ЖУРНАЛ
ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ ФИЗИКИ

Том 34,1994 — N° 5, с. 669–684

УДК 519.853.6

БАРЬЕРНО-ПРОЕКТИВНЫЕ МЕТОДЫ РЕШЕНИЯ ЗАДАЧ НЕЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ[1]

Ю.Г. ЕВТУШЕНКО, В.Г. ЖАДАН

117967 Москва, ул. Вавилова, 40, ВЦ РАН

(Пересмотрена 26 апреля 2002 г.)

Рассматривается класс барьерно-проективных методов решения задач нелинейного программирования. Развивается общий подход к их построению, основанный на преобразовании пространств. Основное внимание уделено асимптотически устойчивым вариантам методов. Доказывается сходимость непрерывных и дискретных вариантов методов, приводятся оценки скорости сходимости.

Введение

Метод проекции градиента был одним из первых численных методов решения задач линейного и нелинейного программирования [1]–[5]. Впоследствии возник подход, в котором объединялись идеи проектирования градиента и метода барьерных функций. В [6, 7] он был использован для решения задач линейного и квадратичного программирования. В работах [8]–[12] данный подход был распространен на случай общих задач нелинейного программирования и исследования операций. Интерес к этому направлению значительно усилился после публикации в 1984 г. статьи Кармаркара [13]. В появившихся вслед за ней статьях [14] – [16] были вновь предложены варианты методов внутренней точки для задач линейного программирования.

Настоящая работа является развитием результатов, полученных в [8]– [12]. Здесь описывается единый подход к построению барьерно-проективных методов, основанный на переходе к новым пространствам, в которых структура допустимого множества существенно проще, чем в исходном пространстве. Это позволяет использовать для отыскания решений в преобразованном пространстве метод проекции градиента в чистом виде. После возвращения в исходное пространство получаются различные варианты методов, названные в [12] **барьерно-проективными**. При этом основное внимание уделяется их устойчивым вариантам, для которых не требуется, чтобы начальное приближение принадлежало допустимому множеству. В случае когда начальные приближения для этих методов взяты из допустимого множества, они ведут себя, как релаксационные методы внутренней точки, т. е. порождаемые ими траектории не покидают допустимого множества и минимизируемая функция убывает вдоль траекторий. Для допустимых множеств достаточно общего

[1]Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (код проекта 94-01-01379).

вида показано, что если выбранное преобразование пространств удовлетворяет определенным условиям, то для всех непрерывных вариантов барьерно-проективных методов решение поставленной задачи нелинейного программирования является асимптотически устойчивым положением равновесия, а их дискретные аналоги локально сходятся к этому решению со скоростью геометрической прогрессии.

§ 1. Устойчивый вариант метода проекции градиента для решения задач с ограничениями типа равенства

Пусть x — вектор из n-мерного евклидова пространства $\mathbb{E}^n$. Всюду на $\mathbb{E}^n$ определена скалярная функция $f(x)$ и вектор-функция $g(x) : \mathbb{E}^n \to \mathbb{E}^m$. Предполагаем, что f и g непрерывны вместе со своими первыми производными. Всюду считаем, что все поставленные ниже оптимизационные задачи имеют решения. Пусть ищется

$$f_* = \min_{x \in X} f(x), \qquad X = \{x \in \mathbb{E}^n : g(x) = 0_m\}. \tag{1.1}$$

Здесь и ниже 0_{ij} — нулевая матрица размером $i \times j$, 0_i — нулевой i-мерный вектор, I_i — единичная матрица порядка i. В тех случаях, когда это очевидно, индексы при нулевых и единичных матрицах и векторах будем опускать.

Введем вектор двойственных переменных $u \in \mathbb{E}^m$ и, составив функцию Лагранжа, вычислим ее градиент:

$$L(x,u) = f(x) + u^\top g(x), \qquad L_x(x,u) = f_x(x) + g_x^\top(x)u.$$

Здесь g_x — матрица первых производных размером $m \times n$.

