ЖУРНАЛ
ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ ФИЗИКИ

Том 36,1996 N° 7, с. 30–45

УДК 519.6:519.852

© 1996 г. *Ю.Г. ЕВТУШЕНКО, В.Г. ЖАДАН*

ДВОЙСТВЕННЫЕ БАРЬЕРНО-ПРОЕКТИВНЫЕ И БАРЬЕРНО-НЬЮТОНОВСКИЕ МЕТОДЫ ДЛЯ ЗАДАЧ ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ[1]

117967 Москва, ул. Вавилова, 40, ВЦ РАН

(Пересмотрена 20 февраля 2002 г.)

Рассматривается двойственная задача линейного программирования. Для ее решения предлагаются барьерно-проективные и барьерно-ньютоновские методы. Дается обоснование сходимости непрерывных и дискретных вариантов методов и приводятся оценки скорости сходимости.

Введение

В последнее время, особенно после опубликования статьи [1], много внимания уделяется методам внутренней точки для решения задач линейного программирования. Были разработаны целые классы методов, основанные на идеях проектирования, масштабирования и центрального пути. Подробный обзор таких методов содержится в [2]. Большинство проективных методов внутренней точки, таких как метод Кармаркара или аффинный масштабирующий метод (а.м.м.) (см., например, [3]–[8]), предназначены для решения задач линейного программирования в стандартной или в специальной канонической форме. В [8, 9] рассматривались также варианты а.м.м., позволяющего решать двойственную задачу, в которой допустимое множество задается ограничениями типа неравенства.

Авторы на протяжении ряда лет [10]–[15] развивали иной подход к построению численных методов, обладающих свойствами методов внутренней точки. Этот подход основан на использовании сюръективных отображений, и он позволяет строить модификации таких известных методов нелинейного программирования, как метод проекции градиента и метод Ньютона для решения задач выпуклого и общего нелинейного программирования, в которых среди ограничений имеются множества "простой структуры". В [13] эти методы названы, соответственно, барьерно-проективными и барьерно-ньютоновскими. Их варианты, предназначенные для решения задач линейного программирования, приведены в [14]. При специальном выборе сюръективного преобразования и начальных приближений барьерно-проективный метод из [14] совпадает с а.м.м. [3].

Цель настоящей статьи — применить указанный подход для решения двойственной задачи линейного программирования и разработать тем самым семейство двойственных барьерно-проективных и барьерно-ньютоновских методов. Сюръективные преобразования применяются в данной работе для освобождения от требования неотрицательности дополнительных двойственных переменных.

[1]Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (код проекта 96-01-01047).

Основная идея построения алгоритмов излагается в § 1. Здесь с привлечением устойчивого метода проекции градиента [16] строится двойственный барьерно-проективный метод. Исследуются непрерывный и дискретный варианты метода, и выясняется влияние вида преобразования на локальную сходимость.

В § 2 рассматриваются два других барьерно-проективных метода, получающихся при использовании иных, отличных от приведенных в § 1 форм представления двойственной задачи. Показывается, что в качестве частного случая к этим вариантам методов принадлежит а.м.м. из [9].

В § 3 исследуется глобальная сходимость одного из методов при специальном выборе шага спуска. Наконец, в § 4 предлагается двойственный барьерно-ньютоновский метод. При обосновании сходимости методов применяется теория устойчивости Ляпунова.

§ 1. Устойчивый вариант двойственного метода

Рассмотрим задачу линейного программирования

$$\min_{x \in X} c^\top x, \qquad X = \{x \in \mathbb{R}^n : Ax = b, \;\; x \geq 0_n\}. \tag{1.1}$$

Здесь c есть n-мерный вектор, b есть m-мерный вектор, A — матрица размера $m \times n$ полного ранга, в которой $m < n$, символ 0_n обозначает нулевой n-мерный вектор. Столбцами матрицы A являются m-мерные векторы a_i, $i \in \{1, 2, \dots, n\}$.

Двойственной к (1.1) будет задача

$$\max_{u \in U} b^\top u, \qquad U = \{u \in \mathbb{R}^m : v = c - A^\top u \geq 0_n\}. \tag{1.2}$$

Всюду ниже предполагается, что решения обеих задач (1.1), (1.2) существуют и что множество $U_0 = \{u \in \mathbb{R}^m : v = c - A^\top u > 0_n\}$ не пусто.

В задаче (1.2) имеются ограничения типа неравенства. Чтобы освободиться от них, воспользуемся подходом, основанном на сюръективном преобразовании пространств. Для исходной задачи (1.1) он применялся в [14]. Введем в рассмотрение непрерывно дифференцируемую n-мерную вектор-функцию $\varphi(w)$, определенную на $\mathbb{R}^n$, у которой замыкание образа всего пространства $\mathbb{R}^n$ совпадает с неотрицательным ортантом $\mathbb{R}^n_+$. Для простоты считаем, что эта функция имеет покомпонентный вид

$$\varphi(w) = [\varphi^1(w^1), \dots, \varphi^n(w^n)]^\top.$$

Пусть $w^i = \psi^i(v^i)$ — функция, обратная к $\varphi^i(w^i)$. Она существует по крайней мере в тех точках $v^i = \varphi^i(w^i)$, где $\dot{\varphi}^i(w^i) \neq 0$. Обозначим

$$\theta(v) = [\theta^1(v^1), \dots, \theta^n(v^n)]^\top, \qquad G(v) = D(\theta(v)),$$

где

$$\theta^i(v^i) = (\gamma^i(v^i))^2, \quad \gamma^i(v^i) = \dot{\varphi}^i(\psi^i(v^i)), \qquad 1 \leq i \leq n,$$

$D(y)$ — диагональная матрица с i-м диагональным элементом, равным y^i.

На преобразование $\varphi(w)$ наложим два условия.

Условие 1. *Функции $\theta^i(v^i)$, $1 \leq i \leq n$, определены и непрерывны в некоторой окрестности $\mathbb{R}^1_+$ и $\theta^i(v^i) = 0$ в том и только том случае, когда $v^i = 0$.*

Условие 2. *Функции $\theta^i(v^i)$, $1 \leq i \leq n$, непрерывно дифференцируемы в некоторой окрестности $\mathbb{R}^1_+$ и $\dot{\theta}^i(0) > 0$.*

Укажем простейшие примеры преобразования $\varphi(w)$ с соответствующими им функциями $\theta(v)$ и матрицами $G(v)$:

$$\varphi(w) = \frac{1}{4}D(w)w, \quad \theta(v) = v, \qquad G(v) = D(v), \tag{1.3}$$

$$\varphi(w) = e^{-w}, \qquad \theta(v) = D(v)v, \quad G(v) = D^2(v). \tag{1.4}$$

Здесь символ e^{-w} обозначает вектор-функцию с компонентами e^{-w^i}, $1 \le i \le n$. Условие 1 выполняется для обоих преобразований (1.3) и (1.4), условие 2 — только для первого из них.

