Ю. Г. Евтушенко, В. Г. Жадан, А. П. Черенков, Применение метода Ньютона к решению задач линейного программирования, *Ж. вычисл. матем. и матем. физ.*, 1995, том 35, номер 6, 850–866

ЖУРНАЛ
ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ ФИЗИКИ

Том 35, 1995 № 6

УК 519.852

© 1995 г. Ю. Г. ЕВТУШЕНКО, В. Г. ЖАДАН, А. П. ЧЕРЕНКОВ

(*Москва*)

ПРИМЕНЕНИЕ МЕТОДА НЬЮТОНА К РЕШЕНИЮ ЗАДАЧ ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ [1)]

Рассматриваются непрерывные и дискретные варианты барьерно-ньютоновского метода решения задач линейного программирования. Метод является прямо-двойственным, и в его основе лежит идея отыскания с помощью метода Ньютона тех точек в прямом и двойственном пространстве, которые удовлетворяют совместной системе условий оптимальности. Исследуются локальные и нелокальные свойства метода. Для дискретных вариантов метода предлагается использовать разные шаги в прямом и двойственном пространствах. Показано, что при специальных регулировках шагов метод сходится со сверхлинейной и квадратичной скоростью. Рассматривается вариант метода, в котором шаги выбираются из условия наискорейшего спуска, и выделена область начальных условий, при которых метод находит решение не более чем за две итерации.

Введение

Метод Ньютона является одним из наиболее эффективных способов решения систем нелинейных уравнений и задач оптимизации [1]—[3]. В последнее время появились многочисленные варианты метода, предназначенные для решения задач линейного программирования (л.п.) (см., например, [4]—[6]). Подробный обзор таких методов содержится в [7], причем особый интерес среди них представляют прямо-двойственные алгоритмы, в которых методом Ньютона решается параметризованная система уравнений, дающих в пределе условия оптимальности для прямой и двойственной задач [8]—[12]. Эти методы позволяют сочетать достаточно высокую локальную скорость сходимости с полиномиальностью алгоритмов. В настоящей работе также предлагается прямо-двойственный метод, основанный на решении системы уравнений, задающих условия оптимальности в задаче л.п., а именно условия дополняющей нежесткости и условия допустимости (см. [13]).

Если воспользоваться преобразованием пространств для освобождения от требования неотрицательности переменных [14], [15] и применить метод Ньютона для отыскания точек, удовлетворяющих условиям Куна — Таккера, то можно получить целое семейство различных методов. Такие методы для общей задачи нелинейного программирования рассматривались в [15], [16] и для задачи л.п. — в [17].

Данная статья посвящена исследованию специального класса методов, соответствующих таким покомпонентным преобразованиям пространств, которые приводят к умножению правых частей систем обыкновенных дифференциальных уравнений, описывающих метод, на диагональные матрицы, играющие роль

[1)] Работа выполнена при поддержке Российского фонда фундаментальных исследований (коды проектов 93-012-450, 94-01-01379).

барьеров и не позволяющие траекториям пересекать границы положительных ортантов как в исходном пространстве, так и в пространстве двойственных дополнительных переменных. В отличие от [17], здесь основное внимание уделяется выбору шагов в прямом и двойственном пространствах, которые, как и в [18], могут быть различными. Рассматриваются отдельно варианты метода с малыми шагами, а также с шагами, близкими к единице или выбираемыми из решения вспомогательных оптимизационных задач.

§ 1. Постановка задачи, основные идеи метода

Пусть $x = [x^1, \dots, x^n]$, $u = [u^1, \dots, u^m]$ — векторы из евклидовых пространств $\mathbb{R}^n$ и $\mathbb{R}^m$ соответственно. Рассмотрим прямую и двойственную задачи л.п., заданные в следующей форме:

$$\min_{x \in X} c^{\mathrm{T}} x, \qquad X = \{x \in \mathbb{R}^n : b - Ax = 0_m,\ x \geq 0_n\}, \tag{1.1}$$

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

Здесь и ниже A — матрица $m \times n$, в которой $m < n$, 0_i — нулевой i-мерный вектор, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$. Символ 0_{ij} будет обозначать нулевую матрицу $i \times j$.

Введем в рассмотрение следующие множества:

$$\mathbb{R}^n_+ = \{x \in \mathbb{R}^n : x \geq 0_n\}, \qquad \mathbb{R}^n_{++} = \{x \in \mathbb{R}^n : x > 0_n\},$$

$$\operatorname{int} U = \{u \in \mathbb{R}^m : v = c - A^{\mathrm{T}} u > 0_n\}, \qquad \operatorname{ri} X = \{x \in \mathbb{R}^n_{++} : Ax = b\}.$$

Всюду предполагаем, что ранг матрицы A равен m, множества $\operatorname{int} U$ и $\operatorname{ri} X$ непусты, задача (1.1) имеет единственное решение x_*, которое не вырождено. Тогда двойственная задача (1.2) также имеет единственное решение u_* и оно не вырождено, вектор x_* имеет m ненулевых компонент, вектор $v_* = c - A^{\mathrm{T}} u_*$ имеет m нулевых компонент. При этом выполняются условие дополняющей нежесткости $x^j_* v^j_* = 0$, $1 \leq j \leq n$, и условие строгой дополняющей нежесткости, т. е. из $x^j_* = 0$ следует, что $v^j_* > 0$.

Для решения задач линейного и нелинейного программирования в [16], [17] был использован метод Ньютона. В результате было получено целое семейство численных методов. Здесь ограничимся рассмотрением только одной численной схемы, в которой итерации строятся по формуле

$$W(x_k, u_k, \lambda_k) \begin{bmatrix} x_{k+1} - x_k \\ u_{k+1} - u_k \end{bmatrix} = - \begin{bmatrix} \alpha_k D(x_k) v_k \\ \tau_k (Ax_k - b) \end{bmatrix}. \tag{1.3}$$

Здесь нижний индекс k обозначает номер итерации, $D(z)$ — диагональную матрицу, у которой на главной диагонали находится вектор z, α_k, τ_k, λ_k — некоторые положительные коэффициенты, $\lambda_k = \alpha_k / \tau_k$, n-мерный вектор v имеет вид $v = v(u) = c - A^{\mathrm{T}} u$, W — квадратная матрица порядка $n + m$,

$$W(x, u, \lambda) = \begin{bmatrix} \lambda D(v) & -D(x) A^{\mathrm{T}} \\ A & 0_{mm} \end{bmatrix}.$$

В тех точках, где векторы x_k и v_k имеют все компоненты, не равные нулю, система (1.3) может быть записана в виде

$$(1.4) \qquad x_{k+1} = D(x_k)\,[e_n + \tau_k(\eta_k - e_n)], \qquad u_{k+1} = u_k + \alpha_k\mu_k,$$

где e_n есть n-мерный вектор, все компоненты которого равны единице, векторы $\eta \in \mathbb{R}^n$, $\mu \in \mathbb{R}^m$ определяются по формулам

$$(1.5) \qquad \eta(x, u) = D^{-1}(v)\,A^{\mathrm{T}}\mu(x, u), \qquad \mu(x, u) = M^{-1}(x, u)\,b.$$

Здесь $M(x, u) = AD(x)\,D^{-1}(v)\,A^{\mathrm{T}}$ — матрица Грама.

Используя связь $v = c - A^{\mathrm{T}}u$ между векторами u и v, метод (1.4) можно переписать в переменных x, v:

$$(1.6) \qquad x_{k+1} = D(x_k)\,[e_n + \tau_k(\eta_k - e_n)], \qquad v_{k+1} = D(v_k)\,(e_n - \alpha_k\eta_k).$$

Введем матрицу $\Lambda \in \mathbb{R}^{d\times n}$, где $d = n - m$ — дефект матрицы A. Столбцы матрицы Λ^{T} образуют базис нуль-пространства матрицы A, т. е. $A\Lambda^{\mathrm{T}} = 0_{md}$. Введем в рассмотрение множество $V = \{v \in \mathbb{R}^n : \Lambda(v - c) = 0_d\}$. Если $v \in V$, то вектор $v - c$ лежит в пространстве строк матрицы A и для вектора v существует такой m-мерный вектор u, что $v = c - A^{\mathrm{T}}u$. Из второго соотношения (1.6) следует $\Lambda(v_{k+1} - c) = \Lambda(v_k - c) - \alpha_k\Lambda A^{\mathrm{T}}\mu_k = \Lambda(v_k - c)$. Таким образом, если $v_0 \in V$, то $v_k \in V$ для всех k и соотношения (1.4) эквивалентны (1.6). Итерации можно проводить как в пространстве x, u, так и в пространстве x, v.

