярной. Поэтому если пара $[x_{k+1}, u_{k+1}]$ оказалась регулярной, то произошла активная итерация. Так как размерность вспомогательной линейной задачи дополнительности (38), которую приходится решать в нерегулярной паре точек, не превышает степень неоптимальности данной пары, то в ходе итеративного процесса максимально возможная размерность таких вспомогательных задач может лишь убывать.

Лемма 8. *Пусть* $[x_k, u_k]$ – *регулярная допустимая пара, и пусть точка* u_k *является вершиной допустимого множества* U, *не совпадающей с оптимальной вершиной* u_*. *Тогда новая пара* $[x_{k+1}, u_{k+1}]$ *является нерегулярной и*

$$J_B^N(x_{k+1}, u_{k+1}) \subseteq J_B^B(x_k, u_k), \quad J_N^B(x_{k+1}, u_{k+1}) = J_N^B(x_k, u_k).$$

Доказательство. Так как $|J_B(u_k)| = m$ и $J_B^N(x_k, u_k) = \varnothing$, то $J_B(u_k) = J_B^B(x_k, u_k)$, $|J_B^B(x_k, u_k)| = m$. Поэтому, на основании формул (22),

$$\Delta u = 0, \quad \Delta x_N^B = -x_N^B, \quad \Delta x_B^B = (A_B^B)^{-1}A_N^B x_N^B.$$

Отсюда сразу получаем, что $u_{k+1} = u_k$. Кроме того, выполняется равенство $\Delta v_N = 0$. Следовательно, $y_N = 0$.

Вычислим y_B^B. Из равенства (27) после его умножения на матрицу $(A_B^B)^{-1}$ получаем

$$x_B^B + (A_B^B)^{-1}A_N^B x_N^B = (A_B^B)^{-1}b.$$

Таким образом, $\Delta x_B^B = (A_B^B)^{-1}b - x_B^B$ и, на основании (61), $y_B^B = D^{-1}(x_B^B)(A_B^B)^{-1}b$.

Так как столбцы матрицы A_B^B образуют допустимый базис в двойственной задаче (2), причем не являющийся базисом оптимальной точки, то они не могут одновременно быть базисом допустимой вершины множества X (в силу предположения о невырожденности обеих задач (1) и (2)). Поэтому среди компонент вектора $(A_B^B)^{-1}b$ найдется по крайней мере одна отрицательная и $y_*^B = y^{i_*}$, где $i_* \in J_B^B(x_k, u_k)$. Поскольку в этом случае, согласно (66), $\alpha_k = (1 - y^{i_*})^{-1} < 1$, то в новой точке x_{k+1} стать равными нулю могут лишь компоненты x_{k+1}^i, индексы которых принадлежат множеству $J_B^B(x_k, u_k)$.

Лемма 9. *Пусть выполнено предположение о сильной невырожденности задачи* (1). *Пусть, кроме того, допустимая пара* $[x_k, u_k]$ *такова, что точка* u_k *является неоптимальной вершиной множества* U. *Тогда за конечное число шагов либо произойдет активная итерация, либо процесс* (59) *покинет данную вершину.*

Доказательство от противного. Предположим, что все последующие итерации являются неактивными и $u_s = u_k$ при $s > k$. Тогда пары $[x_s, u_s]$ оказываются нерегулярными и множества индексов $J_B(u_s)$, $J_N(u_s)$ и $J_N^B(x_s, u_s)$ не меняются от итерации к итерации, т.е. $J_B(u_s) = J_B(u_k)$, $J_N(u_s) = J_N(u_k)$, $J_N^B(x_s, u_s) = J_N^B(x_k, u_k)$. Отсюда, в частности, следует, что $J_N^N(x_s, u_s) = J_N^N(x_k, u_k)$.

Из предположения $\Delta u_s = 0$ при $s \ge k$ вытекает также, что на всех этих итерациях $Hb_1 = 0$, т.е. вектор b_1 принадлежит подпространству, порожденному столбцами a_i матрицы A с индексами $i \in J_B^B(x_s, u_s)$. Кроме того, согласно (29), выполнено $\Delta x_N^B = -x_N^B$, поэтому

$$(x_s)_N^B = \prod_{j=k}^{s-1}(1-\alpha_j)(x_k)_N^B. \tag{68}$$

Имеет место представление

$$b = \sum_{i \in J_B^B(x_k, u_k)} x_k^i a_i + z_k q, \quad q = \sum_{i \in J_N^B(x_k, u_k)} x_k^i a_i, \tag{69}$$