С помощью преобразования $\varphi(w)$ задача (1.2) может быть сведена к следующей:

$$\max b^\top u, \tag{1.5}$$

$$\varphi(w) - c + A^\top u = 0_n. \tag{1.6}$$

Применим для ее решения устойчивый вариант метода проекции градиента [16]. Обозначая через

$$\tilde{L}(u, w, x) = b^\top u - x^\top(\varphi(w) - c + A^\top u)$$

функцию Лагранжа для задачи (1.5), (1.6), приходим к системе обыкновенных дифференциальных уравнений

$$\frac{du}{dt} = \tilde{L}_u(u, w, x(u, w)), \qquad \frac{dw}{dt} = \tilde{L}_w(u, w, x(u, w)), \tag{1.7}$$

в которой зависимость $x(u, w)$ находится из решения системы линейных уравнений:

$$\tilde{L}_{xu}(u, w, x)\dot{u} + \tilde{L}_{xw}(u, w, x)\dot{w} = -\tau\tilde{L}_x(u, w, x), \quad \tau > 0. \tag{1.8}$$

Так как $v = \varphi_w \dot{w}$, то в пространстве векторов $z = [u, v] \in \mathbb{R}^{m+n}$ метод (1.7), (1.8) принимает вид

$$\frac{du}{dt} = b - Ax(z), \quad \frac{dv}{dt} = -G(v)x(z), \tag{1.9}$$

$$\Phi(v)x(z) = A^\top b + \tau(v + A^\top u - c), \tag{1.10}$$

где $\Phi(v) = G(v) + A^\top A$. Обозначим $v(u) = c - A^\top u$.

Лемма 1. *Пусть преобразование $\varphi(w)$ удовлетворяет условию 1. Тогда $\Phi(v(u))$ — неособая матрица для любого $u \in U_0$.*

Доказательство. В силу условия 1 матрица $G(v(u))$ положительно определена на U_0. Матрица $A^\top A$ является матрицей Грама и, следовательно, неотрицательно определена. Поэтому вся матрица $\Phi(v(u))$ положительно определена. $\square$

Лемма 2. *Пусть выполнено предположение предыдущей леммы. Пусть, кроме того, точка $u \in U$ представима в виде*

$$u = \sum_{j=1}^{s} \alpha_j u_j, \qquad \alpha_j > 0, \quad 1 \le j \le s, \quad \sum_{j=1}^{s} \alpha_j = 1, \tag{1.11}$$

где u_j, $1 \le j \le s$, — угловые точки множества U. Тогда если по крайней мере одна точка u_j не вырождена, то $\Phi(v(u))$ — неособая матрица.

Доказательство. Матрица $\Phi(v)$, где $v = v(u)$, будет неособой, если удастся показать, что равенство
$$\Phi(v)\bar{x} = G(v)\bar{x} + A^\top A\bar{x} = 0_n \tag{1.12}$$
имеет место в том и только том случае, когда $\bar{x} = 0_n$.

Действительно, умножая (1.12) слева на $\bar{x}^\top$, получаем
$$\bar{x}^\top G(v)\bar{x} + \bar{x}^\top A^\top A\bar{x} = 0. \tag{1.13}$$
Так как оба члена в (1.13) неотрицательны, то должно выполняться
$$\bar{x}^\top G(v)\bar{x} = 0, \qquad \bar{x}^\top A^\top A\bar{x} = 0. \tag{1.14}$$

Обозначим
$$S_j = \{1 \le i \le n : \alpha_j^\top u_j = c^i\}, \qquad S = \bigcap_{j=1}^{s} S_j.$$
Если $S = 0$, то $v > 0_n$, и из первого равенства (1.14) получаем, что $\bar{x} = 0_n$. Рассмотрим теперь случай, когда $S \neq 0$. Не ограничивая общности, можно считать, что $S = \{1, 2, \ldots, \ldots, k\}$. Пусть B — подматрица матрицы A, составленная из первых k столбцов A, N — подматрица A, составленная из оставшихся $n-k$ столбцов. В соответствии с разбиением A представим также векторы $\bar{x}$ и v как $\bar{x} = [\bar{x}_B, \bar{x}_N]$, $v = [v_B, v_N]$. Так как по крайней мере одна угловая точка u_j не вырождена, то $k \le m$ и матрица B имеет полный ранг. Кроме того, согласно (1.11) выполняется $v_B = 0_k$, $v_N > 0_{n-k}$. Из первого равенства (1.14) тогда следует, что $\bar{x}_N = 0_{n-k}$. Поэтому второе равенство (1.14) сводится к $\bar{x}_B^\top B^\top B\bar{x}_B = 0$. Но это означает, что $B\bar{x}_B = 0_m$. Поскольку B — матрица полного ранга, то отсюда следует, что $\bar{x}_B = 0_k$ и, стало быть, у вектора $\bar{x}$ все компоненты равны нулю. $\square$

Следствие 1. *Если угловая точка и множества U является невырожденной, то $\Phi(v(u))$ — неособая матрица.*

Следствие 2. *Пусть все угловые точки ограниченного множества U не вырождены. Тогда $\Phi(v(u))$ — неособая матрица для любого $u \in U$.*

Введем множества
$$W = \{v \in \mathbb{R}^n : v = v(u),\ \ u \in \mathbb{R}^m\}, \qquad V = \{v \in \mathbb{R}^n : v = v(u),\ \ u \in U\}. \tag{1.15}$$
Если множество U является выпуклым многогранником со всеми невырожденными угловыми точками, то согласно утверждению следствия 2 у матрицы $\Phi(v)$ существует обратная, когда $v \in V$. В силу непрерывности она будет существовать и в некоторой окрестности V. Для точек v из этой окрестности имеем
$$x(u,v) = [G(v) + A^\top A]^{-1}[A^\top b + \tau(v + A^\top u - c)]. \tag{1.16}$$
Подставляя (1.16) в (1.9), приходим к другой форме записи метода (1.9), (1.10):
$$\begin{aligned} \frac{du}{dt} &= \quad b - A[G(v) + A^\top A]^{-1}[A^\top b + \tau(v + A^\top u - c)], \\ \frac{dv}{dt} &= -G(v)[G(v) + A^\top A]^{-1}[A^\top b + \tau(v + A^\top u - c)]. \end{aligned}$$

Пусть $[u(t,z_0), v(t,z_0)]$ — решение системы (1.9), удовлетворяющее начальному условию $u(t,z_0) = u_0$, $v(t,z_0) = v_0$, $z_0^\top = [u_0^\top, v_0^\top]$. Обозначим $y(u,v) = c - A^\top u - v$. Условие (1.10) можно записать в виде
$$\frac{dy(u,v)}{dt} = y_u^\top(u,v)\dot{u} + y_v^\top(u,v)\dot{v} = -\tau y.$$

Отсюда следует, что система (1.9) имеет первый интеграл

$$c - A^\top u(t, z_0) - v(t, z_0) = (c - A^\top u_0 - v_0)e^{-\tau t}. \tag{1.17}$$

Таким образом, $c - A^\top u(t, z_0) - v(t, z_0) \to 0_n$ при $t \to +\infty$. Кроме того, вдоль траекторий системы согласно (1.10) выполняется

$$\begin{aligned} b^\top \frac{du}{dt} &= b^\top(b - Ax(z)) = \|b - Ax(z)\|^2 + x^\top(z)A^\top(b - Ax(z)) = \\ &= \|b - Ax(z)\|^2 + x^\top(z)G(v)x(z) + \tau x^\top(z)(c - A^\top u - v). \end{aligned} \tag{1.18}$$