Лемма 1. *Пусть x_* и u_* — невырожденные решения задач* (1.1) *и* (1.2), *тогда матрица $W(x_*, u_*, \lambda)$ не вырождена.*

Доказательство. Не нарушая общности, считаем, что ненулевыми являются первые m компонент вектора x_*. Тогда векторы x_*, v_* и матрицы A и W можно представить в виде

$$(1.7\text{а}) \qquad x_* = \begin{bmatrix} x_*^B \\ x_*^N \end{bmatrix}, \quad v_* = \begin{bmatrix} v_*^B \\ v_*^N \end{bmatrix}, \quad x_*^B > 0_m, \quad x_*^N = 0_d, \quad v_*^B = 0_m, \quad v_*^N > 0_d,$$

$$(1.7\text{б}) \qquad A = [B\,|\,N], \qquad W(x_*, u_*, \lambda) = \begin{bmatrix} 0_{mm} & 0_{md} & -D(x_*^B)\,B^{\mathrm{T}} \\ 0_{dm} & \lambda D(v_*^N) & 0_{dm} \\ B & N & 0_{mm} \end{bmatrix}.$$

Здесь $B \in \mathbb{R}^{m\times m}$, $N \in \mathbb{R}^{m\times d}$.

Для доказательства достаточно показать, что следующая система однородных алгебраических уравнений имеет только нулевое решение:

$$D(x_*^B)\,B^{\mathrm{T}}\bar{u} = 0_m, \qquad \lambda D(v_*^N)\,\bar{x}^N = 0_d, \qquad B\bar{x}^B + N\bar{x}^N = 0_m,$$

где $\bar{x}^B \in \mathbb{R}^m$, $\bar{x}^N \in \mathbb{R}^d$, $\bar{u} \in \mathbb{R}^m$. Но это очевидно, так как B — невырожденная матрица.

Лемма 2. *Для всяких $x \in \mathbb{R}^n_{++}$, $u \in \operatorname{int} U$, $\lambda \in \mathbb{R}^1_{++}$ матрица $W(x, u, \lambda)$ не вырождена.*

Поскольку на рассматриваемых множествах матрица $M(x, u)$ не вырождена, то доказательство леммы 2 следует из формулы Фробениуса для обратной матрицы:

$$W^{-1} = \begin{bmatrix} \lambda^{-1}D^{-1}(v)\,[I_n - D(x)A^{\mathrm{T}}M^{-1}AD^{-1}(v)] & D^{-1}(v)\,D(x)\,A^{\mathrm{T}}M^{-1} \\ -M^{-1}AD^{-1}(v) & \lambda M^{-1} \end{bmatrix}.$$

Здесь и ниже I_s — единичная матрица $s \times s$.

Объединим векторы x и u одним символом, положив $z^{\mathrm{T}} = [x^{\mathrm{T}}, u^{\mathrm{T}}] \in \mathbb{R}^{n+m}$. Согласно лемме 2, правые части соотношений (1.6) однозначно определены, если все компоненты векторов x и v (u) строго положительны. Векторы z, у которых $z \in \mathbb{R}^n_{++}$, $u \in \operatorname{int} U$, назовем *внутренними*. Те варианты метода (1.3), в которых на всех итерациях вектор z остается внутренним, будем называть *методами внутренних точек*.

Введем в рассмотрение скаляр $\Phi_k = x_k^{\mathrm{T}} v_k$. Из (1.4), (1.6) следует, что

(1.8) $$Ax_{k+1} - b = (1 - \tau_k)\,(Ax_k - b),$$

(1.9) $$\Phi_{k+1} = (1 - \tau_k)\,\Phi_k + (\tau_k - \alpha_k)\,\mu_k^{\mathrm{T}} A x_k + \alpha_k \tau_k \mu_k^{\mathrm{T}} (Ax_k - b).$$

Пусть $x_k > 0_n$, $v_k > 0_n$. Тогда для того, чтобы гарантировать неотрицательность векторов x_{k+1} и v_{k+1}, шаги α_k, τ_k должны удовлетворять условиям $e_n \geq a_k \eta_k$, $e_n \geq \tau_k (e_n - \eta_k)$. Легко видеть, что эти условия имеют место, если

(1.10) $$\alpha_k \leq \alpha_k^{\diamond} = 1/[\eta_k^{\diamond}]_+, \quad 0 < \tau_k \leq \tau_k^{\diamond} = 1/[1 - \eta_{\diamond}^{k}]_+,$$

где $[\alpha]_+ = \max\,[0, \alpha]$, $\eta_k^{\diamond}$ и $\eta_{\diamond}^{k}$ — соответственно, максимальная и минимальная компоненты вектора η_k. Условимся считать, что $\alpha_k^{\diamond} = +\infty$, если $\eta_k^{\diamond} \leq 0$, и что $\tau_k^{\diamond} = +\infty$, если $\eta_{\diamond}^{k} \geq 1$.

Числа $\tau_k^{\diamond}$ и $\alpha_k^{\diamond}$ определяют максимально возможные шаги по прямым и двойственным переменным вдоль ньютоновского направления, сохраняющие неотрицательность всех компонент векторов x и v на k-й итерации.

Лемма 3. *Пусть* $x_k > 0_n$, $v_k > 0_n$, $v_k \in V$ *и* $b \neq 0_m$, *тогда* $\eta_k^{\diamond} > 0$. *Если, кроме того,* $x_k \in X$, *множество* X *ограничено и вектор* c *не принадлежит пространству строк матрицы* A, *то* $\eta_{\diamond}^{k} < 1$.

Доказательство. Предположим противное: пусть $\eta_k^{\diamond} \leq 0$. Тогда из (1.5) следует $A^{\mathrm{T}}\mu_k \leq 0_n$ и, согласно (1.4)—(1.6), выполняются неравенства $b^{\mathrm{T}}u_{k+1} - b^{\mathrm{T}}u_k = \alpha_k b^{\mathrm{T}} M^{-1}(x_k, u_k)\, b > 0$, $v_{k+1} > 0_n$. Таким образом, двигаясь вдоль ньютонова направления, получаем при $\alpha_k \to \infty$, что вектор v_{k+1} имеет все неотрицательные компоненты и целевая функция двойственной задачи стремится к бесконечности. Это противоречит существованию ограниченного решения задачи (1.2). Поэтому $\eta_k^{\diamond} > 0$ и, следовательно, максимальный шаг $\alpha_k^{\diamond}$ ограничен.

Пусть $x_k \in X$. Тогда, согласно (1.8), точка $x_{k+1} \in X$. Если от противного предположить, что $\eta_{\diamond}^{k} \geq 1$, то в случае, когда $\eta_k^{\diamond} > 1$, должна существовать такая j-я компонента вектора η_k, что $\eta_k^j > 1$. Следовательно, в силу (1.4), $x_{k+1}^j \to \infty$ при $\tau_k \to \infty$. Но это невозможно из-за ограниченности X. Если $\eta_{\diamond}^{k} = \eta_k^{\diamond} = 1$, то

$\alpha_k^{\circ} = 1$ и при $\alpha_k = 1$ из (1.6) следует, что $v_{k+1} = 0_n$, $c = A^{\tau} u_k$, что противоречит условиям леммы. Поэтому $\eta_{\circ}^{k} < 1$. Лемма доказана.

Для полного описания численных методов (1.3) и (1.5) необходимо определить правила выбора шагов α_k и τ_k. Укажем три класса методов, которые порождаются разными способами регулировки шагов.

1. Шаги α, τ фиксированы и достаточно малы. В этом случае процесс (1.3) близок к непрерывному варианту метода, который будет рассмотрен в следующем параграфе.

2. Шаги α и τ близки к единице. В этом случае метод близок к методу Ньютона.

3. На каждой итерации шаги α_k и τ_k выбираются из решения вспомогательных оптимизационных задач. Эти варианты можно называть *наискорейшим спуском.*

Ниже остановимся на некоторых вариантах методов из этих классов.