где $z_k = 1$. Так как $J_N^B(x_s, u_s) = J_N^B(x_k, u_k)$, то на основании (68) на всех последующих итерациях получаем

$$b = \sum_{i \in J_B^B(x_s, u_s)} x_s^i a_i + z_s q, \quad z_s = \prod_{j=k}^{s-1}(1-\alpha_j). \tag{70}$$

Обозначим через $\gamma = \{x \in X: x^i = 0, i \in J_N^N(x_k, u_k)\}$ грань множества X размерности $(d - n_N^N)$, которой принадлежит точка x_k. Из неизменности множества индексов $J_N^N(x_s, u_s)$ следует, что и все последующие точки x_s также принадлежат данной грани, причем находятся на ее границе. Пусть $i_j \in J^B(x_k)$, $j = 1, 2, \dots, r$, и пусть $l_{i_1, \dots, i_r}$ есть грань (точнее – подгрань грани) γ, определяемая как

$$l_{i_1, \dots, i_r} = \{x \in \gamma: x^{i_j} = 0,\ j = 1, 2, \dots, r\}.$$

В ходе итераций в прямом пространстве происходит переход с одной такой грани на другую. Так как число данных граней конечно, то можно выделить по крайней мере одну грань, на которой точки из последовательности $\{x_s\}$ оказываются дважды. Предположим, для определенности, $x_{s_1} \in l_{i_1, \dots, i_r}$ и $x_{s_2} \in l_{i_1, \dots, i_r}$, где $s_2 > s_1 > k$. При этом, не умаляя общности, можно считать, что x_{s_1} и x_{s_2} принадлежат относительной внутренности грани $l_{i_1, \dots, i_r}$. Тогда $J_B^B(x_{s_1}, u_{s_1}) = J_B^B(x_{s_2}, u_{s_2}) = J_B^B$. Поскольку все пары $[x_s, u_s]$ при $s > k$ являются нерегулярными, то число индексов в множестве J_B^B меньше, чем m.

Из предположения о сильной невырожденности задачи (1) следует, что вектор b не может быть представлен линейной комбинацией менее чем m столбцов матрицы A. Поэтому векторы a_i $(i \in J_B^B)$ и q линейно независимы и разложение (70) при $s = s_1$ и $s = s_2$ единственно. Но $z_{s_1} > z_{s_2}$ на основании (70), что не может иметь места. Таким образом, за конечное число шагов либо произойдет активная итерация, либо на неактивной итерации процесс (59) покинет вершину u_k. Лемма доказана.

Аналогичные результаты имеют место для вершин множества X.

Лемма 10. *Пусть $[x_k, u_k]$ – регулярная допустимая пара, и пусть точка x_k является вершиной допустимого множества X, не совпадающей с оптимальной вершиной x_*. Тогда новая пара $[x_{k+1}, u_{k+1}]$ является нерегулярной и*

$$J_B^N(x_{k+1}, u_{k+1}) \subseteq J_N^N(x_k, u_k), \quad J_N^B(x_{k+1}, u_{k+1}) = J_N^B(x_k, u_k).$$

Лемма 11. *Пусть выполнено предположение о сильной невырожденности задачи* (2). *Пусть, кроме того, допустимая пара $[x_k, u_k]$ такова, что точка x_k является неоптимальной вершиной множества X. Тогда за конечное число шагов либо произойдет активная итерация, либо процесс* (59) *покинет данную вершину.*

Рассмотрим частный случай процесса (59), когда начальная пара $[x_0, u_0]$ имеет степень неоптимальности, равную единице. При этом предположении любая активная итерация приводит к тому, что преобразованная пара будет оптимальной. Поэтому процесс будет протекать как последовательность неактивных итераций, которая может прерваться только получением оптимальных решений. На неактивных итерациях все пары, за исключением, быть может, начальной, являются нерегулярными и текущие точки являются вершинами допустимых множеств. Покажем, что при этом происходит попеременное движение то в прямом, то в двойственном пространствах.