Из второго уравнения (1.9) следует, что если преобразование $\varphi(w)$ удовлетворяет условию 1, то каждая компонента вектора $v(t, z_0)$ не меняет знак. Поэтому если $v_0 > 0$, то и вдоль всей траектории $v(t, z_0) > 0$. Отсюда с учетом того, что $y(u(t, z_0), v(t, z_0)) \equiv 0_n$ при $y(u_0, v_0) = 0_n$, получаем, что в случае, когда $u_0 \in U$, можно освободиться от уравнения для v и тем самым упростить систему (1.9). Вместо (1.9), (1.10) приходим к

$$\frac{du}{dt} = b - Ax(u), \tag{1.19}$$

$$[G(v(u)) + A^\top A]x(u) = A^\top b, \tag{1.20}$$

где $u(0, u_0) = u_0 \in U$. Для данной системы вместо (1.18) справедлива формула

$$b^\top \frac{du}{dt} = \|b - Ax(u)\|^2 + x^\top(u)G(v(u))x(u) \geq 0,$$

т.е. целевая функция двойственной задачи (1.2) монотонно возрастает на допустимом множестве. Метод (1.19), (1.20) был впервые предложен в 1977 г. в [11].

Применяя метод Эйлера для интегрирования системы (1.9), (1.10), получаем

$$u_{k+1} = u_k + \alpha_k(b - Ax_k), \qquad v_{k+1} = v_k - \alpha_k G(v_k)x_k, \tag{1.21}$$

$$[G(v_k) + A^\top A]x_k = A^\top b + \tau(v_k + A^\top u_k - c). \tag{1.22}$$

Соответственно для системы (1.19), (1.20) имеем

$$u_{k+1} = u_k + \alpha_k(b - Ax_k), \quad [G(v_k) + A^\top A]x_k = A^\top b, \quad v_k = v(u_k). \tag{1.23}$$

Оба эти варианта метода решают прямую и двойственную задачи (1.1), (1.2) одновременно.

Теорема 1. *Пусть x_* и u_* являются невырожденными решениями соответственно задач (1.1), (1.2) и $v_* = v_*(u_*)$. Пусть, кроме того, преобразование $\varphi(w)$ удовлетворяет условиям 1 и 2. Тогда верно следующее:*

а) точка $z_*^\top = [u_*^\top, v_*^\top]$ является асимптотически устойчивым положением равновесия для системы (1.9), (1.10);

б) решения $u(t, z_0)$, $v(t, z_0)$ системы (1.9), (1.10) локально экспоненциально сходятся к точке z_*, а соответствующая функция $x(z(t, z_0))$ сходится к x_*;

в) существует $\alpha_* > 0$ такое, что для любого фиксированного $0 < \alpha_k < \alpha_*$ последовательность $\{[u_k, v_k]\}$, генерируемая процессом (1.21), (1.22), локально сходится к z_* с линейной скоростью; соответствующая последовательность $\{x_k\}$ сходится к x_*;

г) решения $u(t, u_0)$ системы (1.19), (1.20) локально экспоненциально сходятся к u_* на U, а соответствующая функция $x(u(t, u_0))$ сходится к x_*;

д) существует такое $\alpha_* > 0$, что для любого фиксированного $0 < \alpha_k < \alpha_*$ последовательность $\{u_k\}$, генерируемая процессом (1.23), локально сходится к u_* на U с линейной скоростью; соответствующая последовательность $\{x_k\}$ сходится к x_*.

Доказательство. Составим уравнение в вариациях для системы (1.9), (1.10):

$$\begin{aligned}
\delta\dot{u} &= -A\left[\frac{\partial x(z_*)}{\partial u}\delta u + \frac{\partial x(z_*)}{\partial v}\delta v\right],\\
\delta\dot{v} &= \pm\left\{G(v_*)\frac{\partial x(z_*)}{\partial u}\delta u + \left[D(\dot{\theta}(v_*))D(x_*) + G(v_*)\frac{\partial x(z_*)}{\partial v}\right]\delta v\right\},
\end{aligned}$$

или в матричном виде, введя обозначение $\delta z^\top = [\delta u^\top, \delta v^\top]$:

$$\delta\dot{z} = -Q(z_*)\delta z.$$

Здесь

$$Q(z_*) = \begin{bmatrix} \tau A\Phi_*^{-1}A^\top & A\Phi_*^{-1}[\tau I_n - D(\dot{\theta}(v_*))D(x_*)] \\ \tau G(v_*)\Phi_*^{-1}A^\top & [I_n - G(v_*)\Phi_*^{-1}]D(\dot{\theta}(v_*))D(x_*) + \tau G(v_*)\Phi_*^{-1} \end{bmatrix}, \tag{1.24}$$

где $\Phi_* = G(v_*) + A^\top A$, I_n — единичная матрица порядка n. При вычислении матрицы (1.24) были использованы соотношения, следующие из (1.10):

$$\begin{aligned}
&[G(v) + A^\top A]\frac{\partial x(z)}{\partial u} = \tau A^\top,\\
&D(\dot{\theta}(v))D(x) + [G(v) + A^\top A]\frac{\partial x(z)}{\partial v} = \tau I_n.
\end{aligned}$$

Предположим для определенности, что базис точки x_* состоит из первых m столбцов матрицы A. Тогда для векторов x_*, v_* и матриц A, $G(v_*)$ имеют место представления

$$x_* = \begin{bmatrix} x_*^B \\ x_*^N \end{bmatrix}, \qquad v_* = \begin{bmatrix} v_*^B \\ v_*^N \end{bmatrix}, \qquad A = [B\ N], \qquad G(v_*) = \begin{bmatrix} 0_{mm} & 0_{md} \\ 0_{dm} & G_N \end{bmatrix},$$

где $x_*^B > 0_m$, $v_*^B = 0_m$, $x_*^N = 0_d$, $v_*^N > 0_d$, $d = n - m$, $G_N = D(\theta(v_*^N))$ — правая нижняя квадратная подматрица матрицы $G(v_*)$ порядка d, 0_{ks} — нулевая матрица размера $k \times s$.

Так как при сделанном предположении

$$\Phi_* = \begin{bmatrix} B^\top B & B^\top N \\ N^\top B & G_N + N^\top N \end{bmatrix},$$

то на основании формулы Фробениуса получаем

$$\Phi_*^{-1} = \begin{bmatrix} B^{-1}[I_m + NG_N^{-1}N^\top](B^\top)^{-1} & -B^{-1}NG_N^{-1} \\ -G_N^{-1}N^\top(B^\top)^{-1} & G_N^{-1} \end{bmatrix}.$$

Отсюда вытекает, что

$$\Phi_*^{-1}A^\top = \begin{bmatrix} B^{-1} \\ 0_{dm} \end{bmatrix}.$$

С помощью этого соотношения матрицу Q можно привести к виду

$$Q = \begin{bmatrix} \tau I_m & Q_2 \\ 0_{nm} & Q_1 \end{bmatrix}, \qquad Q_1 = \begin{bmatrix} D(\dot{\theta}(v_*^B))D(x_*^B) & 0_{md} \\ Q_3 & \tau I_d \end{bmatrix}, \tag{1.25}$$

где вид матриц Q_2 и Q_3 несуществен.