§ 2. Первый класс методов

В работе [17] метод (1.3) был получен из непрерывного варианта, в котором решение задач (1.1) и (1.2) сводилось к отысканию предельных точек (при $t \to \infty$) решений следующей системы обыкновенных дифференциальных уравнений:

$$\lambda D(v)\frac{dx}{dt} - D(x)A^{\tau}\frac{du}{dt} = -\alpha D(x)v, \tag{2.1а}$$

$$A\frac{dx}{dt} = -\tau(Ax - b). \tag{2.1б}$$

Здесь $x(t, z_0)$, $u(t, z_0)$ — решения задачи Коши (2.1) с вектором начальных условий $z_0^{\tau} = [x_0^{\tau}, u_0^{\tau}]$.

Пусть $\alpha > 0$, $\tau > 0$ и $\lambda = 1$, тогда метод (2.1) обладает свойством локальной сходимости. В [17] доказана

Теорема 1. *Пусть $x_{\circ}$ и $u_{\circ}$ — изолированные невырожденные решения задач* (1.1) *и* (1.2). *Тогда пара $[x_{\circ}, u_{\circ}]$ является асимптотически устойчивым положением равновесия для системы* (2.1). *Дискретный вариант метода* (1.3) *локально сходится по крайней мере линейно при фиксированных параметрах λ_k, α_k и τ_k таких, что $\lambda_k = 1$, $0 < \alpha_k < 2$ и $0 < \tau_k < 2$.*

Система (2.1) обладает $n + m$ первыми интегралами:

$$D^{\lambda}(x(t, z_0))\, v(t, z_0) = e^{-\alpha t} D^{\lambda}(x_0)\, v_0, \tag{2.2а}$$

$$Ax(t, z_0) - b = e^{-\tau t}(Ax_0 - b). \tag{2.2б}$$

Отсюда следует, что на траекториях системы (2.1) компоненты векторов x и v не меняют знак. Поэтому если взять $x_0 > 0_n$, $v_0 > 0_n$ и траектории (2.1) ограничены и продолжимы при $t \to \infty$, то в предельных точках $x_{\circ}$, $v_{\circ}$ будут выполнены условия оптимальности Куна — Таккера для задачи (1.1):

$$Ax_{\circ} = b, \quad x_{\circ} \ge 0_n, \quad v_{\circ} \ge 0_n, \quad D(x_{\circ})\, v_{\circ} = 0_n.$$

Таким образом, для решения задач (1.1) и (1.2) можно либо отыскивать предельные точки решения задачи Коши для системы (2.1), либо находить пределы решения нелинейной системы (2.2) при $t \to \infty$. Ниже будем ориентироваться на первый подход. Если все компоненты векторов x и v ненулевые, то систему (2.1) можно разрешить относительно производных и мы получим эквивалентную систему

(2.3) $$dx/dt = \tau D(x)\,[\eta(x,u) - e_n], \qquad du/dt = \alpha\mu(x,u).$$

Правые части системы (2.3) определены во всех тех точках, где вектор $z^{\mathsf T}(t,z_0) = [x^{\mathsf T}(t,z_0),\ u^{\mathsf T}(t,z_0)]$ внутренний. Вместе с тем правые части не определены в точке $z^{\mathsf T}_{\star} = [x^{\mathsf T}_{\star}, u^{\mathsf T}_{\star}]$. Покажем, что если $z(t, z_0)$ приближается к $z_{\star}$, оставаясь внутренним, то вектор-функции $\mu(x(t,z_0), u(t,z_0))$ и $\eta(x(t,z_0), u(t,z_0))$ имеют конечные пределы. Аналогично (1.7) вектор η представим в виде $\eta^{\mathsf T} = [(\eta^B)^{\mathsf T}, (\eta^N)^{\mathsf T}]$, $\eta^B \in \mathbb{R}^m$, $\eta^N \in \mathbb{R}^d$.

Лемма 4. *Пусть точки $x_{\star}$ и $u_{\star}$ являются невырожденными решениями, соответственно, задач* (1.1) *и* (1.2), *причем точка $x_{\star}$ представима в виде* (1.7). *Тогда для произвольных $x > 0$ и $u \in \operatorname{int} U$ имеют место представления*

(2.4) $$\eta^B = e_m - D^{-1}(x^B_{\star})\,\delta x^B + s_1(\delta z),$$

(2.5) $$\eta^N = D^{-1}(v^N_{\star})\,\delta v^N + s_2(\delta z),$$

(2.6) $$\mu = -\delta u + s_3(\delta z),$$

где $\delta x = x - x_{\star}$, $\delta u = u - u_{\star}$, $\delta z^{\mathsf T} = [\delta x^{\mathsf T}, \delta u^{\mathsf T}]$, $\delta v = -A^{\mathsf T}\delta u$; $\|s_i(\delta z)\| = o(\|\delta z\|)$, $i = 1, 2, 3$; *разбиение векторов δx и δv на, соответственно, компоненты δx^B, δx^N и δv^B, δv^N осуществляется аналогично* (1.7).

Доказательство. В соответствии с разбиением вектора x_{*} матрицу $M(x, u)$ можно представить в виде

(2.7) $$M(x,u) = M^B(x,u) + M^N(x,u).$$

Здесь

(2.8) $$M^B(x,u) = BD(x^B)\,D^{-1}(v^B(u))\,B^{\mathsf T}, \qquad M^N(x,u) = ND(x^N)\,D^{-1}(v^N(u))\,N^{\mathsf T}.$$

Если $M^B(x,u)$ — невырожденная матрица, то наряду с (2.7) справедливо равенство

$$M(x,u) = M^B(x,u)\,\{I_m + [M^B(x,u)]^{-1} M^N(x,u)\}.$$

Поэтому

(2.9) $$M^{-1} = [I_m + (M^B)^{-1} M^N]^{-1}(M^B)^{-1} = [M^B(x,u)]^{-1} + S(x,u),$$

где $S = -(M^B)^{-1} M^N [I_m - (M^B)^{-1} M^N + \ldots](M^B)^{-1}$.

На основании (2.8) имеем

(2.10) $$[M^B(x,u)]^{-1} = (B^{\mathsf T})^{-1} D(v^B(u))\,D^{-1}(x^B)\,B^{-1}.$$

Поскольку, в силу невырожденности решений прямой и двойственной задач, $x^N_{\star} = 0_n$, $v^B_{\star} = 0_m$, $v^N_{\star} > 0_d$, то из (2.9) и (2.10) следует, что $[M(x_{*}, u_{*})]^{-1} = [M^B(x_{\star}, u_{\star})]^{-1} = 0_{mm}$.

Имеют место очевидные представления

(2.11а) $b = Bx_{\star}^{B} = Bx^{B} - B\delta x^{B}$,

(2.11б) $\delta v^{N} = -N^{\mathrm{T}}\delta u = N^{\mathrm{T}}(B^{\mathrm{T}})^{-1}\delta v^{B}$, $\|S(x, u)\| = o(\|\delta z\|)$.

Подставляя (2.9) и (2.11а) в выражения (1.5), получаем

(2.12) $\eta(x, u) = D^{-1}(v(u))A^{\mathrm{T}}M^{-1}(x, u)b = \eta_{1}(x, u) + s(\delta z)$.

Здесь $\eta_{1}(x, u) = D^{-1}(v(u))A^{\mathrm{T}}(M^{B})^{-1}B(x^{B} - \delta x^{B})$, $\|s(\delta z)\| = o(\|\delta z\|)$.

Если воспользоваться соотношением (2.10) и первым равенством (2.11б), то приходим к (2.6) и к следующим формулам:

$$\eta_{1}^{B} = D^{-1}(v^{B})B^{\mathrm{T}}(B^{\mathrm{T}})^{-1}D(v^{B})D^{-1}(x^{B})B^{-1}B(x^{B} - \delta x^{B}) =$$
$$= D^{-1}(x^{B})(x^{B} - \delta x^{B}) = e_{m} - D^{-1}(x_{\star}^{B})\delta x^{B} + \theta^{B}(\delta z),$$
$$\eta_{1}^{N} = D^{-1}(v^{N})N^{\mathrm{T}}(B^{\mathrm{T}})^{-1}D(v^{B})D^{-1}(x^{B})B^{-1}B(x^{B} - \delta x^{B}) =$$
$$= D^{-1}(v^{N})N^{\mathrm{T}}(B^{\mathrm{T}})^{-1}v^{B} + \theta^{N}(\delta z) = D^{-1}(v^{N})N^{\mathrm{T}}(B^{\mathrm{T}})\delta v^{B} +$$
$$+ \theta^{N}(\delta z) = D^{-1}(v_{\star}^{N})\delta v^{N} + \theta_{1}^{N}(\delta z),$$

где $\|\theta^{B}(\delta z)\| = o(\|\delta z\|)$, $\|\theta_{1}^{N}(\delta z)\| = o(\|\delta z\|)$. Так как $v_{\star}^{B} = 0$, то отсюда и из (2.12) заключаем, что справедливы равенства (2.4), (2.5). Лемма доказана.

