Ю. Г. Евтушенко, В. Г. Жадан, Двойственные барьерно-проективные и барьерно-ньютоновские методы для задач линейного программирования, *Ж. вычисл. матем. и матем. физ.*, 1996, том 36, номер 7, 30–45

ЖУРНАЛ
ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ ФИЗИКИ

Том 36, 1996 № 7

УДК 519.6:519.852

© 1996 г. Ю.Г. ЕВТУШЕНКО, В.Г. ЖАДАН

(Москва)

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

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 двойствен-

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

ных барьерно-проективных и барьерно-ньютоновских методов. Сюръективные преобразования применяются в данной работе для освобождения от требования неотрицательности дополнительных двойственных переменных.

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

$$\min_{x\in X} c^{\mathrm{T}}x, \quad X=\{x\in\mathbb{R}^n:\ Ax=b, x\geqslant 0_n\}. \tag{1.1}$$

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

$$\max_{u\in U} b^{\mathrm{T}}u, \quad U=\{u\in\mathbb{R}^m:\ v=c-A^{\mathrm{T}}u\geqslant 0_n\}. \tag{1.2}$$

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

$$\varphi(w)=[\varphi^1(w^1),\dots,\varphi^n(w^n)]^{\mathrm{T}}.$$

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

$$\theta(v)=[\theta^1(v^1),\dots,\theta^n(v^n)]^{\mathrm{T}}, \quad G(v)=D(\theta(v)),$$

где

$$\theta^i(v^i)=(\gamma^i(v^i))^2, \ \gamma^i(v^i)=\dot\varphi^i(\psi^i(v^i)), \ 1\leqslant i\leqslant n,$$

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

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

У с л о в и е 1. Функции $\theta^i(v^i)$, $1 \leqslant i \leqslant n$, определены и непрерывны в некоторой окрестности $\mathbb{R}^1_+$ и $\theta^i(v^i) = 0$ в том и только том случае, когда $v^i = 0$.

У с л о в и е 2. Функции $\theta^i(v^i)$, $1 \leqslant i \leqslant n$, 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 $G(v)$:

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

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

Здесь символ e^{-w} обозначает вектор-функцию с компонентами e^{-w^i}, $1 \leqslant i \leqslant n$. 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:

$$(1.5а)\qquad \max b^{\mathrm{T}}u,$$

$$(1.5б)\qquad \varphi(w) - c + A^{\mathrm{T}}u = 0_n.$$

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

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

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

$$(1.6)\qquad du/dt = \tilde{L}_u(u,w,x(u,w)),\quad dw/dt = \tilde{L}_w(u,w,x(u,w)),$$

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

$$(1.7)\qquad \tilde{L}_{xu}(u,w,x)\dot{u} + \tilde{L}_{xw}(u,w,x)\dot{w} = -\tau\tilde{L}_x(u,w,x),\quad \tau > 0.$$

Так как $\dot{v} = \varphi_w\dot{\omega}$, то в пространстве векторов $z = [u, v] \in \mathbb{R}^{m+n}$ метод (1.6), (1.7) принимает вид

$$(1.8а)\qquad du/dt = b - Ax(z),\quad dv/dt = -G(v)x(z),$$

$$(1.8б)\qquad \Phi(v)x(z) = A^{\mathrm{T}}b + \tau(v + A^{\mathrm{T}}u - c),$$

где $\Phi(v) = G(v) + A^{\mathrm{T}}A$. Обозначим $v(u) = c - A^{\mathrm{T}}u$,

Л е м м а 1. *Пусть преобразование* $\varphi(w)$ *удовлетворяет условию* 1. *Тогда* $\Phi(v(u))$ – *неособая матрица для любого* $u \in U_0$.

Д о к а з а т е л ь с т в о. В силу условия 1, матрица $G(v(u))$ положительно определена на U_0. Матрица $A^{\mathrm{T}}A$ является матрицей Грама и, следовательно, неотрицательно определена. Поэтому вся матрица $\Phi(v(u))$ положительно определена.