Из (1.25) следует, что матрица Q имеет n собственных значений, равных τ, и m собственных значений $\dot{\theta}^i(0)x_*^i$, $1 \leq i \leq m$. Поскольку преобразование $\varphi(w)$ удовлетворяет условию 2, то все они строго положительны. Поэтому по теореме Ляпунова об устойчивости по первому приближению положение равновесия (точка z_*) является асимптотически устойчивым, а решения системы (1.9), (1.10) локально экспоненциально сходятся к z_*.

Сходимость дискретного варианта (1.21), (1.22) для достаточно малых постоянных α_k следует из [12, теорема 2.3.7].

Так как решения системы (1.19), (1.20) при $u_0 \in U$ совпадают с соответствующими решениями более общей системы (1.9), (1.10), если в последней положить $v_0 = v(u_0)$, то решения (1.19), (1.20) локально экспоненциально сходятся к u_* на U. По той же причине последовательность $\{u_k\}$, генерируемая процессом (1.23), локально сходится к u_* на U. Теорема доказана. □

Обозначим через η^* и η_* наибольшее и наименьшее собственные значения матрицы Q:

$$\eta^* = \max\left[\tau, \max_{1\leq i\leq m} \dot{\theta}^i(0)x_*^i\right], \qquad \eta_* = \min\left[\tau, \max_{1\leq i\leq m} \dot{\theta}^i(0)x_*^i\right].$$

Проводя стандартные рассуждения (см. [12]), можно показать, что фигурирующая в утверждении (в) величина $\alpha_* = 2/\eta^*$. При этом скорость сходимости будет наибольшей, если в (1.21), (1.22) шаги взять равными $\alpha_k = 2/(\eta_* + \eta^*)$. Тогда для выполнения условия $\|z_k - z_*\| \leq \varepsilon$ достаточно сделать $\ln(\varepsilon/\|z_* - z_0\|)/\ln q$ итераций, где $q = (\eta^* - \eta_*)/(\eta^* + \eta_*)$.

Для того чтобы определить правую часть в системе (1.19), необходимо найти вектор $x(u)$ и, следовательно, решить систему n линейных уравнений. Если $u \in U_0$, то матрица $G(v(u))$ не вырождена и можно воспользоваться формулой Шермана–Моррисона–Вудбери для определения обратной матрицы:

$$[G(v) + A^\top A]^{-1} = G^{-1}(v)[I_n - A^\top(I_m + AG^{-1}(v)A^\top)^{-1}AG^{-1}(v)].$$

Отсюда приходим к системе

$$\frac{du}{dt} = [I_m + AG^{-1}(v(u))A^\top]^{-1}b, \qquad u(t,0) = u_0 \in U_0. \tag{1.26}$$

Локальная сходимость метода (1.26) и его дискретного варианта к решению задачи (1.2) на U_0 опять же следует из более общих утверждений (б) и (г) теоремы 1.

Отметим также, что условие (1.10) в методе (1.9), (1.10) можно было бы заменить на любое другое, обеспечивающее стремление компонент вектора $y(u,v)$ к нулю. Например, вместо (1.10) можно было взять

$$[G(v) + A^\top A]x(z) = A^\top b + \tau D(v + A^\top u - c)(v + A^\top u - c), \qquad \tau > 0.$$

Тогда система (1.9), (1.10) вместо (1.17) имела бы первый интеграл

$$c - A^\top u(t,z_0) - v(t,z_0) = D^{-1}(c - A^\top u_0 - v_0 + \tau t)(c - A^\top u_0 - v_0).$$

Утверждение теоремы 1 при этом полностью сохраняется.

§ 2. Другие варианты двойственных барьерно-проективных методов

Согласно сделанному предположению ранг матрицы A равен m и ее нуль-пространство имеет размерность $d = n - m$. Пусть P — такая матрица полного ранга, что $AP^\top = 0_{md}$. Так как строки матрицы P линейно независимы, то они образуют базис в нуль-пространстве матрицы A. Если матрица A представима в блочном виде $A = [B\ N]$, где квадратная матрица B не вырождена, то в качестве P можно взять, например, матрицу

$$P = \left[\ -N(B^\top)^{-1}\ \middle|\ I_d\ \right].$$

Определения (1.15) множеств W и V с помощью матрицы P можно переписать в виде

$$W = \{v \in \mathbb{R}^n : P(v - c) = 0_d\}, \qquad V = \{v \in \mathbb{R}^n_+ : P(v - c) = 0_d\}.$$

Пусть $\bar{x} \in \mathbb{R}^n$ — произвольный вектор, удовлетворяющий условию $A\bar{x} = b$. Тогда

$$\max_{u\in U} b^\top u = \max_{u\in U} \bar{x}^\top A^\top u = \max_{v\in V} \bar{x}^\top (c - v) = \bar{x}^\top c - \min_{v\in V} \bar{x}^\top v.$$

Следовательно, решение двойственной задачи (1.2) может быть заменено решением эквивалентной задачи минимизации

$$\min_{v\in V} \bar{x}^\top v. \tag{2.1}$$

Устойчивый вариант барьерно-проективного метода из [14], примененный к задаче (2.1), приводит к формулам

$$\frac{dv}{dt} = -G(v)[\bar{x} - P^\top x(v)], \tag{2.2}$$

$$PG(v)P^\top x(v) = PG(v)\bar{x} + \tau P(c - v). \tag{2.3}$$

В тех точках $v \in \mathbb{R}^n$, где матрица $PG(v)P^\top$ неособая, разрешая уравнение (2.3), получаем

$$x(v) = [PG(v)P^\top]^{-1}[PG(v)\bar{x} + \tau P(c - v)].$$

Обозначим $H(v) = G^{1/2}(v)$. Кроме того, введем в рассмотрение правую псевдообратную матрицу $(PH)^+ = (PH)^\top (PGP^\top)^{-1}$ и матрицу проектирования $(PH)^\# = (PH)^+ PH$. Тогда метод (2.2), (2.3) можно записать в следующей проективной форме:

$$\frac{dv}{dt} = H[\tau (PH)^+ P(c - v) - (I_n - (PH)^\#)H\bar{x}]. \tag{2.4}$$

Первый вектор, стоящий в квадратных скобках, принадлежит нуль-пространству матрицы AH^{-1}, второй вектор принадлежит пространству строк матрицы AH^{-1}. Имеют место формулы

$$P\frac{dv}{dt} = \tau P(c - v), \qquad P(c - v(t, v_0)) = P(c - v_0)e^{-\tau t},$$

откуда видно, что $v(t, v_0)$ приближается к множеству W при $t \to \infty$.