Утверждение леммы объясняет, почему вектор $\eta(x, u)$ часто называют *индикаторным* (см., например, [19]). Все компоненты вектора $\eta(x_{\star}, u_{\star})$ состоят из нулей и единиц, причем базисным компонентам вектора $x_{\star}$ соответствуют единичные компоненты вектора $\eta(x_{\star}, u_{\star})$. Нулевым компонентам вектора $x_{\star}$ соответствуют нулевые компоненты $\eta(x_{\star}, u_{\star})$.

Введем в рассмотрение функцию Ляпунова $F(x, u)$ и лебегово множество Ω_{0} по формулам

$$F(x, u) = \|D^{\lambda}(x)(c - A^{\mathrm{T}}u)\| + \|Ax - b\|,$$
$$\Omega_{0} = \{[x, u]: F(x, u) \le F(x_{0}, u_{0}), \quad x \ge 0_{n}, \quad v \ge 0_{n}, \quad b^{\mathrm{T}}u_{0} \le b^{\mathrm{T}}u\}.$$

Теорема 2. *Пусть* $[x_{\star}, u_{\star}]$ — *пара невырожденных решений задач* (1.1) *и* (1.2). *Пусть, кроме того, лебегово множество* Ω_{0} *ограничено. Тогда для любого внутреннего вектора начальных условий* z_{0} *имеют место следующие свойства:*

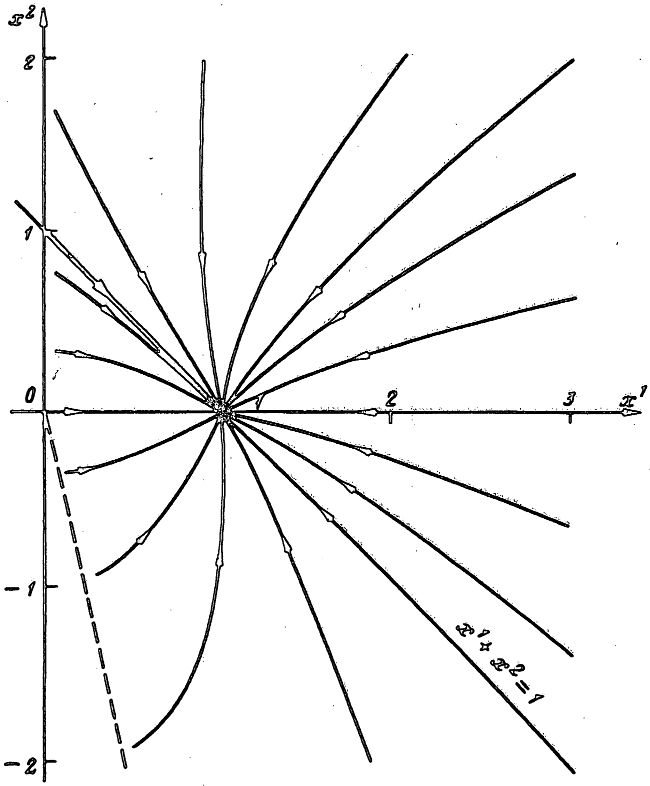
1) *матрица* $M(x(t, z_{0}), u(t, z_{0}))$ *не вырождена при любых* $t \ge 0$,

2) $z(t, z_{0}) \in \Omega_{0}$ *и* $v(t, z_{0}) \in V$ *при всех* $t \ge 0$,

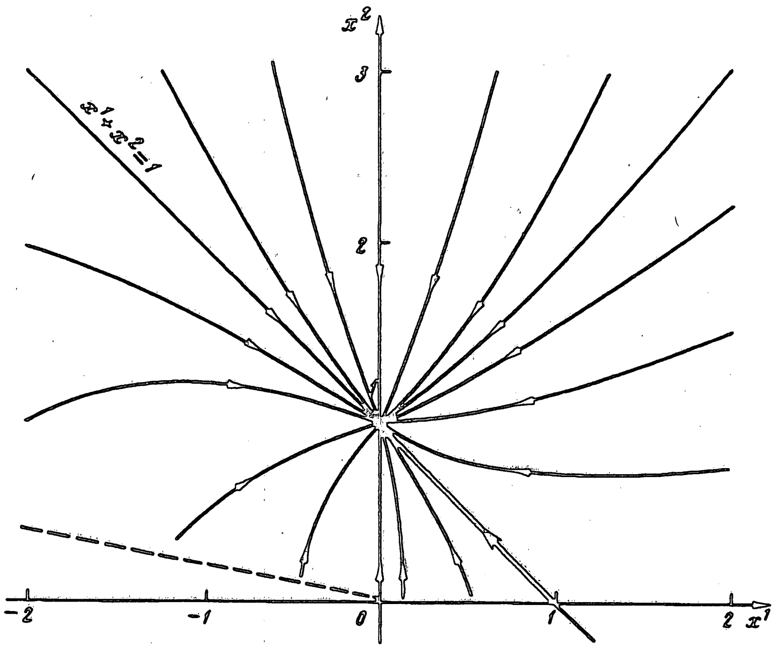
3) *целевая функция двойственной задачи* $b^{\mathrm{T}}u$ *монотонно возрастает на траекториях системы* (2.3),

4) *пара решений* $[x(t, z_{0}), u(t, z_{0})]$ *системы* (2.3) *ограничена, продолжима и сходится к паре* $[x_{\star}, u_{\star}]$ *при* $t \to \infty$.

Для иллюстрации свойств метода рассмотрим простейший пример, в котором $n = 2$, $m = 1$, $A = [1, 1]$, $b = 1$, $c^{\mathrm{T}} = [-2, 1]$. Очевидно, что $x_{\star}^{\mathrm{T}} = [1, 0]$, $u_{\star} = -2$, $v_{\star}^{\mathrm{T}} = [0, 3]$, $f_{\star} = c^{\mathrm{T}}x_{\star} = -2$.



Фиг. 1



Фиг. 2

На фиг. 1 показаны фазовые траектории системы (2.1) в плоскости x^1, x^2. В качестве начального вектора двойственных переменных взят $u_0 = -3$. При этом $v_0 = [1, \ 4]$. Все траектории, начинающиеся из точек, имеющих строго положительные компоненты, сходятся к точке x_*. Штриховой линией обозначен луч $x^1v_0^1 + x^2v_0^2 = 0$, $x^1 \geq 0$, $x^2 \leq 0$, соответствующий множеству точек, где матрица $M(x, \ u_0)$ вырождена. В процессе итераций x и v изменяются и этот луч вращается, приближаясь к вертикальной оси при $t \to \infty$. Из рисунка видно, что сходимость к x_* имеет место при любых x_0, лежащих выше этого луча.

На фиг. 2 показан фазовый портрет в случае, когда $u_0 = 2$ и $v_0 = [-4, -1]$. Все траектории, начинающиеся из $\mathbb{R}^2_{++}$, сходятся к точке $x^* = [0, 1]$, являющейся решением задачи поиска максимума функции $c^{\mathrm{T}}x$ на X. Точка x_* в этом случае неустойчива, а точка x^* является аттрактором.

Интегрируя систему (2.1) по схеме Эйлера, приходим к методу (1.4). Итеративный процесс (1.4) будет близок к процессу, описываемому системой дифференциальных уравнений (2.1), если шаги α_k, τ_k достаточно малы. Свойства этого варианта методов были исследованы Г. В. Смирновым; в частности, он показал полиномиальность дискретного варианта метода при малых α_k и τ_k. Вместе с тем очевидно, что метод (1.4) более эффективен, если шаги α_k и τ_k берутся достаточно большими. Анализ этих вариантов метода проведен ниже.