В нерегулярных парах $[x_k, u_k]$ степени неоптимальности единица множество $J_B^B(x_k, u_k)$ состоит из $m - 1$ индексов, а множества $J_N^B(x_k, u_k)$ и $J_B^N(x_k, u_k)$ – из одного индекса. На выбор ньютоновских направлений Δx_k и Δu_k основное влияние оказывает решение вспомогательной задачи (38), размерность которой равна единице. Поэтому либо $\Delta v_B^N > 0$ и $\Delta x_B^N = 0$, либо $\Delta v_B^N = 0$ и $\Delta x_B^N > 0$; это зависит от того, какое значение (положительное или отрицательное) принимает величина

$$p = \frac{v_B^N \langle a_B^N, h\rangle \langle b, h\rangle}{x_N^B \langle a_N^B, h\rangle^2} = \frac{v_N^B \langle a_B^N, h\rangle}{\langle a_N^B, h\rangle}, \tag{71}$$

где h – ненулевой вектор, ортогональный ко всем столбцам матрицы A_B^B, и a_B^N, a_N^B – единственные столбцы, составляющие матрицы A_B^N и A_N^B соответственно.

Лемма 12. *Пусть степень неоптимальности нерегулярной пары точек* $[x, u]$ *равняется единице. Тогда если* $\Delta x_B^N > 0$, *то* $\Delta u = 0_m$ *и, наоборот,* $\Delta x = 0_n$, *если* $\Delta x_B^N = 0$.

Доказательство. Рассмотрим сначала случай, когда $p > 0$ и, следовательно, $\Delta x_B^N > 0$. Из (38) имеем

$$\Delta x_B^N = \frac{x_N^B \langle a_N^B, h\rangle}{\langle a_B^N, h\rangle} = \frac{\langle b, h\rangle}{\langle a_B^N, h\rangle}.$$

Поэтому

$$b_1 = b - \frac{\langle b, h\rangle}{\langle a_B^N, h\rangle} a_B^N$$

и, стало быть, $\langle h, b_1\rangle = 0$. Отсюда получаем, что $\Delta u = 0_m$.

Допустим теперь, что $p < 0$ и число $\Delta x_B^N = 0$ является решением задачи (38). Тогда, согласно (21) и (29), имеет место равенство

$$A_B^B \Delta x_B^B + A_N^B \Delta x_N^B = 0.$$

Так как x – вершина множества X, то столбцы матрицы $A^B = [A_B^B, A_N^B]$ линейно независимы. Поэтому $\Delta x_B^B = 0$, $\Delta x_N^B = 0$. Отсюда с учетом (29) получаем, что $\Delta x = 0$. Лемма доказана.

Случай, когда $p = 0$, невозможен, поскольку при этом предположении, согласно (71), $\langle a_B^N, h\rangle = 0$ и, следовательно, вектор a_B^N принадлежит подпространству S_B^B. Поэтому столбцы a_i $(i \in J_B(u))$, образующие базис вершины u, являются линейно-зависимыми, что противоречит предположению о невырожденности задачи (2).

Пусть пара $[x_k, u_k]$ степени неоптимальности единица является нерегулярной. Тогда x_k и u_k – вершины своих допустимых множеств. Предположим для определенности, что точка x_k принадлежит ребру $l(x_k, u_k)$ допустимого множества X, определяемому условием

$$l(x_k, u_k) = \{x \in X: x^i = 0, i \in J_N^N(x_k, u_k)\}.$$

У данного ребра может быть не более двух вершин, одна из которых совпадает с точкой x_k. Если ребро ограничено (не является крайним лучом), то имеется другая вершина $\bar{x}$. Если $\bar{x}$ не совпадает с оптимальным решением x_* в задаче (1), то у точки $\bar{x}$ одна из компонент $\bar{x}^i = 0$, где $i \in$ $\in J_B^B(x_k, u_k)$. Кроме того, обязательно $\bar{x}^j > 0$, где j – единственный индекс, содержащийся в множестве $J_B^N(x_k, u_k)$.

В случае когда на k-й итерации $\Delta x_B^N > 0$, согласно утверждению леммы 12, имеем $u_{k+1} = u_k$ и $x_{k+1} = \bar{x}$. Так как $\Delta x_N^N = 0$, то движение в x-пространстве происходит по ребру $l(x_k, u_k)$. На следующей, $(k + 1)$-й итерации обязательно меняется u_{k+1} и остается прежним x_{k+1}. Действительно, иначе опять движение в прямом пространстве происходило бы по ребру $l(x_k, u_k)$ из одной вершины в другую, т.е. из точки x_{k+1} в точку x_k. Но, в силу теоремы 4, при переходе из x_k в x_{k+1} происходит убывание значения целевой функции $\langle c, x\rangle$. При обратном переходе происходило бы ее возрастание, что противоречит опять же утверждению теоремы 4.