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

$$(1.9)\qquad u = \sum_{j=1}^{s}\alpha_j u_j,\quad \alpha_j > 0,\quad 1 \leqslant j \leqslant s,\quad \sum_{j=1}^{s}\alpha_j = 1,$$

где u_j, $1 \leqslant j \leqslant s$, – 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

$$\Phi(v)\bar{x} = G(v)\bar{x} + A^{\mathrm{T}}A\bar{x} = 0_n \tag{1.10}$$

имеет место в том и только том случае, когда $\bar{x} = 0_n$.

Действительно, умножая (1.10) слева на $\bar{x}^{\mathrm{T}}$, получаем

$$\bar{x}^{\mathrm{T}}G(v)\bar{x} + \bar{x}^{\mathrm{T}}A^{\mathrm{T}}A\bar{x} = 0. \tag{1.11}$$

Так как оба члена в (1.11) неотрицательны, то должно выполняться

$$\bar{x}^{\mathrm{T}}G(v)\bar{x} = 0, \quad \bar{x}^{\mathrm{T}}A^{\mathrm{T}}A\bar{x} = 0. \tag{1.12}$$

Обозначим

$$S_j = \{1 \leqslant i \leqslant n\colon\ \alpha_i^{\mathrm{T}}u_j = c^i\}, \quad S = \bigcap_{j=1}^{s} S_j.$$

Если $S = 0$, то $v > 0_n$, и из первого равенства (1.12) получаем, что $\bar{x} = 0_n$. Рассмотрим теперь случай, когда $S \neq 0$. Не ограничивая общности, можно считать, что $S = \{1, 2, \ldots, k\}$. 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 $k \leqslant m$ и матрица B имеет полный ранг. Кроме того, согласно (1.9), выполняется $v_B = 0_k$, $v_N > 0_{n-k}$. Из первого равенства (1.12) тогда следует, что $\bar{x}_N = 0_{n-k}$. Поэтому второе равенство (1.12) сводится к $\bar{x}_B^{\mathrm{T}}B^{\mathrm{T}}B\bar{x}_B = 0$. 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.

Следствие 1. Если угловая точка u 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

$$W = \{v \in \mathbb{R}^n\colon\ v = v(u),\ u \in \mathbb{R}^m\}, \quad V = \{v \in \mathbb{R}^n\colon\ v = v(u),\ u \in U\}. \tag{1.13}$$

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

$$x(u, v) = [G(v) + A^{\mathrm{T}}A]^{-1}[A^{\mathrm{T}}b + \tau(v + A^{\mathrm{T}}u - c)]. \tag{1.14}$$

Подставляя (1.14) в (1.8а), приходим к другой форме записи метода (1.8):

$$du/dt = b - A[G(v) + A^{\mathrm{T}}A]^{-1}[A^{\mathrm{T}}b + \tau(v + A^{\mathrm{T}}u - c)],$$

$$dv/dt = -G(v)[G(v) + A^{\mathrm{T}}A]^{-1}[A^{\mathrm{T}}b + \tau(v + A^{\mathrm{T}}u - c)].$$

Пусть $[u(t, z_0), v(t, z_0)]$ – решение системы (1.8а), удовлетворяющее начальному условию $u(t, z_0) = u_0$, $v(t, z_0) = v_0$, $z_0^{\mathrm{T}} = [u_0^{\mathrm{T}}, v_0^{\mathrm{T}}]$. Обозначим $y(u, v) = c - A^{\mathrm{T}}u - v$. Условие (1.8б) можно записать в виде

$$dy(u, v)/dt = y_u^{\mathrm{T}}(u, v)\dot{u} + y_v^{\mathrm{T}}(u, v)\dot{v} = -\tau y.$$

Отсюда следует, что система (1.8а) имеет первый интеграл

(1.15) $$c - A^{\mathrm{T}}u(t, z_0) - v(t, z_0) = (c - A^{\mathrm{T}}u_0 - v_0)e^{-\tau t}.$$

Таким образом, $c - A^{\mathrm{T}}u(t, z_0) - v(t, z_0) \to 0_n$ при $t \to +\infty$. Кроме того, вдоль траекторий системы, согласно (1.8б), выполняется

(1.16) $$b^{\mathrm{T}}\frac{du}{dt} = b^{\mathrm{T}}(b - Ax(z)) = \|b - Ax(z)\|^2 + x^{\mathrm{T}}(z)A^{\mathrm{T}}(b - Ax(z)) =$$

$$= \|b - Ax(z)\|^2 + x^{\mathrm{T}}(z)G(v)x(z) + \tau x^{\mathrm{T}}(z)(c - A^{\mathrm{T}}u - v).$$