§ 3. Локальные свойства метода

Из формул (1.5) следует, что вектор $D(x)\eta$ является одним из решений системы линейных алгебраических уравнений $AD(x)\eta = b$. Если $x = x_*$, то у вектора η компоненты, соответствующие базисным компонентам вектора x_*, будут равны единице, а компоненты, соответствующие небазисным компонентам, будут равны нулю. Поэтому $\eta_* = 0$, $\eta^* = 1$, $\alpha^* = \tau^* = 1$. Таким образом, в окрестности решения максимально допустимые шаги, гарантирующие неотрицательность x и v, близки к единице. Рассмотрим простейший вариант выбора α_k, τ_k. Будем считать, что

$$\alpha_k = (1 - \rho_k)\,\alpha_k^*, \qquad \tau_k = (1 - \rho_k)\,\tau_k^*, \tag{3.1}$$

где $0 < \rho_k < 1$. Укажем три правила выбора ρ_k:

$$0 < \rho_k < 1, \quad \lim_{k\to\infty} \rho_k = 0, \tag{3.2}$$

$$\rho_k = \max\,[\varkappa\Phi_k, 1 - \delta], \quad 0 < \varkappa, \quad 0 < \delta < 1, \tag{3.3}$$

$$\rho_k = \varkappa\Phi_k/(1 + \varkappa\Phi_k), \quad \varkappa > 0. \tag{3.4}$$

При всех этих способах гарантируется, что $0 < 1 - \rho_k < 1$. Поэтому если $x_k > 0_n$, $v_k > 0_n$, то на следующей итерации эти векторы также строго положительны.

Теорема 3. *Пусть выполнены предположения леммы 4 и на каждой итерации шаги* α_k *и* τ_k *выбираются согласно* (3.1)—(3.4). *Тогда метод* (1.4) *локально по меньшей мере сверхлинейно сходится к* $[x_*, u_*]$, *т. е.*

$$(3.5)\qquad \overline{\lim_{k\to\infty}} \frac{|x^i_{k+1} - x^i_*|}{|x^i_k - x^i_*|} = 0, \quad \overline{\lim_{k\to\infty}} \frac{|u^j_{k+1} - u^j_*|}{|u^j_k - u^j_*|} = 0, \quad 1 \le i \le n, \quad 1 \le j \le m.$$

$$(3.6)\qquad \lim_{k\to\infty} \alpha^*_k = \lim_{k\to\infty} \tau^*_k = 1, \quad \lim_{k\to\infty} \mu_k = 0_m, \quad \lim_{k\to\infty} \eta^*_k = 1, \quad \lim_{k\to\infty} \eta^k_* = 0.$$

При использовании регулировок (3.3) *или* (3.4) *имеет место квадратичная скорость сходимости.*

Доказательство. Обозначим $\Delta x_k = x_k - x_*$, $\Delta v_k = v_k - v_*$, $\Delta u_k = u_k - u_*$, $\Delta z_k = [\Delta x_k, \Delta u_k]$. Тогда метод (1.4) можно записать в виде

$$(3.7а)\qquad \Delta x_{k+1} = [I_n - \tau_k D(e_n - \eta_k)]\, \Delta x_k - \tau_k D(x_*)(e_n - \eta_k),$$

$$(3.7б)\qquad \Delta v_{k+1} = [I_n - \alpha_k D(\eta_k)]\, \Delta v_k - \alpha_k D(v_*)\, \eta_k,$$

$$(3.7в)\qquad \Delta u_{k+1} = \Delta u_k + \alpha_k \mu_k.$$

Предположим, что пара векторов $[x_k, u_k]$ близка к $[x_*, u_*]$ и норма вектора Δz_k является малой величиной порядка ε. Предположим также для определенности, что базис в точке x_* образуют первые m столбцов матрицы A, т. е. имеет место представление (1.7). Из (2.4) и (2.5) тогда следует, что

$$(3.8а)\qquad \eta^B_k = e_m - D^{-1}(x^B_*)\, \Delta x^B_k + s_1(\Delta z_k),$$

$$(3.8б)\qquad \eta^N_k = D^{-1}(v^N_*)\, \Delta v^N_k + s_2(\Delta z_k),$$

$$(3.8в)\qquad \mu_k = -\Delta u_k + s_3(\Delta z_k),$$

где $\| s_i(\Delta z_k) \| = O(\varepsilon^2)$, $i = 1, 2, 3$.

Подставляя эти соотношения в правые части (3.7), получаем

$$(3.9а)\qquad \Delta x_{k+1} = (1 - \tau_k)\, \Delta x_k + \theta_1(\Delta z_k),$$

$$(3.9б)\qquad \Delta v_{k+1} = (1 - \alpha_k)\, \Delta v_k + \theta_2(\Delta z_k),$$

$$(3.9в)\qquad \Delta u_{k+1} = (1 - \alpha_k)\, \Delta u_k + \theta_3(\Delta z_k).$$

Здесь $\| \theta_i(\Delta z_k) \| = O(\varepsilon^2)$, $i = 1, 2, 3$; α_k и τ_k определяются из условий (3.1). Имеем при сделанных предположениях

$$1/\alpha^*_k = \eta^*_k = 2 - \min_{1\le i\le m} [x^i_k / x^i_*] + O(\varepsilon)^2,$$

$$1/\tau^k_* = \eta^k_* = 2 - \min_{m<i\le n} [v^i_k / v^i_*] + O(\varepsilon^2).$$

Отсюда следует, что

$$(3.10)\qquad \alpha^*_k = 1 + O(\varepsilon), \qquad \tau^*_k = 1 + O(\varepsilon).$$

Эти формулы и (3.1) подставим в правую часть (3.9), после ряда преобразований получим

$$\Delta x_{k+1} = \rho_k \Delta x_k + \widetilde{\theta}_1 (\Delta z_k), \qquad \Delta u_{k+1} = \rho_k \Delta u_k + \widetilde{\theta}_2 (\Delta z_k),$$

где $\|\widetilde{\theta}_i (\Delta z_k)\| = O(\varepsilon^2)$, $i = 1, 2$. С учетом (3.2) приходим к (3.5). Равенства (3.6) следуют из (3.8) и (3.10).

Если имеет место регулировка (3.3) или (3.4), то вблизи решения

$$\rho_k = \gamma \left(\sum_{i=1}^{m} x_*^i \Delta v_k^i + \sum_{i=m+1}^{n} v_*^i \Delta x_k^i \right) + O(\varepsilon^2).$$

Отсюда и из (3.9) заключаем, что $\|\Delta z_{k+1}\| = O(\varepsilon^2)$, т. е. имеет место квадратичная скорость сходимости. Теорема доказана.

Важным свойством метода (1.4) является возможность в некоторых случаях решать задачи (1.1) и (1.2) за конечное число шагов. Рассмотрим это свойство. Введем индексные множества, зависящие от векторов x и v:

$$\sigma(x) = \{1 \le i \le n : x^i = 0\}, \qquad \sigma(v) = \{1 \le i \le n : v^i = 0\}.$$

Если все компоненты векторов x и v не равны нулю, то $\sigma(x) = \varnothing$, $\sigma(v) = \varnothing$.

Обозначим через T_k множество начальных пар $[x_0, u_0]$ таких, что алгоритм (1.3) при $\tau_k = \alpha_k = 1$ дает решение обеих задач (1.1) и (1.2) за k итераций. Определим множества

$$\Omega_1 = \{[x, u] : x = x_*, \ \ \sigma(x) \cap \sigma(v) = \varnothing\},$$

$$\Omega_2 = \{[x, u] : u = u_*, \ \sigma(x) \cap \sigma(v) = \varnothing\},$$

$$\Omega_3 = \{[x, u] : \sigma(x) = \sigma(x_*), \ \sigma(x) \cap \sigma(v) = \varnothing\}.$$

Теорема 4. *Предположим, что задачи* (1.1) *и* (1.2) *имеют невырожденные решения x_* и u_*. Предположим также, что в методе* (1.3) *парамтры выбираются следующим образом:* $\lambda_k = \tau_k = \alpha_k = 1$; *тогда* $\Omega_1 \subseteq T_1$, $\Omega_2 \subseteq T_1$, $\Omega_3 \subseteq T_2$.

Доказательство. Метод (1.3) при $\lambda_k = \tau_k = \alpha_k = 1$ упрощается и имеет вид

$$D(v_k)\, x_{k+1} - D(x_k)\, A^{\mathrm{T}} u_{k+1} = -D(x_k)\, A^{\mathrm{T}} u_k, \qquad A x_{k+1} = b.$$

Поэтому если пара $[x_0, u_0] \in T_1$, то должно выполняться равенство

$$D(v_0)\, x_* = D(x_0)\, A^{\mathrm{T}} (u_* - u_0). \tag{3.11}$$

Легко показать, что любая пара начальных условий $[x_0, u_0]$ из Ω_1 или Ω_2 удовлетворяет (3.11).