Аналогичные рассуждения приводят к тому, что если на k-й итерации $\Delta x_k = 0$ и $\Delta u_k \neq 0$, то на следующей итерации $\Delta x_{k+1} \neq 0$ и $\Delta u_{k+1} = 0$. Таким образом, происходит попеременное движение в прямом и двойственном пространствах по вершинам, причем с убыванием значений целевой функции в прямой задаче (1) и возрастанием – в двойственной (2).

Теорема 6. *Пусть начальная пара* $[x_0, u_0]$ *совпадает с парой вершин допустимых множеств* X *и* U. *Тогда за конечное число итераций будет получена оптимальная пара.*

Доказательство. Результат теоремы вытекает непосредственно из приведенных выше рассуждений, так как число вершин конечно.

Если начальная пара $[x_0, u_0]$ имеет степень неоптимальности, равную двум, то все последующие точки x_k и u_k до тех пор, пока не произойдет активная итерация, принадлежат либо ребрам допустимых множеств, либо являются их вершинами, причем любую неоптимальную вершину процесс покидает самое большее в течение двух последовательных итераций. В [18] показано, что если точки x_0 и u_0 принадлежат ребрам допустимых множеств, прилегающим к оптимальным вершинам, то метод находит решение за одну или две итерации.

СПИСОК ЛИТЕРАТУРЫ

1. *Kojima M., Mizuno S., Yoshise A.* A primal-dual interior point method for linear programming // Progress Math. Program. Interior Point and Related Methods. Berlin: Springer, 1989. P. 29–47.
2. *Monteiro R.C., Adler I.* Interior path-following primal-dual algorithms. Part I: linear programming // Math. Program. 1989. V. 44. P. 27–41.
3. *Todd M.J., Ye Y.* A centered projective algorithm for linear programming // Math. Operat. Res. 1990. V. 15. P. 508–529.
4. *Mizuno S., Todd M.J., Ye Y.* On adaptive step primal-dual interior-point algorithms for linear programming: Techn. Rept 944. Scool Operat. Res. and Industr. Engng. New York: Cornell Univ., Ithaca, 1990. 14853-3801.
5. *Gonzaga C.* Path following methods for linear programming // SIAM Rev. 1992. V. 34. № 2. P. 167–224.
6. *Choi I.C., Monma I.C., Shanno D.F.* Futher development of a primal-dual interior point method for linear programming // ORSA J. Comput. 1990. V. 2. P. 304–311.
7. *Jansen B., Roos C., Terlaky T., Vial J.-Ph.* Primal-dual target-following algorithms for linear programming: Rept № 93-107. Delft Univ. Technol., 1993.
8. *Lustig I.J., Marsten R.E., Shanno D.F.* Computational experience with a primal-dual interior point method for linear programming // Linear Algebra and Appl. 1989. V. 152. P. 191–222.
9. *Lustig I.J., Marsten R.E., Shanno D.F.* On implementing Mehrotra's predictor-corrector interior point method for linear programming // SIAM J. Optimizat. 1992. V. 2. P. 435–449.
10. *Mehrotra S.* Quadratic convergence in primal-dual methods // Math. Operat. Res. 1993. V. 18. P. 741–751.
11. *Roos C., Terlaky T., Vial J.-Ph.* Theory and algorithms for linear optimization. An interior point approach. New York etc.: John Wiley & Sons, 1997.
12. *Wright S.J.* Primal-dual interior-point methods. Philadelphia: SIAM, 1997.
13. *Zhang Y., Tapia R.A.* Superlinear and quadratic convergence of primal-dual interior point methods for linear programming // SIAM J. Optimizat. 1993. V. 3. P. 118–133.
14. *Evtushenko Yu., Zhadan V.* Stable barrier-projection and barrier-Newton methods in nonlinear programming // Optimizat. Meth. and Software. 1994. V. 3. P. 237–356.
15. *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.
16. *Станевичюс А.-И.А., Щербак Л.В.* Новые варианты барьерно-ньютоновских методов для решения задач линейного программирования // Ж. вычисл. матем. и матем. физ. 1995. Т. 35. № 12. С. 1796–1807.
17. *El-Bakry A.S., Tapia R.A., Zhang Y.* A study of indicators for identifying zero variables in interior-point methods // SIAM Rev. 1994. V. 36. № 1. P. 45–72.
18. *Жадан В.Г.* Метод Ньютона с наискорейшим спуском для задач линейного программирования // Сообщ. по вычисл. матем. М.: ВЦ РАН, 1997.