Из второго уравнения (1.8а) 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 (1.8а). Вместо (1.8) приходим к

(1.17а) $$du/dt = b - Ax(u),$$

(1.17б) $$[G(v(u)) + A^{\mathrm{T}}A]x(u) = A^{\mathrm{T}}b,$$

где $u(0, u_0) = u_0 \in U$. Для данной системы вместо (1.16) справедлива формула

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

т.е. целевая функция двойственной задачи (1.2) монотонно возрастает на допустимом множестве. Метод (1.17) был впервые предложен в 1977 г. в [11].

Применяя метод Эйлера для интегрирования системы (1.8), получаем

(1.18а) $$u_{k+1} = u_k + \alpha_k(b - Ax_k), \quad v_{k+1} = v_k - \alpha_k G(v_k)x_k,$$

(1.18б) $$[G(v_k) + A^{\mathrm{T}}A]x_k = A^{\mathrm{T}}b + \tau(v_k + A^{\mathrm{T}}u_k - c).$$

Соответственно, для системы (1.17) имеем

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

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

Т е о р е м а 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 $z_*^{\mathrm{T}} = [u_*^{\mathrm{T}}, v_*^{\mathrm{T}}]$ является асимптотически устойчивым положением равновесия для системы* (1.8);

б) *решения $u(t, z_0)$, $v(t, z_0)$ системы* (1.8) *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* (1.18), *локально сходится к z_* с линейной скоростью; соответствующая последовательность $\{x_k\}$ сходится к x_**;

г) *решения $u(t, u_0)$ системы* (1.17) *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* (1.19), *локально сходится к* u_* на U *с линейной скоростью; соответствующая последовательность* $\{x_k\}$ *сходится к* x_*.

Доказательство. Составим уравнение в вариациях для системы (1.8):

$$\delta\dot{u} = -A\left[\frac{\partial x(z_*)}{\partial u}\delta u + \frac{\partial x(z_*)}{\partial \upsilon}\delta\upsilon\right],$$

$$\delta\dot{\upsilon} = \pm\left\{G(\upsilon_*)\frac{\partial x(z_*)}{\partial u}\delta u + \left[D(\dot{\theta}(\upsilon_*))D(x_*) + G(\upsilon_*)\frac{\partial x(z_*)}{\partial \upsilon}\right]\delta\upsilon\right\},$$

или в матричном виде, введя обозначение $\delta z^{\mathrm{T}} = [\delta u^{\mathrm{T}}, \delta\upsilon^{\mathrm{T}}]$:

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

Здесь

$$Q(z_*) = \begin{bmatrix} \tau A\Phi_*^{-1}A^{\mathrm{T}} & A\Phi_*^{-1}[\tau I_n - D(\dot{\theta}(\upsilon_*))D(x_*)] \\ \tau G(\upsilon_*)\Phi_*^{-1}A^{\mathrm{T}} & [I_n - G(\upsilon_*)\Phi_*^{-1}]D(\dot{\theta}(\upsilon_*))D(x_*) + \tau G(\upsilon_*)\Phi_*^{-1} \end{bmatrix}, \tag{1.20}$$

где $\Phi_* = G(\upsilon_*) + A^{\mathrm{T}}A$, I_n – единичная матрица порядка n. При вычислении матрицы (1.20) были использованы соотношения, следующие из (1.8б):

$$[G(\upsilon) + A^{\mathrm{T}}A]\frac{\partial x(z)}{\partial u} = \tau A^{\mathrm{T}},$$

$$D(\dot{\theta}(\upsilon))D(x) + [G(\upsilon) + A^{\mathrm{T}}A]\frac{\partial x(z)}{\partial \upsilon} = \tau I_n.$$

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

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

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

Так как при сделанном предположении

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

то на основании формулы Фробениуса получаем

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

Отсюда вытекает, что

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

С помощью этого соотношения матрицу Q можно привести к виду

$$Q = \begin{bmatrix} \tau I_m & Q_2 \\ 0_{nm} & Q_1 \end{bmatrix}, \quad Q_1 = \begin{bmatrix} D(\dot\theta(v_*^B))D(x_*^B) & 0_{md} \\ Q_3 & \tau I_d \end{bmatrix}, \tag{1.21}$$

где вид матриц Q_2 и Q_3 несуществен.