Если $v_0 \in V_0$, где $V_0 = \{v \in V : v > 0_n\}$, то (2.4) обладает свойствами метода внутренней точки; целевая функция $\bar{x}^\top v(v_0, t)$ монотонно убывает и $v(t, v_0) \in V_0$ для всех $t \geq 0$. В этом случае метод (2.4) может быть переписан в виде

$$\frac{dv}{dt} = -G(v)[I_n - P^\top (PG(v)P^\top)^{-1} PG(v)]\bar{x}, \qquad v_0 \in V_0. \tag{2.5}$$

Теорема 2. *Пусть выполнены предположения теоремы* 1. *Тогда верно следующее:*

а) точка v_* является асимптотически устойчивым положением равновесия для системы (2.2), (2.3);

б) решения системы (2.2), (2.3) локально экспоненциально сходятся к точке v_*;

в) существует такое $\alpha_* > 0$, что для любых постоянных $0 < \alpha_k < \alpha_*$ дискретный вариант метода

$$v_{k+1} = v_k - \alpha_k G(v_k)(\bar{x} - P^\top x_k), \qquad x_k = x(v_k),$$

локально сходится к v_* с линейной скоростью.

Доказательство этой теоремы аналогично доказательству теоремы 1, только теперь у матрицы Q будет d собственных значений, равных τ, и m собственных значений, равных $\dot{\theta}^i(0)x_*^i$, $1 \le i \le m$. Выбор вектора $\bar{x}$ в (2.1), таким образом, не влияет на скорость сходимости.

Так как для системы (2.5) имеет место равенство $P\dot{v} = 0$, то вектор $\dot{v}$ в данном методе принадлежит нуль-пространству матрицы P, которое совпадает с пространством строк матрицы A. Поэтому наряду с (2.5) имеет место представление

$$\dot{v} = A^\top \lambda \tag{2.6}$$

для некоторого вектора $\lambda \in \mathbb{R}^m$. Если $v > 0_n$, то после умножения обеих частей равенства (2.6) на матрицу $AG^{-1}(v)$ с учетом того, что $\dot{v}$ имеет вид (2.5), получаем

$$\lambda = -[AG^{-1}(v)A^\top]^{-1}A\bar{x} = -[AG^{-1}(v)A^\top]^{-1}b. \tag{2.7}$$

Подставляя (2.7) в (2.6), приходим к другой форме записи метода (2.5):

$$\frac{dv}{dt} = -A^\top[AG^{-1}(v)A^\top]^{-1}b, \qquad v_0 \in V_0. \tag{2.8}$$

В пространстве переменных u метод (2.8) принимает вид

$$\frac{du}{dt} = [AG^{-1}(v(u))A^\top]^{-1}b, \qquad u_0 \in U_0. \tag{2.9}$$

Если воспользоваться преобразованиями (1.3) и (1.4), то из (2.9) получаем соответственно

$$\frac{du}{dt} = [AD^{-1}(v(u))A^\top]^{-1}b, \qquad u_0 \in U_0, \tag{2.10}$$

$$\frac{du}{dt} = [AD^{-2}(v(u))A^\top]^{-1}b, \qquad u_0 \in U_0. \tag{2.11}$$

Формула (2.11) совпадает с непрерывным вариантом двойственного а.м.м., предложенного в [9].

Как следует из утверждения теоремы 2, решения системы (2.2), (2.3) локально экспоненциально сходятся к $v_* = v(u_*)$ на множестве V. Поэтому решения системы (2.10) также локально экспоненциально сходятся к u_* на множестве U_0.

Если рассмотреть дискретный аналог метода (2.10)

$$u_{k+1} = u_k + \alpha_k[AD^{-1}(v_k)A^\top]^{-1}b, \qquad u_0 \in U_0, \tag{2.12}$$

где $v_k = v(u_k)$, то согласно [12] из экспоненциальной сходимости непрерывного варианта (2.10) следует локальная сходимость итеративного процесса (2.12) с линейной скоростью к u_* для достаточно малых постоянных α_k.

Еще один вариант двойственного барьерно-проективного метода получается, если представить (1.2) как задачу с $2n$ ограничениями типа равенства:

$$\max b^\top u, \tag{2.13}$$
$$c - A^\top u - v = 0_n, \tag{2.14}$$
$$v - \varphi(w) = 0_n. \tag{2.15}$$

Обозначим $\tilde z^\top = [u^\top, v^\top, w^\top]$ и составим для (2.13), (2.14), (2.15) функцию Лагранжа

$$\tilde L(\tilde z, x, y) = b^\top u + x^\top (c - A^\top u - v) + y^\top (v - \varphi(w)).$$

Аналогом метода (1.7) для задачи (2.13), (2.14), (2.15) является

$$\dot u = \tilde L_u = b - Ax(\tilde z), \qquad \dot v = \tilde L_v = y(\tilde z) - x(\tilde z), \qquad \dot w = \tilde L_w = -\varphi_w^\top(w) y(\tilde z), \tag{2.16}$$

где зависимости $x(\tilde z)$ и $y(\tilde z)$ находятся из условий

$$\tilde L_{xu}\dot u + \tilde L_{xv}\dot v + \tilde L_{xw}\dot w = -\tau \tilde L_x, \qquad \tilde L_{yu}\dot u + \tilde L_{yv}\dot v + \tilde L_{yw}\dot w = -\tau \tilde L_y. \tag{2.17}$$

В переменных u, v и $p = \varphi(w)$ система (2.16), (2.17) может быть переписана в виде

$$\frac{du}{dt} = b - Ax(z), \qquad \frac{dv}{dt} = y(z) - x(z), \qquad \frac{dp}{dt} = -G(p)y(z), \tag{2.18}$$
$$(I_n + A^\top A)x(z) - y(z) = A^\top b - \tau(c - A^\top u - v), \tag{2.19}$$
$$(I_n + G(p))y(z) - x(z) = \tau(p - v), \tag{2.20}$$

где $z^\top = [u^\top, v^\top, p^\top]$. Отсюда, если взять $p_0 = v_0$, то вдоль всей траектории $z(t, z_0)$ выполняется $p(t, z_0) \equiv v(t, z_0)$ и из (2.20) следует, что $y(z) = [I_n + G(p)]^{-1}x(z)$. Поэтому в этом случае метод (2.18), (2.19), (2.20) упрощается: вместо (2.18), (2.19), (2.20) имеем

$$\frac{du}{dt} = b - Ax(u, v), \qquad \frac{dv}{dt} = -G(v)[I_n + G(v)]^{-1}x(u, v), \tag{2.21}$$
$$\{G(v)(1 + G(v))^{-1} + A^\top A\}x(u, v) = A^\top b - \tau(c - A^\top u - v). \tag{2.22}$$

Кроме того, если дополнительно предположить, что $v_0 = v(u_0)$, то $v(t, z_0) \equiv v(u(t, z_0))$ и метод (2.21), (2.22) с помощью формулы Шермана–Моррисона–Вудбери на множестве U_0 принимает вид

$$\frac{du}{dt} = \{I_n + A[I_n + G^{-1}(v(u))]A^\top\}^{-1} b, \qquad u_0 \in U_0. \tag{2.23}$$

Локальная сходимость метода (2.23) и его дискретного варианта

$$u_{k+1} = u_k + \alpha_k \{I_n + A[I_n + G^{-1}(v(u_k))]A^\top\}^{-1} b$$

для случая, когда $u_0 \in U_0$, $G(v) = D(v)$ и шаг α_k постоянный и достаточно малый, следует из асимптотической устойчивости точки $[u_*, v_*, p_*]$, где $p_* = v_* = \varphi(w_*)$, для более общей системы (2.18), (2.19), (2.20).