Введем в рассмотрение следующую систему из n обыкновенных дифференциальных уравнений:

$$\frac{dx}{dt} = -L_x(x, u(x)), \qquad x(x_0, 0) = x_0. \tag{1.2}$$

Функцию $u(x)$ в правой части (1.2) выберем таким образом, чтобы все траектории системы при $t \to +\infty$ приближались к допустимому множеству X. Для этого потребуем, чтобы

$$\frac{dg}{dt} = -\tau g(x), \qquad \tau > 0. \tag{1.3}$$

Дифференцируя функцию $g(x)$ в силу системы (1.2) и используя (1.3), получаем

$$\frac{dg}{dt} = -g_x(x)[f_x(x) + g_x^\top(x)u(x)] = -\tau g(x). \tag{1.4}$$

Если матрица Грама $\Gamma(x) = g_x(x)g_x^\top(x)$ не вырождена, то из этого условия определяем функцию $u(x)$:

$$u(x) = \Gamma^{-1}(x)[\tau g(x) - g_x(x)f_x(x)].$$

Подставив эту функцию в правую часть (1.2), получим

$$\frac{dx}{dt} = -\{f_x(x) + g_x^\top(x)\Gamma^{-1}(x)[\tau g(x) - g_x(x)f_x(x)]\}. \tag{1.5}$$

Производная функции $f(x)$ в силу этой системы равна

$$\frac{df}{dt} = -||L_x(x, u(x))||^2 + \tau u^\top(x)g(x). \tag{1.6}$$

Здесь и ниже $\|\cdot\|$ — евклидова норма в $\mathbb{E}^n$, $\langle\cdot,\cdot\rangle$ — евклидово скалярное произведение.

Если точка $x_0 \in X$ или $\tau = 0$, то система (1.2), (1.4) переходит в метод проекции градиента, в котором

$$dx/dt = -[f_x(x) + g_x^\top(x)u(x)], \qquad x_0 \in X, \tag{1.7}$$

$$g_x(x)g_x^\top(x)u(x) + g_x(x)f_x(x) = 0. \tag{1.8}$$

Согласно (1.6), в данном случае функция $f(x(x_0,t))$ монотонно убывает. В общем случае эта функция может изменяться немонотонно, и лишь на множестве X и вблизи множества X, где норма $\|g(x)\|$ мала, эта функция монотонно убывает. Система (1.7), (1.8) является нейтрально устойчивой по отношению к допустимому множеству, так как если $g(x_0) = c$, $\|c\| \neq 0$, то $g(x(x_0,t)) \equiv c$. Система (1.5) согласно (1.4) является асимптотически устойчивой по отношению к ограничениям. Действительно, если решения системы (1.5) продолжимы при $t \to \infty$, то

$$g(x(x_0,t)) = g(x_0)e^{-\tau t}, \qquad \lim_{t\to\infty} g(x(x_0,t)) = 0.$$

Следовательно, при $t \to +\infty$ траектории системы приближаются к точкам допустимого множества X.

Определение 1. *В точке $x \in \mathbb{E}^n$ выполнено* **условие регулярности ограничений** *(у.р.о.) для задачи* (1.1), *если столбцы матрицы $g_x^\top(x)$ линейно независимы.*

Обозначим через $K^\perp(x)$ векторное подпространство, порожденное векторами $g_x^1(x),\dots$, $\dots, g_x^m(x)$. Если в точке x выполнено у.р.о., то размерность подпространства $K^\perp$ равна m, ортогональным дополнением к нему будет подпространство $K(x) = \{\bar{x} \in \mathbb{E}^n : g_x(x)\bar{x} = 0_m\}$ размерности $d = n - m$.

Пусть W — матрица $m \times n$, ее ранг максимальный, равен m. Матрица W^+ размера $n \times m$ обозначает правую псевдообратную для матрицы W, т.е. $W^+ = W^\top(WW^\top)^{-1}$, $WW^+ = I_m$. Через $\pi(W)$ обозначим матрицу $\pi(W) = I_n - W^+W$. Систему (1.5) можно записать в проективном виде:

$$\frac{dx}{dt} = -\pi(g_x(x))f_x(x) - \tau(g_x(x))^+g(x). \tag{1.9}$$

Первый вектор, стоящий в правой части, является проекцией антиградиента функции $f(x)$ на касательное подпространство $K(x)$ к многообразию $g(x) = const$, второй вектор лежит в ортогональном подпространстве $K^\perp(x)$.