Из (1.21) следует, что матрица Q имеет n собственных значений, равных τ, и m собственных значений $\dot\theta^i(0)x_*^i$, $1 \leqslant i \leqslant m$. 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 (1.8) локально экспоненциально сходятся к z_*.

Сходимость дискретного варианта (1.18) для достаточно малых постоянных α_k следует из теоремы 2.3.7, приведенной в [12].

Так как решения системы (1.17) при $u_0 \in U$ совпадают с соответствующими решениями более общей системы (1.8), если в последней положить $v_0 = v(u_0)$, то решения (1.17) локально экспоненциально сходятся к u_* на U. По той же причине последовательность $\{u_k\}$, генерируемая процессом (1.19), локально сходится к u_* на U. Теорема доказана.

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

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

Проводя стандартные рассуждения (см. [12]), можно показать, что фигурирующая в утверждении в) величина $\alpha_* = 2/\eta^*$. При этом скорость сходимости будет наибольшей, если в (1.18) шаги взять равными $\alpha_k = 2/(\eta_* + \eta^*)$. Тогда для выполнения условия $\| z_k - z_* \|) \leqslant \varepsilon$ достаточно сделать $\ln(\varepsilon/\| z_* - z_0 \|)/\ln q$ итераций, где $q = (\eta^* - \eta_*)/(\eta^* + \eta_*)$.

Для того чтобы определить правую часть в системе (1.17а), 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:

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

Отсюда приходим к системе

$$\frac{du}{dt} = [I_m + AG^{-1}(v(u))A^{\mathrm T}]^{-1}b, \quad u(t,0) = u_0 \in U_0. \tag{1.22}$$

Локальная сходимость метода (1.22) 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 (1.8б) в методе (1.8) можно было бы заменить на любое другое, обеспечивающее стремление компонент вектора $y(u, v)$ к нулю. Например, вместо (1.8б) можно было взять

$$[G(v) + A^{\mathrm T}A]x(z) = A^{\mathrm T}b + \tau D(v + A^{\mathrm T}u - c)(v + A^{\mathrm T}u - c), \quad \tau > 0.$$

Тогда система (1.8) вместо (1.15) имела бы первый интеграл

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

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 $AP^{\mathrm{T}} = 0_{md}$. 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

$$P = \left[\begin{array}{c|c} -N(B^{\mathrm{T}})^{-1} & I_d \end{array}\right].$$

Определения (1.13) множеств W и V с помощью матрицы P можно переписать в виде

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

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

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

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

(2.1) $$\min_{v \in V} \bar{x}^{\mathrm{T}}v.$$

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

(2.2а) $$dv/dt = -G(v)[\bar{x} - P^{\mathrm{T}}x(v)],$$

(2.2б) $$PG(v)P^{\mathrm{T}}x(v) = PG(v)\bar{x} + \tau P(c - v).$$

В тех точках $v \in \mathbb{R}^n$, где матрица $PG(v)P^{\mathrm{T}}$ неособая, разрешая уравнение (2.2б), получаем

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

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

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

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

$$Pdv/dt = \tau P(c - v), \quad P(c - v(t, v_0)) = P(c - v_0)e^{-\tau t},$$

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

Если $v_0 \in V_0$, где $V_0 = \{v \in V\colon\ v > 0_n\}$, то (2.3) обладает свойствами метода внутренней точки; целевая функция $\bar{x}^{\mathrm{T}} v(v_0, t)$ монотонно убывает и $v(t, v_0) \in V_0$ для всех $t \geqslant 0$. В этом случае метод (2.3) может быть переписан в виде

$$dv/dt = -G(v)[I_n - P^{\mathrm{T}}(PG(v)P^{\mathrm{T}})^{-1}PG(v)]\bar{x}, \quad v_0 \in V_0. \tag{2.4}$$

Т е о р е м а 2. *Пусть выполнены предположения теоремы* 1. *Тогда верно следующее*:

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

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

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

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

локально сходится к v_* *с линейной скоростью.*

Д о к а з а т е л ь с т в о этой теоремы аналогично доказательству теоремы 1, только теперь у матрицы Q будет d собственных значений, равных τ, и m собственных значений, равных $\dot{\theta}^i(0)\overset{i}{x}_*$, $1 \leqslant i \leqslant m$. Выбор вектора $\bar{x}$ в (2.1), таким образом, не влияет на скорость сходимости.