Докажем последнее утверждение — о том, что $\Omega_3 \subseteq T_2$. Пусть для точки x_0 имеет место представление, аналогичное (1.7), т. е. $x_0^B \neq 0_m$, $x_0^N = 0_d$, тогда

$$D(v_0^B)\, x_1^B - D(x_0^B)\, B^{\mathrm{T}} u_1 = -D(x_0^B)\, B^{\mathrm{T}} u_0, \ \ D(v_0^N)\, x_1^N = 0_d, \ \ B x_1^B + N x_1^N = b.$$

Все компоненты вектора v_0^B не равны нулю, поэтому

$$x_1^N = 0_d, \ \ x_1^B = B^{-1} b = x_*^B, \ \ u_1 = u_0 + (B^{\mathrm{T}})^{-1} D^{-1}(x_0^B)\, D(v_0^B)\, x_*^B.$$

Таким образом, за один шаг получаем точный вектор $x = x_*$.

Полагая $k = 1$, находим

$$D(v_1^B)\, x_2^B - D(x_1^B)\, B^{\mathrm{T}} u_2 = -D(x_1^B)\, B^{\mathrm{T}} u_1, \quad D(v_1^N)\, x_2^N = 0_d, \quad B x_2^B + N x_2^N = b.$$

Решение этой системы очевидно:

$$x_2^B = x_*^B, \quad x_2^N = 0_d, \quad u_2 = (B^{\mathrm{T}})^{-1} c^B, \quad v_2^B = 0_m,$$

$$v_2^N = c^N - N^{\mathrm{T}} (B^{\mathrm{T}})^{-1} c^B = v_*^N > 0_d,$$

следовательно, найдено точное решение обеих задач. Теорема доказана.

Полученный результат имеет глобальный характер. Если, например, $[x_0, u_0] \in \Omega_2$, то вектор x_0 может иметь даже отрицательные компоненты.

§ 4. Метод внутренней точки с наискорейшим спуском

Правила регулирования шагов α_k и τ_k во внутренней точке $z_k^{\mathrm{T}} = [x_k^{\mathrm{T}}, v_k^{\mathrm{T}}]$ будем определять из решения вспомогательных задач. Чтобы точка z_{k+1}^{T} была внутренней, потребуем выполнения условий

(4.1) $$0 \le \alpha_k \le \omega \alpha_k^* = \bar{\alpha}_k, \quad 0 \le \tau_k \le \omega \tau_k^* = \bar{\tau}_k,$$

где $0 < \omega < 1$. Для простоты предположим, что α_k^* и τ_k^* конечны. Из (4.1) и (1.10) следует неравенство

(4.2) $$1/\bar{\alpha}_k + 1/\bar{\tau}_k \ge 1/\omega.$$

Шаги α_k и τ_k при условии (4.1) целесообразно выбирать так, чтобы по возможности минимизировать величину Φ_{k+1} и норму невязки $\|Ax_{k+1} - b\|$, определяемые по формулам (1.9), (1.8). Опуская для упрощения записи индекс k и вводя обозначения $L = \mu^{\mathrm{T}} A x$, $M = \mu^{\mathrm{T}} b$, $N = \|Ax - b\|$, приходим к двум критериям:

$$\varphi_1(\alpha, \tau) = \Phi - L\alpha + (L - \Phi)\,\tau + (L - M)\,\alpha\tau, \qquad \varphi_2(\tau) = N\,|1 - \tau|.$$

Объединим оба критерия в один, используя их линейную свертку. Получим вспомогательную задачу: найти

(4.3) $$\varphi^0(\bar{\alpha}, \bar{\tau}) = \min_{0 \le \alpha \le \bar{\alpha},\ 0 \le \tau \le \bar{\tau}} \varphi(\alpha, \tau),$$

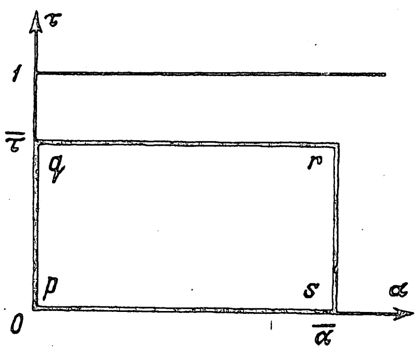
(4.4) $$\varphi(\alpha, \tau) = \varphi_1(\alpha, \tau) + \varphi_2(\tau).$$

Решение задачи (4.3) обозначим α^0, τ^0. Считаем, что $b \ne 0_m$, поэтому $M > 0$. Кроме того, будем принимать во внимание, что на рассматриваемом множестве $\varphi_1(\alpha, \tau) > 0$.

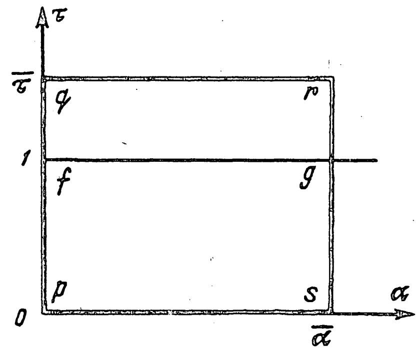
В общем случае функция $\varphi(\alpha, \tau)$ кусочно-билинейна. Поэтому среди экстремальных точек хотя бы одна будет вершиной одного из прямоугольников, являющихся областями ее билинейности.

В случае $\bar{\tau} < 1$ функция $\varphi(\alpha, \tau)$ билинейна и выражается формулой

(4.5) $$\varphi(\alpha, \tau) = \Phi + N - L\alpha + (L - \Phi - N)\,\tau + (L - M)\,\alpha\tau,$$



Фиг. 3



Фиг. 4

областью ее определения является прямоугольник $pqrs$ (фиг. 3). Запишем значения $\varphi(\alpha, \tau)$ в его вершинах:

(4.6а) $\varphi_p = \Phi + N, \qquad \varphi_q = (\Phi + N)(1 - \bar{\tau}) + L\bar{\tau},$

(4.6б) $\varphi_r = (\Phi + N)(1 - \bar{\tau}) + L(\bar{\tau} - \bar{\alpha}) + (L - M)\bar{\alpha}\bar{\tau}, \qquad \varphi_s = \Phi + N - L\bar{\alpha}.$

В точке p не может достигаться $\min\varphi$, так как при $L > 0$ верно $\varphi_s < \varphi_p$, а если $L \leq 0$, то $\varphi_q < \varphi_p$. Вид решения зависит от поведения функции φ на отрезках qr и rs. Учитывая, что $\partial\varphi(\alpha, \bar{\tau})/\partial\alpha = -L(1 - \bar{\tau}) - M\bar{\tau}$, $\partial\varphi(\bar{\alpha}, \tau)/\partial\tau = -(\Phi + N + M\bar{\alpha}) + L(1 + \bar{\alpha})$, находим решение задачи. Оно отражено в табл. 1, значения $\varphi^0(\bar{\alpha}, \bar{\tau})$ определяются по формулам (4.6).

Таблица 1

$\partial\varphi/\partial\tau$ \ $\partial\varphi/\partial\alpha$	$\frac{\partial\varphi}{\partial\alpha}\Big\vert_{\tau=\bar{\tau}} > 0$	$\frac{\partial\varphi}{\partial\alpha}\Big\vert_{\tau=\bar{\tau}} = 0$	$\frac{\partial\varphi}{\partial\alpha}\Big\vert_{\tau=\bar{\tau}} < 0$
$\frac{\partial\varphi}{\partial\tau}\Big\vert_{\alpha=\bar{\alpha}} < 0$	точка q	отрезок qr	точка r
$\frac{\partial\varphi}{\partial\tau}\Big\vert_{\alpha=\bar{\alpha}} = 0$	не реализуется	не реализуется	отрезок rs
$\frac{\partial\varphi}{\partial\tau}\Big\vert_{\alpha=\bar{\alpha}} > 0$	не реализуется	не реализуется	точка s

В случае $\bar{\tau} \geq 1$ при $\tau \leq 1$ функция $\varphi(\alpha, \tau)$ имеет вид (4.5), а при $\tau \geq 1$

$$\varphi(\alpha, \tau) = \Phi - N - L\alpha + (L - \Phi + N)\tau + (L - M)\alpha\tau.$$

Функция $\varphi(\alpha, \tau)$ кусочно-билинейна. Подобластями билинейности являются прямоугольники $pfgas$ и $fqrg$ (фиг. 4). Найдем значения $\varphi(\alpha, \tau)$ в их вершинах:

(4.7а) $\varphi_p = \Phi + N, \qquad \varphi_f = L, \qquad \varphi_q = (\Phi - N)(1 - \bar{\tau}) + L\bar{\tau},$

(4.7б) $\varphi_r = (\Phi - N + L\bar{\alpha})(\bar{\tau} - 1) + (L - M\bar{\alpha})\bar{\tau}, \quad \varphi_g = L - M\bar{\alpha}, \quad \varphi_s = \Phi + N - L\bar{\alpha},$

Вершина p не может быть точкой минимума $\varphi(\alpha, \tau)$, так как, ввиду $\varphi_1(\alpha, \tau) > 0$, должно быть $L > 0$ и поэтому $\varphi_s < \varphi_p$. Минимум $\varphi(\alpha, \tau)$ не может достигать я и в вершине f, так как $\varphi_g < \varphi_f$. Анализируя поведение $\varphi(\alpha, \tau)$ на сторонах qr и rs, используя соотношения

$$(\partial\varphi/\partial\alpha)_{\tau=\bar{\tau}} = L(\bar{\tau} - 1) - M\bar{\tau},$$

$$(\partial\varphi/\partial\tau)_{\substack{\alpha=\bar{\alpha}\\ \tau>1}} = L(1 + \bar{\alpha}) + N - (\Phi + M\bar{\alpha}) \geq (\partial\varphi/\partial\tau)_{\substack{\alpha=\bar{\alpha}\\ \tau<1}} =$$

$$= L(1 + \bar{\alpha}) - (\Phi + N + M\bar{\alpha}),$$

получаем возможные варианты решения задачи. Они отражены в табл. 2, значения $\varphi^0(\bar{\alpha}, \bar{\tau})$ находятся по формулам (4.7).

Таблица 2

$\partial\varphi/\partial\tau$ \ $\partial\varphi/\partial\alpha$	$\frac{\partial\varphi}{\partial\alpha}\Big\vert_{\tau=\bar{\tau}} > 0$	$\frac{\partial\varphi}{\partial\alpha}\Big\vert_{\tau=\bar{\tau}} = 0$	$\frac{\partial\varphi}{\partial\alpha}\Big\vert_{\tau=\bar{\tau}} < 0$
$\frac{\partial\varphi}{\partial\tau}\Big\vert_{\alpha=\bar{\alpha},\, \tau>1} < 0$	точка q	отрезок qr	точка r
$\frac{\partial\varphi}{\partial\tau}\Big\vert_{\substack{\alpha=\bar{\alpha}\\ \tau<1}} < 0 = \frac{\partial\varphi}{\partial\tau}\Big\vert_{\substack{\alpha=\bar{\alpha}\\ \tau>1}}$		ломаная qrg	отрезок rg
$\frac{\partial\varphi}{\partial\tau}\Big\vert_{\substack{\alpha=\bar{\alpha}\\ \tau<1}} = 0 = \frac{\partial\varphi}{\partial\tau}\Big\vert_{\substack{\alpha=\bar{\alpha}\\ \tau>1}}$	не реализуется		отрезок rs
$\frac{\partial\varphi}{\partial\tau}\Big\vert_{\substack{\alpha=\bar{\alpha}\\ \tau<1}} < 0 < \frac{\partial\varphi}{\partial\tau}\Big\vert_{\substack{\alpha=\bar{\alpha}\\ \tau>1}}$	arg min $\{\varphi_q, \varphi_g\}$	точка g	
$\frac{\partial\varphi}{\partial\tau}\Big\vert_{\substack{\alpha=\bar{\alpha}\\ \tau<1}} = 0 < \frac{\partial\varphi}{\partial\tau}\Big\vert_{\substack{\alpha=\bar{\alpha}\\ \tau>1}}$	arg min $\{\varphi_q, \varphi_{sg}\}$	отрезок sg	
$0 < \frac{\partial\varphi}{\partial\tau}\Big\vert_{\alpha=\bar{\alpha},\, \tau<1}$	arg min $\{\varphi_q, \varphi_s\}$	точка s	

Таким образом, если следовать полученным оптимальным стратегиям выбора шагов α_k и τ_k, то в любом случае шаг α_k можно принять равным 0 либо $\bar{\alpha}$. Шаг τ_k в случае $\bar{\tau} \geq 1$, помимо крайних значений 0 и $\bar{\tau}$, может также принимать промежуточное значение 1. Метод внутренней точки (1.4) с шагами α_k и τ_k, выбираемыми из решения вспомогательной задачи (4.4), назовем *методом наискорейшего спуска.*

Оценим величину $\varphi^0(\bar{\alpha}, \bar{\tau})$ в предположении, что

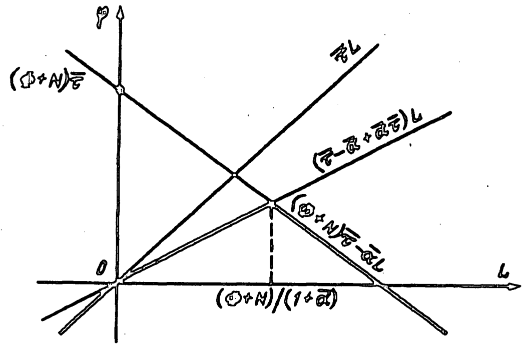
$$\|\eta_k\|_\infty = \max_{1\leq i\leq n} |\eta_k^i| \leq C.$$

В этом случае, согласно (1.10) и (4.1), имеем

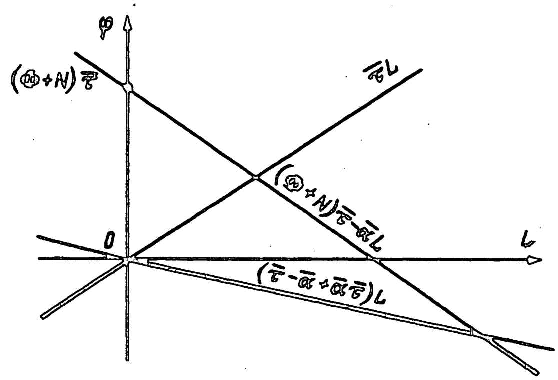
(4.8) $\bar{\alpha} \geq \omega/C, \quad \bar{\tau} \geq \omega/(1 + C).$

Найдем

$$\psi^* = \sup_{\bar{\alpha}\geq\omega/C,\ \bar{\tau}\geq\omega/(1+C)} \varphi^0(\bar{\alpha}, \bar{\tau}).$$



Фиг. 5



Фиг. 6

Воспользуемся тем, что $\varphi^0(\bar\alpha, \bar\tau)$ не должно превышать значение $\varphi(\alpha, \tau)$ в любой из допустимых точек $[\alpha, \tau]$.

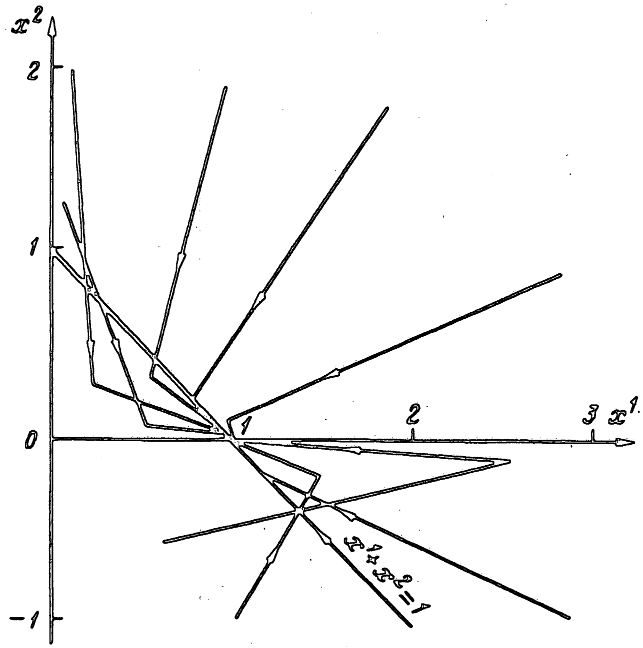
Для случая $\bar\tau \geq 1$ в качестве таких точек возьмем s и f:

(4.9) $$\varphi^0(\bar\alpha, \bar\tau) \leq \min\{\varphi_s, \varphi_f\} \leq$$
$$\leq \max_L \min\{L, \Phi + N - \bar\alpha L\} = (\Phi + N)/(1 + \bar\alpha).$$