Те точки $x \in \mathbb{E}^n$, в которых правые части (1.9) обращаются в нуль, называются **стационарными**.

Лемма 1. *Точка x_*, в которой выполнено у.р.о., является стационарной тогда и только тогда, когда пара $[x_*, u_*]$, где $u_* = u(x_*)$, образует точку Куна–Таккера, т.е.*

$$L_x(x_*,u_*) = 0_n, \qquad g(x_*) = 0_m. \tag{1.10}$$

Доказательство. Достаточность очевидна, поэтому докажем только необходимость. Пусть x_* — стационарная точка. Тогда с учетом ортогональности векторов, стоящих в правой части (1.9), получаем, что

$$\pi(g_x(x_*))f_x(x_*) = 0_n, \qquad (g_x(x_*))^+g(x_*) = 0_n.$$

Здесь первые n соотношений совпадают с первыми n соотношениями (1.10). В силу у.р.о. вторые n соотношений могут быть удовлетворены только при $g(x_*) = 0$. Таким образом, $[x_*, u_*]$ является точкой Куна–Таккера, что и требовалось доказать. $\square$

Если в стационарной точке x_* выполнены достаточные условия второго порядка изолированного локального минимума в задаче (1.1), приведенные в [17], то траектории системы (1.9) локально экспоненциально сходятся к x_*, траектории системы (1.7), (1.8) локально сходятся к X_* на допустимом множестве X. Справедливость этого утверждения следует из более общего результата, полученного в § 3.

§ 2. Учет дополнительных ограничений простой структуры

Изложенный в предыдущем параграфе подход можно использовать для задач, имеющих более сложные ограничения. Пусть решается задача об отыскании

$$f_* = \min_{x \in X} f(x), \quad X = \{x \in \mathbb{E}^n : g(x) = 0_m, \;\; x \in \Pi\}, \tag{2.1}$$

где $\Pi \subset \mathbb{E}^n$ — выпуклое замкнутое множество, имеющее непустую внутренность.

Введем новое n-мерное евклидово пространство $\mathbb{E}^n$ с координатами $[y^1, \ldots, y^n]$. Осуществим переход от этого пространства к исходному с помощью преобразования $x = \xi(y)$. Это преобразование построим так, чтобы оно было сюръекцией из $\mathbb{E}^n$ в Π или по крайней мере из $\mathbb{E}^n$ в $int\,\Pi$. Тогда каждый элемент из Π (или $int\,\Pi$) есть образ не менее чем одного элемента из $\mathbb{E}^n$ и замыкание образа множества $\mathbb{E}^n$ совпадает с Π. Предполагаем также, что отображение $\xi(y)$ непрерывно дифференцируемо всюду на $\mathbb{E}^n$. Исходную задачу (2.1) заменим следующей: найти

$$\tilde{f}_* = \inf_{y \in Y} \tilde{f}(y), \qquad Y = \{y \in \mathbb{E}^n : \tilde{g}(y) = 0_m\}. \tag{2.2}$$

Здесь $\tilde{f}(y) = f(\xi(y))$, $\tilde{g}(y) = g(\xi(y))$, причем $\tilde{f}_* = f_*$.

Предположим, что точка y_* является решением задачи (2.2). Тогда точка $x_* = \xi(y_*)$ будет решением задачи (2.1). Поэтому можно применить для решения задачи (2.2) описанный в предыдущем параграфе метод, после этого совершить обратный переход к координатам $[x^1, \ldots, x^n]$ и получить таким образом метод решения задачи (2.1) в исходном пространстве. В пространстве $y \in \mathbb{E}^n$ метод (1.2), (1.4) принимает следующий вид:

$$dy/dt = -[\tilde{f}_y(y) + \tilde{g}_y^\top(y)\tilde{u}(y)], \qquad y(y_0, 0) = y_0 \in \mathbb{E}^n, \tag{2.3}$$

$$\tilde{g}_y(y)\tilde{g}_y^\top(y)\tilde{u}(y) + \tilde{g}_y(y)\tilde{f}_y(y) = \tau \tilde{g}(y). \tag{2.4}$$