Так как для системы (2.4) 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 (2.4) имеет место представление

$$\dot{v} = A^{\mathrm{T}}\lambda \tag{2.5}$$

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 (2.4), получаем

$$\lambda = -[AG^{-1}(v)A^{\mathrm{T}}]^{-1}A\bar{x} = -[AG^{-1}(v)A^{\mathrm{T}}]^{-1}b. \tag{2.6}$$

Подставляя (2.6) в (2.5), приходим к другой форме записи метода (2.4):

$$dv/dt = -A^{\mathrm{T}}[AG^{-1}(v)A^{\mathrm{T}}]^{-1}b, \quad v_0 \in V_0. \tag{2.7}$$

В пространстве переменных u метод (2.7) принимает вид

$$du/dt = [AG^{-1}(v(u))A^{\mathrm{T}}]^{-1}b, \quad u_0 \in U_0. \tag{2.8}$$

Если воспользоваться преобразованиями (1.3) и (1.4), то из (2.8) получаем, соответственно,

$$du/dt = [AD^{-1}(v(u))A^{\mathrm{T}}]^{-1}b, \quad u_0 \in U_0, \tag{2.9}$$

$$du/dt = [AD^{-2}(v(u))A^{\mathrm{T}}]^{-1}b, \quad u_0 \in U_0. \tag{2.10}$$

Формула (2.10) совпадает с непрерывным вариантом двойственного а.м.м., предложенного в [9].

Как следует из утверждения теоремы 2, решения системы (2.2) локально экспоненциально сходятся к $v_* = v(u_*)$ на множестве V. Поэтому решения системы (2.9) также локально экспоненциально сходятся к u_* на множестве U_0.

Если рассмотреть дискретный аналог метода (2.9)

$$u_{k+1} = u_k + \alpha_k[AD^{-1}(v_k)A^{\mathrm{T}}]^{-1}b, \quad u_0 \in U_0, \tag{2.11}$$

где $v_k = v(u_k)$, то, согласно [12], из экспоненциальной сходимости непрерывного варианта (2.9) следует локальная сходимость итеративного процесса (2.11) 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:

(2.12а) $\max b^{\mathrm{T}}u$,

(2.12б) $c - A^{\mathrm{T}}u - v = 0_n$,

(2.12в) $v - \varphi(w) = 0_n$.

Обозначим $\tilde{z}^{\mathrm{T}} = [u^{\mathrm{T}}, v^{\mathrm{T}}, w^{\mathrm{T}}]$ и составим для (2.12) функцию Лагранжа

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

Аналогом метода (1.6) для задачи (2.12) является

(2.13) $\dot{u} = \tilde{L}_u = b - Ax(\tilde{z}),\quad \dot{v} = \tilde{L}_v = y(\tilde{z}) - x(\tilde{z}),\quad \dot{w} = \tilde{L}_w = -\varphi_w^{\mathrm{T}}(w)y(\tilde{z}),$

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

(2.14) $\tilde{L}_{xu}\dot{u} + \tilde{L}_{xv}\dot{v} + \tilde{L}_{xw}\dot{w} = -\tau\tilde{L}_x,\quad \tilde{L}_{yu}\dot{u} + \tilde{L}_{yv}\dot{v} + \tilde{L}_{yw}\dot{w} = -\tau\tilde{L}_y.$

В переменных u, v и $p = \varphi(w)$ система (2.13), (2.14) может быть переписана в виде

(2.15а) $du/dt = b - Ax(z),\ dv/dt = y(z) - x(z),\ dp/dt = -G(p)y(z),$

(2.15б) $(I_n + A^{\mathrm{T}}A)x(z) - y(z) = A^{\mathrm{T}}b - \tau(c - A^{\mathrm{T}}u - v),$

(2.15в) $(I_n + G(p))y(z) - x(z) = \tau(p - v),$

где $z^{\mathrm{T}} = [u^{\mathrm{T}}, v^{\mathrm{T}}, p^{\mathrm{T}}]$. 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 (2.15в) следует, что $y(z) = [I_n + G(p)]^{-1}x(z)$. Поэтому в этом случае метод (2.15) упрощается: вместо (2.15) имеем