§ 3. Глобальная сходимость методов

Рассмотрим вопрос о глобальной сходимости метода (2.12) на множестве U_0.

Предположим, что задача (1.1) такова, что

$$A\bar{e} = 0_m, \tag{3.1}$$

где $\bar{e}^\top = [1, \dots, 1] \in \mathbb{R}^n$. Считаем также, что в задаче (1.2) существует единственное решение u_*. Тогда обязательно $C = c^\top \bar{e} > 0$. Пусть $v_* = v(u_*)$ и $J_*^N = \{i \in \{1, 2, \dots, n\} : v_*^i > 0\}$. Введем в рассмотрение функцию Ляпунова

$$F(u) = \sum_{i \in J_*^N} v_*^i [\ln v_*^i - \ln v^i(u)]. \tag{3.2}$$

Функция $F(u)$ определена, непрерывно дифференцируема и неотрицательна на множестве $U_1 = \{u \in U : v^i(u) > 0, \ \ i \in J_*^N\}$. Действительно, так как согласно (3.1)

$$\sum_{i \in J_*^N} v_*^i = \bar{e}^\top v_* = \bar{e}^\top v(u) = \bar{e}^\top c = C > 0, \tag{3.3}$$

то из неравенства между средним арифметическим и средним геометрическим следует, что для любого $u \in U_1$

$$F(u) = -C \sum_{i \in J_*^N} \frac{v_*^i}{C} \ln \frac{v^i(u)}{v_*^i} = -C \ln \prod_{i \in J_*^N} \left[\frac{v^i(u)}{v_*^i}\right]^{v_*^i/C} \geq -C \ln \sum_{i \in J_*^N} \frac{v^i(u)}{C} = 0,$$

причем равенство возможно в том и только том случае, когда $u = u_*$.

Вычислим производную функции (3.2) в силу системы (2.10). Имеем

$$\frac{dF(u)}{dt} = F_u^\top \dot{u} = v_*^\top D^{-1}(v(u)) A^\top [A D^{-1}(v(u)) A^\top]^{-1} b. \tag{3.4}$$

Обозначим

$$p(u) = [A D^{-1}(v(u)) A^\top]^{-1} b, \qquad x(u) = D^{-1}(v(u)) A^\top p(u).$$

Определенный таким образом вектор $x(u)$ удовлетворяет равенству $Ax(u) = b$. Кроме того, согласно (3.1) $x^\top(u) v(u) = \bar{e}^\top A^\top p(u) = 0$. Поэтому

$$x^\top(u) c = x^\top(u)(v(u) + A^\top u) = u^\top A x(u) = b^\top u.$$

Отсюда и из (3.4) получаем

$$\frac{dF(u)}{dt} = v_*^\top x(u) = x^\top(u)(c - A^\top u_*) = b^\top u - b^\top u_* \leq 0, \tag{3.5}$$

причем равенство возможно только тогда, когда $u = u_*$.

Для произвольного $u_0 \in U_0$ обозначим $Q = \{u \in U_1 : F(u) \leq F(u_0)\}$. Это множество компактно, так как в силу (3.3) компактно множество V, а следовательно и множество U. Кроме того, множество Q не содержит угловых точек U, за исключением u_*. Из неравенства (3.5) следует, что $u(t, u_0) \in Q$ для всех $t \geq 0$.

Положим

$$K = \inf_{u \in Q} \frac{\langle b, u_* - u \rangle}{F(u)}. \tag{3.6}$$

Тогда на основании (3.5) в (3.6) получаем, что

$$F(u(t,u_0)) \le F(u_0)e^{-Kt}, \qquad t \ge 0.$$

Лемма 3. *Пусть в задаче* (1.2) *существует единственное невырожденное решение* u_*. *Тогда для величины* K *имеет место оценка*

$$K \ge \frac{1-\exp\left[-F(u_0)/C\right]}{F(u_0)} \min_{1\le j\le m} s_j > 0, \tag{3.7}$$

где $s_j = b^\top(u_* - u_j)$, u_j — *смежная с* u_* *вершина многогранника* U, $1 \le j \le m$.

Доказательство. Введем переменные $z = u - u_*$. В этих переменных функция $F(u)$ и формула (3.6) для определения величины K принимают вид

$$\tilde F(z) = -\sum_{i\in J_*^N} v_*^i \ln\left\{1 - \frac{a_i^\top z}{v_i^*}\right\}, \qquad K = -\sup_{z\in Q_1}\frac{\langle b,z\rangle}{\tilde F(z)},$$

где $Q_1 = \{z \in Z : \tilde F(z) \le F(u_0)\}$, $Z = \{z \in \mathbb{R}^m : A^\top z \le v_*\}$. Функция $\tilde F(z)$ выпукла по z на Q_1. Имеем $\tilde F(0) = 0$ и $\tilde F(z) > 0$, $\langle b,z\rangle < 0$ для всех $z \in Z$, $z \ne 0_m$. Поэтому для любой точки $\bar z \in S = \{z \in Q_1 : \tilde F(z) = F(u_0)\}$ и любого $0 < \alpha \le 1$ выполняется $\tilde F(\alpha\bar z) \le \alpha\tilde F(\bar z)$. Отсюда следует, что

$$\frac{\langle b,\alpha\bar z\rangle}{\tilde F(\alpha\bar z)} \le \frac{\langle b,\bar z\rangle}{\tilde F(\bar z)}, \qquad K = -\frac{1}{F(u_0)}\max_{z\in S}\langle b,z\rangle. \tag{3.8}$$

Точка $z = 0$ является вершиной многогранника Z. Пусть z_j — другие вершины этого многогранника, смежные с ней, и пусть β_j — решение уравнения

$$\sum_{i\in J_*^N} v_*^i \ln(1 - \beta_j q_{ij}) + F(u_0) = 0, \tag{3.9}$$

где $q_{ij} = a_i^\top z_j / v_*^i$. Поскольку $\tilde F(z_j) = +\infty$, то $0 < \beta_j < 1$. Имеем

$$\max_{z\in S}\langle b,z\rangle = \max_{1\le j\le n}\beta_j\langle b,z_j\rangle = -\min_{1\le j\le m}\beta_j s_j < 0. \tag{3.10}$$