Для случая $\bar\tau < 1$ обратимся к точкам q, r, s. Так как $M > 0$, то $\varphi_r < (\Phi + N)(1 - \bar\tau) + (\bar\tau - \bar\alpha + \bar\alpha\,\bar\tau)L$. Поэтому

(4.10) $$\varphi^0(\bar\alpha, \bar\tau) \leq \min\{(\Phi + N)(1 - \bar\tau) + L\bar\tau,\ (\Phi + N)(1 - \bar\tau) +$$
$$+ (\bar\tau - \bar\alpha + \bar\alpha\,\bar\tau)L,\ \Phi + N - \bar\alpha L\} \leq (\Phi + N)(1 - \bar\tau) +$$
$$+ \max_L \min\{\bar\tau L,\ (\bar\tau - \bar\alpha + \bar\alpha\,\bar\tau)L,\ (\Phi + N)\bar\tau - \bar\alpha L\}.$$

Из $\bar\tau \leq 1$ вытекает, что $\bar\tau - \bar\alpha + \bar\alpha\,\bar\tau \leq \bar\tau$. Таким образом, если $\bar\tau - \bar\alpha + \bar\alpha\,\bar\tau \geq 0$, то максимум по L находится из условия $(\bar\tau - \bar\alpha + \bar\alpha\,\bar\tau)L = (\Phi + N)\bar\tau - \bar\alpha L$ и достигается в точке $L = (\Phi + N)/(1 + \bar\alpha)$, как показано на фиг. 5. Подставляя данное значение L в правую часть (4.10), приходим к выводу, что оценка (4.9) сохраняется и при этом предположении. Случаю $\bar\tau - \bar\alpha + \bar\alpha\,\bar\tau < 0$ соответствует



Фиг. 7

ситуация, изображенная на фиг. 6. Максимум определяется из условия $\bar{\tau}L = (\bar{\tau} - \bar{\alpha} + \bar{\alpha}\,\bar{\tau})\,L$ и достигается в нуле. Вместо (4.9) выполняется

$$(4.11)\qquad \varphi^0(\bar{\alpha}, \bar{\tau}) \le (\Phi + N)(1 - \bar{\tau}).$$

Из (4.1), (4.9) и (4.11) следует, что

$$(4.12)\qquad \varphi^{\ast} \le \nu\,(\Phi + N),$$

$$\nu = \max\left\{\max_{\bar{\alpha}\ge\omega/C}\frac{1}{1+\bar{\alpha}},\ \max_{\bar{\tau}\ge\omega/(1+C)}(1-\bar{\tau})\right\} = 1 - \frac{\omega}{1+C}.$$

Полученная оценка недостижима и неулучшаема. Недостижимость ($\varphi^0(\bar{\alpha}, \bar{\tau}) < \varphi^{\ast}$) вытекает из предыдущих рассуждений, а неулучшаемость видна из следующего примера: полагая $\Phi = N = 1$, $L = 0$, $M \to 0$, $\omega = 0.8$, $C = 0.6$, $\alpha^{\ast} = 1.875$, $\tau^{\ast} = 0.625$ и убеждаясь, что выполнены условия (4.2), (4.8), $\bar{\tau} < 1$, что оптимальна точка r, с помощью (4.6) и (4.12) находим $\varphi^0 = 1-3M/4$, $\nu(\Phi + N) = 1$, т. е. $\lim\limits_{M\to 0}\varphi^0 = \nu(\Phi + N)$.

На основании (4.12), используя определение $\varphi^{\ast}$, получаем

$$\Phi_{k+1} + \|Ax_{k+1} - b\| \le \nu\,(\Phi_k + \|Ax_k - b\|).$$

Это неравенство позволяет оценить число шагов, достаточных для попадания точки $[x_k, u_k]$ в некоторую окрестность решения задач (1.1) и (1.2).

Теорема 5. *Пусть $z_0 = [x_0, u_0]$ — внутренняя точка, и пусть последовательности $\{x_k\}$ и $\{u_k\}$, порождаемые методом наискорейшего спуска, таковы,* *что* $\|\eta_k\|_\infty \le C$ *для всех* k. *Тогда для произвольного* $\varepsilon > 0$ *функция* $\Phi(x, u) = v^{\mathrm{T}}(u)\,x + \|Ax - b\|$ *становится меньше* ε *не более чем за*

$$K = \left\lceil \frac{1 + C}{\omega} \ln \frac{\Phi(x_0, u_0)}{\varepsilon} \right\rceil$$

итераций, где $\lceil a \rceil$ — *наименьшее целое, приближающее число* a *сверху.*

На фиг. 7 показано решение примера из § 2 методом наискорейшего спуска. Параметр ω полагался равным 0.9. Отметим, что при приближении ω к единице количество итераций, необходимых для решения задачи с заданной точностью, уменьшалось, доходя для некоторых начальных точек до двух.

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

1. *Ортега Д., Рейнболдт В.* Итерационные методы решения нелинейных уравнений со многими неизвестными. М.: Мир, 1975.
2. *Васильев Ф. П.* Численные методы решения экстремальных задач. М.: Наука, 1980.
3. *Евтушенко Ю. Г.* Методы решения экстремальных задач и их применение в системах оптимизации. М.: Наука, 1982.
4. *Iri M., Imai H.* A multiplicative barrier function method for linear programming//Algorithmica. 1986. № 1. P. 455—482.
5. *De Ghellink G., Vial J.-P.* A polynomial Newton method for linear programming//Algorithmica. 1986. № 1. P. 425—453.
6. *Renegar J.* A polynomial-time algorithm, based on Newton's method for linear programming//Math. Program. 1988. V. 40. № 1. P. 55—94.
7. *Gonzaga C. C.* Path-following methods for linear programming//SIAM Rev. 1992. V. 34. № 2. P. 167—224.
8. *Kojima M., Mizuno S., Yoshise A.* A primal-dual interior point method for linear programming//Progress Math. Program. Interior Point and Relative Methods/Ed. N. Megiddo. Berlin: Springer-Verlag, 1989. Ch. 2.
9. *Monteiro R. C., Adler I.* Interior path-following primal-dual algorithm, part I: Linear programming//Math. Program. 1989. V. 44. P. 43—66.
10. *McShane K., Monma C., Shanno D.* An implementation of a primal-dual interior point method for linear programming//ORSA J. Comput. 1989. № 1. P. 70—83.
11. *Ye Y., Tapia R., Zhang Y.* A superlinear convergent $O(\sqrt{n}\,L)$-iteration algorithm for linear programming: Techn. Rept TR91-22, 1991, Rice Univ. Houston, Texas.
12. *Jansen B., Roos C., T. Terlaky, Vial J.-Ph.* Primal-dual target-following algorithms for linear programming: Techn. Rept 93—107, Fac. Techn. Math. and Informatics, TU Delft, 1993.
13. *Еремин И. И., Астафьев Н. Н.* Введение в теорию линейного и выпуклого программирования. М.: Наука, 1976.
14. *Evtushenko Yu. G., Zhadan V. G.* The space transformation technique in mathematical programming//System Modelling and Optimizat. Ed. P. Kall. Lect. Notes in Control and Information Sci. 180. Proc. 15th IFIP Conf. Springer-Verlag, 1991.
15. *Evtushenko Yu. G., Zhadan V. G.* Stable barrier-projection and barrier-Newton methods in nonlinear programming//Optimizat Meth. and Software. 1994. V. 3. P. 237—256.
16. *Евтушенко Ю. Г., Жадан В. Г.* Барьерно-проективные и барьерно-ньютоновские численные методы оптимизации (случай нелинейного программирования)//Сообщ. по вычисл. матем. М.: ВЦ АН СССР, 1991.
17. *Евтушенко Ю. Г., Жадан В. Г.* Барьерно-проективные и барьерно-ньютоновские численные методы оптимизации (случай линейного программирования)//Сообщ. по вычисл. матем. М.: ВЦ РАН, 1992.
18. *Ye Y.* Line search in potential reduction algorithms for linear programming//Dept. Management Sci. Iowa City, IA: Univ. Iowa, 1989.
19. *El-Bakly A. S., Tapia R. A., Zhang Y.* A study of indicators for identifing zero variables in interior-point methods: Techn. Rept. TR91-15, 1991. Rice Univ., Houston, Texas.

Поступила в редакцию 05.07.94