(2.16а) $du/dt = b - Ax(u, v),\ dv/dt = -G(v)[I_n + G(v)]^{-1}x(u, v),$

(2.16б) $\{G(v)(1 + G(v))^{-1} + A^{\mathrm{T}}A\}x(u, v) = A^{\mathrm{T}}b - \tau(c - A^{\mathrm{T}}u - v).$

Кроме того, если дополнительно предположить, что $v_0 = v(u_0)$, то $v(t, z_0) \equiv v(u(t, z_0))$ и метод (2.16) с помощью формулы Шермана–Моррисона–Вудбери на множестве U_0 принимает вид

(2.17) $du/dt = \{I_n + A[I_n + G^{-1}(v(u))]A^{\mathrm{T}}\}^{-1}b,\quad u_0 \in U_0.$

Локальная сходимость метода (2.17) и его дискретного варианта

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

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 (2.15).

§ 3. Глобальная сходимость методов

Рассмотрим вопрос о глобальной сходимости метода (2.11) на множестве U_0.

Предположим, что задача (1.1) такова, что

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

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

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

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),

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

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

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

$$\geqslant -C \ln \sum_{i \in J_*^N} \frac{v^i(u)}{C} = 0,$$

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

Вычислим производную функции (3.2) в силу системы (2.9). Имеем

$$(3.4) \qquad dF(u)/dt = F_u^{\mathrm{T}} \dot{u} = v_*^{\mathrm{T}} D^{-1}(v(u)) A^{\mathrm{T}} [A D^{-1}(v(u)) A^{\mathrm{T}}]^{-1} b.$$

Обозначим

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

Так определенный вектор $x(u)$ удовлетворяет равенству $Ax(u) = b$. Кроме того, согласно (3.1), $x^{\mathrm{T}}(u) v(u) = \bar{e}^{\mathrm{T}} A^{\mathrm{T}} p(u) = 0$. Поэтому

$$x^{\mathrm{T}}(u) c = x^{\mathrm{T}}(u)(v(u) + A^{\mathrm{T}} u) = u^{\mathrm{T}} A x(u) = b^{\mathrm{T}} u.$$

Отсюда и из (3.4) получаем

$$(3.5) \qquad dF(u)/dt = v_*^{\mathrm{T}} x(u) = x^{\mathrm{T}}(u)(c - A^{\mathrm{T}} u_*) = b^{\mathrm{T}} u - b^{\mathrm{T}} u_* \leqslant 0,$$

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

Для произвольного $u_0 \in U_0$ обозначим $Q = \{u \in U_1: F(u) \leqslant F(u_0)\}$. 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 $t \geqslant 0$.

Положим

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

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

$$F(u(t, u_0)) \leqslant F(u_0) e^{-Kt}, \quad t \geqslant 0.$$

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

$$(3.7)\qquad K \geqslant \frac{1-\exp[-F(u_0)/C]}{F(u_0)} \min_{1\leqslant j\leqslant m} s_j > 0,$$

где $s_j = b^{\mathrm{T}}(u_* - u_j)$, u_j – *смежная с* u_* *вершина многогранника* U, $1 \leqslant j \leqslant m$.

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

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

где $Q_1 = \{z \in Z: \tilde{F}(z) \leqslant F(u_0)\}$, $Z = \{z \in R^m:\ A^{\mathrm{T}} z \leqslant \upsilon_*\}$. 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 $z \in Z$, $z \neq 0_m$. Поэтому для любой точки $\bar{z} \in S = \{z \in Q_1 : \tilde{F}(z) = F(u_0)\}$ и любого $0 < \alpha \leqslant 1$ выполняется $\tilde{F}(\alpha\bar{z}) \leqslant \alpha\tilde{F}(\bar{z})$. Отсюда следует, что

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

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

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

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

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

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

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

и, следовательно, любое решение β_j уравнения (3.9), $1 \leqslant j \leqslant m$, удовлетворяет неравенству $\beta_j \geqslant \bar{\beta}$, где $\bar{\beta}$ – решение уравнения

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

Отсюда

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

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

Обозначим

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

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

$$(3.12)\qquad A^{\mathrm{T}}\{u + \alpha[AD^{-1}(\upsilon)A^{\mathrm{T}}]^{-1} b\} \leqslant c.$$

Таким образом, вектор $u + \alpha[AD^{-1}(v)A^{\mathrm{T}}]^{-1}b$ 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* (2.11) *выбирается из условия*

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

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

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

где величина K *определяется согласно* (3.6).