Так как $A^\top z_j \le v_*$, то $q_{ij} \le 1$ для всех $i \in J_*^N$, причем по крайней мере для одного индекса i выполняется $q_{ij} = 1$. Поэтому

$$\ln(1 - \beta_j q_{ij}) \ge \ln(1 - \beta_j)$$

и, следовательно, любое решение β_j уравнения (3.9), $1 \le j \le m$, удовлетворяет неравенству $\beta_j \ge \bar\beta$, где $\bar\beta$ — решение уравнения

$$\ln(1 - \bar\beta)\sum_{i\in J_*^N} v_*^i + F(u_0) = 0.$$

Отсюда

$$\bar\beta = 1 - \exp[-F(u_0)/C]. \tag{3.11}$$

Из (3.8), (3.10) и (3.11) приходим к оценке (3.7). Лемма доказана. $\square$

Обозначим

$$\mu(u) = \max_{1\le i\le n} x^i(u).$$

Величина $\mu(u) > 0$ для любого $u \in U_0$. Действительно, если $\mu(u) \le 0$, то $x(u) \le 0_n$, причем $x^i(u) < 0$ хотя бы для одного индекса i. Тогда для любого $\alpha > 0$ выполняется $\alpha x(u) \le 0_n < \bar{e}$. После умножения этого неравенства на матрицу $D(v(u))$ получаем $\alpha A^\top[AD^{-1}(v)A^\top]^{-1}b \le v(u)$, или

$$A^\top\{u + \alpha[AD^{-1}(v)A^\top]^{-1}b\} \le c. \tag{3.12}$$

Таким образом, вектор $u+\alpha[AD^{-1}(v)A^\top]^{-1}b$ принадлежит множеству U для любого $\alpha > 0$, что противоречит ограниченности множества U. Из (3.12) следует также, что величина $1/\mu(u)$ является верхней границей для α, при которой $u + \alpha x(u) \in U$.

Теорема 3. *Пусть шаг α_k в (2.12) выбирается из условия*

$$0 < \alpha_k = \gamma/\mu(u_k), \qquad 0 < \gamma < 1. \tag{3.13}$$

Тогда для любого $u_0 \in U_0$ существует такое $0 < \gamma(u_0) < 1$, что для всех $0 < \gamma \le \gamma(u_0)$ и $k \ge 0$ имеет место оценка

$$F(u_{k+1}) \le F(u_k)(1 - 0.5\alpha_k K), \tag{3.14}$$

где величина K определяется согласно (3.6).

Доказательство. Вычислим изменение функции (3.2) за один шаг итеративного процесса. Имеем

$$\begin{aligned} F(u_{k+1}) &= -\sum_{i\in J_*^N} v_*^i \ln\left[1 - \frac{\alpha_i^\top(u_k + \alpha_k p_k - u_*)}{v_i^*}\right] = \\ &= -\sum_{i\in J_*^N} v_*^i \ln\left[\frac{v_k^i}{v_*^i}(1 - \alpha_k x_k)\right] = F(u_k) - \sum_{i\in J_*^N} v_*^i \ln(1 - \alpha_k x_k^i). \end{aligned} \tag{3.15}$$

Здесь $p_k = p(u_k)$, $x_k = x(u_k)$. Обозначим

$$\Delta(u,\alpha) = \alpha^{-1} \sum_{i\in J_*^N} v_*^i \ln[1 - \alpha^i x(u)], \qquad \alpha > 0. \tag{3.16}$$

Разлагая в ряд правую часть (3.16), получаем

$$\Delta(u,\alpha) = -v_*^\top x(u) - \frac{\alpha}{2}\sum_{i\in J_*^N} \frac{v_*^i[x^i(u)]^2}{[1 - \alpha\theta^i(u)x^i(u)]^2},$$

где $0 \le \theta^i(u) \le 1$, $i \in J_*^N$. Отсюда и из (3.5) приходим к неравенству

$$\Delta(u,\alpha) \ge b^\top(u_* - u) - \frac{\gamma}{2(1-\gamma)^2\mu(u)} \sum_{i\in J_*^N} v_*^i[x^i(u)]^2, \tag{3.17}$$

справедливому для любого $\alpha \le \gamma/\mu(u)$.

Обозначим

$$r(u) = \mu(u) b^\top(u_* - u)\left[\sum_{i\in J_*^N} v_*^i[x^i(u)]^2\right]^{-1}, \qquad \bar{r} = \inf_{u\in Q} r(u)$$

и покажем, что $\bar{r} > 0$. Это неравенство действительно имеет место, если инфимум достигается в точке $u \in Q$, не совпадающей с u_*. Покажем, что и в случае, когда существует такая сходящаяся к точке u_* последовательность $\{u_s\}$, что $\bar{r} = \lim_{s\to\infty} r(u_s)$, опять $\bar{r} > 0$.

Пусть для определенности точка u_* такова, что $v_*^\top = [v_*^B, v_*^N]$, где $v_*^B \in \mathbb{R}^m$, $v_*^N \in \mathbb{R}^d$, $v_*^B = 0_m$, $v_*^N > 0_d$. То же самое разбиение будем использовать для произвольного вектора $v(u)$, а также для матрицы $A = [BN]$. Обозначим $\Gamma^B(u) = BD^{-1}(v^B(u))B^\top$, $\Gamma^N(u) = ND^{-1}(v^\top(u))N^\top$. Так как B — невырожденная матрица, то для всех $u \in U_0$

$$\Gamma(u) = AD^{-1}(v(u))A^\top = \Gamma^B(u) + \Gamma^N(u) = \Gamma^B(u)[I + (\Gamma^B(u))^{-1}\Gamma^N(u)].$$

Поэтому

$$\Gamma^{-1}(u) = \{I - (\Gamma^B(u))^{-1}\Gamma^N(u) + [(\Gamma^B(u))^{-1}\Gamma^N(u)]^2 - \ldots\}(\Gamma^B(u))^{-1} = (\Gamma^B(u))^{-1} + \Phi(u),$$

где $\|\Phi(u)\| = o(\|u - u_*\|)$. Отсюда получаем

$$\begin{array}{rcl} x^B(u) & = & D^{-1}(v^B)B^\top(\Gamma^B(u))^{-1}b + D^{-1}(v^B)B^\top\Phi(u)b = x_*^B + \varphi_1(u), \\ x^N(u) & = & D^{-1}(v^N)N^\top\Gamma^{-1}(u)b = \varphi_2(u), \\ \mu(u) & = & \max\limits_{1\le i\le m} x_*^i + \varphi_3(u), \\ \|\varphi_i(u)\| & = & O(\|u - u_*\|), \qquad i = 1, 2, 3. \end{array}$$

Пусть последовательность $u_s \to u_*$, $u_s \in U_0$. Если $\bar{r} = 0$, то $r(u_s) < 1$ для достаточно больших s. Но из того, что $\|x^N(u)\| = O(\|u - u_*\|)$, получаем

$$\sum_{i\in J_*^N} v_*^i(x^i(u)) = o(\|u - u_*\|),$$

и, стало быть, неравенства $r(u_s) < 1$ при больших s не могут иметь места. Полученное противоречие показывает, что $\bar{r} > 0$.

Так как $\bar{r} > 0$, то найдется такое достаточно малое $0 < \gamma(u_0) < 1$, что

$$\gamma(1-\gamma)^{-2}\mu^{-1}(u)\sum_{i\in J_*^N} v_*^i(x^i(u))^2 \le b^\top(u_* - u)$$

для всех $0 < \gamma \le \gamma(u_0)$ и $u \in Q$. Поэтому для этих u, γ и $\alpha \le \gamma/\mu(u)$ согласно (3.17) $\Delta(u, a) \ge b^\top(u_* - u)/2$. Отсюда и из (3.15), учитывая неравенство $b^\top(u_* - u) \le KF(u)$, приходим к искомой оценке (3.14), что и завершает доказательство теоремы. $\square$

Обозначим

$$B(u_0) = \max_{u\in Q(u_0)}\max_{1\le j\le n} x^i(u), \qquad \bar{\alpha}(u_0) = \frac{\gamma}{B(u_0)}.$$

Тогда если шаг α_k выбирается из условия (3.13), то $\alpha_k \ge \bar{\alpha}(u_0)$ для любого $k \ge 0$. Поэтому наряду с (3.14) имеет место неравенство

$$F(u_{k+1}) \le F(u_k)(1 - 0.5\alpha K), \tag{3.18}$$

где $0 < \alpha \le \bar{\alpha}(u_0)$. С помощью (3.18) можно оценить число шагов, необходимых для попадания процесса (2.12) в некоторую окрестность точки u_*. Неравенство (3.18) сохраняется и для процесса (2.12) с постоянным шагом $\alpha_k = \alpha \le \bar{\alpha}(x_0)$.

§ 4. Двойственный барьерно-ньютоновский метод

Если подставить выражение из (1.20) для $x(u)$ в условие допустимости, то придем к следующему уравнению:

$$b - Ax(u) = 0. \tag{4.1}$$

Применим для решения (4.1) метод Ньютона. Его непрерывный вариант приводит к системе

$$\Lambda(u)\frac{du}{dt} = Ax(u) - b, \tag{4.2}$$

где $\Lambda(u)$ — полная производная по u вектор-функции $-Ax(u)$.

Дифференцируя тождество (1.20) по u, получаем

$$-D(\dot\theta(v))D(x)A^\top + (D(\theta(v)) + A^\top A)\frac{dx}{du} = 0.$$

Поэтому

$$\Lambda(u) = -A[D(\theta(v(u))) + A^\top A]^{-1}D(\dot\theta(v(u)))D(x(u))A^\top,$$

и если эта матрица неособая, то метод (4.2) может быть переписан в виде

$$\frac{du}{dt} = [A[D(\theta(v(u))) + A^\top A]^{-1}D(\dot\theta(v(u)))D(x(u))A^\top]^{-1}[b - Ax(u)]. \tag{4.3}$$

Лемма 4. *Пусть решения x_* и u_* обеих задач линейного программирования* (1.1) *и* (1.2) *не вырождены, выполнены условия* 1 *и* 2. *Тогда матрица $\Lambda(u_*)$ — неособая.*

Доказательство. При $v_* = v_*(u_*)$ выполняется $(D(\theta(v_*)) + A^\top A)x_* = A^\top b$ и мы имеем $x(u_*) = x_*$. Предположим для определенности, что базис точки x_* образуют первые m столбцов матрицы A. Тогда имеет место представление $A = [BN]$, где B — квадратная неособая матрица. Учитывая, что в этом случае

$$A[D(\theta(v_*)) + A^\top A]^{-1} = [(B^\top)^{-1} \mid 0_{md}],$$

получаем

$$\Lambda(u_*) = (B^\top)^{-1}D(\dot\theta(v_*^B)D(x_*^B)B. \tag{4.4}$$

Все квадратные матрицы, входящие в правую часть (4.4), не вырождены, поэтому матрица $\Lambda(u_*)$ неособая. □

Утверждение леммы 4 дает возможность сформулировать теорему о локальной сходимости метода (4.3) и его дискретного варианта

$$u_{k+1} = u_k + [A[D(\theta(v_k)) + A^\top A]^{-1}D(\dot\theta(v_k))D(x_k)A^\top]^{-1}(b - Ax_k), \tag{4.5}$$

где $v_k = v(u_k)$, $x_k = x(u_k)$.

Теорема 4. *Пусть выполнены предположения леммы* 4. *Тогда точка u_* является асимптотически устойчивым положением равновесия для системы* (4.3). *Если, кроме того, матрица $\Lambda(u)$ удовлетворяет условию Липшица в некоторой окрестности u_*, то последовательность $\{u_k\}$, генерируемая процессом* (4.5), *локально сходится к u_* с квадратичной скоростью.*

Вид метода (4.3) несколько упрощается, если в качестве $\varphi(w)$ используется преобразование (1.3):

$$\frac{du}{dt} = [A[D(v(u)) + A^\top A]^{-1} D(x(u)) A^\top]^{-1}[b - Ax(u)].$$

Его дискретный вариант аналогичен (4.5).

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

1. *Karmarkar N.* A new polynomial time algorithm for linear programming // Combinatorica. 1984. V.4. P. 373–395.

2. *Gonzaga C.C.* Path-following algorithms for linear programming // SIAM Rev. 1992. V.34. N° 2. P. 167–224.

3. *Дикин И.И.* Итеративное решение задач линейного и квадратичного программирования // Докл. АН СССР. 1967. Т.174. N° 4. С. 747–748.

4. *Barnes E.* A variation on Karmarkar's algorithm for solving linear programming problelms // Math. Program. 1986. V.36. P. 174–182.

5. *Vanderbei R., Meketon M., Freedman B.* A modification of Karmarkar's linear programming algorithm // Algorithmica. 1986. N° 1. P. 395–407.

6. *Bayer D.A., Lagarias J.C.* The nonlinear geometry of linear programming. Affine and projective scaling trajectories // Trans. Amer. Math. Soc. 1989. V.314. N° 2. P. 499–526.

7. *Herzel S., Recchioni M.C., Zirilli F.* A quadratically convergent method for linear programming // Linear Algebra and its Appl. 1991. V.152. P. 255–289.

8. *Зоркальцев В.И.* Проективные алгоритмы оптимизации, использующие множители предыдущих итераций // ЖВМ и МФ. 1994. Т.34. N° 7. С. 1095–1103.

9. *Adler I., Karmarkar N., Resende M.G.C., Velga G.* An implementation of Karmarkar's algorithm for linear programming // Math. Program. 1989. V.44. P. 297–335.

10. *Евтушенко Ю.Г.. Жадан В.Г.* Численные методы решения некоторых задач исследования операций // ЖВМ и МФ. 1973. Т.13. N° 3. С. 583–597.

11. *Евтушенко Ю.Г., Жадан В.Г.* Релаксационный метод решения задач нелинейного программирования // ЖВМ и МФ. 1977. Т.17. N° 4. С. 890–904.

12. *Евтушенко Ю.Г.* Методы решения экстремальных задач и их применение в системах оптимизации. М.: Наука, 1982.

13. *Evtushenko Yu.G., Zhadan V.G.* Stable barrier-projection and barrier-Newton methods in nonlinear programming // Optimizat. Methods and Software. 1994. V.3. N° 1-3. P. 237–256.

14. *Evtushenko Yu.G., Zhadan V.G.* Stable barrier-projection and barrier-Newton methods in linear programming // Comput. Optimizat. and Appl. 1994. V.3. P. 289–303.

15. *Евтушенко Ю.Г., Жадан В.Г.* Барьерно-проективные методы решения задач нелинейного программирования // ЖВМ и МФ. 1994. Т.34. N° 5. С. 669–684.

16. *Tanabe K.* A geometric method in nonlinear programming // J. Optimizat. Theory and Appl. 1980. V.30. N° 2. P. 181–210.