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

$$(3.15)\qquad F(u_{k+1}) = -\sum_{i\in J_*^N} v_*^i \ln\left[1 - \frac{\alpha_i^{\mathrm{T}}(u_k + \alpha_k p_k - u_*)}{v_i^*}\right] =$$

$$= -\sum_{i\in J_*^N} v_*^i \ln\left[\frac{v_k^i}{v_*^i}(1 - \alpha_k x_k)\right] = F(u_k) - \sum_{i\in J_*^N} v_*^i \ln(1 - \alpha_k x_k^i).$$

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

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

Разлагая в ряд правую часть (3.16), получаем

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

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

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

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

Обозначим

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

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 $v_*^{\mathrm{T}} = [v_*^B, v_*^N]$, 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 $\Gamma^B(u) = BD^{-1}(v^B(u))B^{\mathrm{T}}$, $\Gamma^N(u) = ND^{-1}(v^{\mathrm{T}}(u))N^{\mathrm{T}}$. Так как B – невырожденная матрица, то для всех $u \in U_0$

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

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

$$x^B(u) = D^{-1}(v^B)B^{\mathrm{T}}(\Gamma^B(u))^{-1}b + D^{-1}(v^B)B^{\mathrm{T}}\Phi(u)b = x_*^B + \varphi_1(u),$$

$$x^N(u) = D^{-1}(v^N)N^{\mathrm{T}}\Gamma^{-1}(u)b = \varphi_2(u),$$

$$\mu(u) = \max_{1 \leqslant i \leqslant m} x_*^i + \varphi_3(u),$$

$$\|\varphi_i(u)\| = O(\|u - u_*\|), \quad i = 1,2,3.$$

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

$$\sum_{i \in J_*^N} v_*^i (x^i(u))^2 = o(\|u - u_*\|),$$

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

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

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

Обозначим

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

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

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

где $0 < \alpha \leqslant \bar{\alpha}(u_0)$. С помощью (3.18) можно оценить число шагов, необходимых для попадания процесса (2.11) в некоторую окрестность точки u_*. Неравенство (3.18) сохраняется и для процесса (2.11) с постоянным шагом $\alpha_k = \alpha \leqslant \bar{\alpha}(x_0)$.

§ 4. Двойственный барьерно-ньютоновский метод

Если подставить выражение из (1.17б) 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 (1.17б) по u, получаем

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

Поэтому

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

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

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

Л е м м а 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.*

Д о к а з а т е л ь с т в о. При $v_* = v_*(u_*)$ выполняется $(D(\theta(v_*)) + A^{\mathrm{T}}A)x_* = A^{\mathrm{T}}b$ 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

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

получаем

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

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

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

где $v_k = v(u_k)$, $x_k = x(u_k)$.

Т е о р е м а 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):

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

Его дискретный вариант аналогичен (4.5).

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

1. *Karmarkar N.* A new polinomial 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. № 2. P. 167–224.
3. *Дикин И.И.* Итеративное решение задач линейного и квадратичного программирования // Докл. АН СССР. 1967. Т. 174. № 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. № 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. № 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 // Ж. вычисл. матем. и матем. физ. 1994. Т. 34. № 7. С. 1095–1103.
9. *Adler I., Karmarkar N., Resende M.G.C., Veiga 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 // Ж. вычисл. матем. и матем. физ. 1973. Т. 13. № 3. С. 583–597.
11. *Евтушенко Ю.Г., Жадан В.Г.* Релаксационный метод решения задач нелинейного программирования // Ж. вычисл. матем. и матем. физ. 1977. Т. 17. № 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. № 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 // Ж. вычисл. матем. и матем. физ. 1994. Т. 34. № 5. С. 669–684.
16. *Tanabe K.* A Geometric method in nonlinear programming // J. Optimizat. Theory and Appl. 1980. V. 30. № 2. P. 181–210.

Поступила в редакцию 14.04.95