Градиенты функций $\tilde{f}$, $\tilde{g}$ и f, g связаны очевидными соотношениями

$$\tilde{f}_y(y) = \tilde{J}^\top(y) f_x(\xi(y)), \qquad \tilde{g}_y^\top(y) = \tilde{J}^\top(y) g_x^\top(\xi(y)).$$

Здесь $\tilde{J}(y) = d\xi(y)/dy$ — матрица Якоби. В неособых точках преобразования $x = \xi(y)$, где якобиан отличен от нуля, существует обратное преобразование $y = \xi^{-1}(x)$. Если воспользоваться этим преобразованием и взять в качестве аргумента у матрицы Якоби вектор x, то получим матрицу $J(x) = \tilde{J}(\xi^{-1}(x))$, зависящую уже от x. Используя соотношения

$$\frac{dx}{dt} = \frac{d\xi(y)}{dy}\frac{dy}{dt} = \tilde{J}(y)\frac{dy}{dt} = J(x)\frac{dy}{dt},$$

можно найти дифференциальные уравнения, определяющие в пространстве x траектории, которые соответствуют траекториям (2.3), (2.4). При этом надо иметь в виду, что если для (2.3), (2.4) условие $x_0 = \xi(y_0) \in \Pi$ выполняется автоматически, то в пространстве x надо потребовать, чтобы $x_0 \in \Pi$. Из (2.3), (2.4) получим

$$\frac{dx}{dt} = -G(x)L_x(x, u(x)), \qquad x(x_0, 0) = x_0 \in \Pi, \tag{2.5}$$

$$\Gamma(x)u(x) + g_x(x)G(x)f_x(x) = \tau g(x). \tag{2.6}$$

Здесь введены две матрицы Грама: $\Gamma(x) = g_x(x)G(x)g_x^\top(x)$, $G(x) = J(x)J^\top(x)$. Аналогом формулы (1.6) будет

$$\frac{df}{dt} = -||J^\top(x)L_x(x, u(x))||^2 + \tau u^\top(x)g(x). \tag{2.7}$$

При $\tau = 0$ метод переходит в следующий:

$$\frac{dx}{dt} = -G(x)L_x(x, u(x)), \qquad x(x_0, 0) = x_0 \in X, \tag{2.8}$$

$$\Gamma(x)u(x) + g_x(x)G(x)f_x(x) = 0. \tag{2.9}$$

Методы типа (2.5), (2.6), (2.8), (2.6) назовем **барьерно-проективными**.

Обозначим, соответственно, через $K(x|\Pi)$ и $K^*(x|\Pi)$ конус допустимых направлений в точке x относительно множества Π и двойственный к нему:

$$\begin{array}{rcl} K(x|\Pi) & = & \{z \in \mathbb{E}^n : \exists \lambda(z) > 0 \ \text{ такое, что } \ x + \lambda z \in \Pi \quad \forall\, 0 < \lambda < \lambda(z)\}, \\ K^*(x|\Pi) & = & \{z \in \mathbb{E}^n : \langle z, y\rangle \geq 0 \quad \forall y \in K(x|\Pi)\}. \end{array}$$

Пусть $S(x|\Pi)$ — линейная оболочка конуса $K^*(x|\Pi)$.

Определение 2. *В точке $x \in \Pi$ выполнено у.р.о. для задачи* (2.1)*, если все векторы $g_x^i(x)$, $1 \leq i \leq m$, и произвольный ненулевой вектор $p \in S(x|\Pi)$ линейно независимы.*

Наложим на преобразование $\xi(y)$ следующее условие.

Условие 1. *В каждой точке $x \in \Pi$ матрица $J(x)$ определена и ядро $\ker J^\top(x)$ совпадает с $S(x|\Pi)$.*

Из этого условия, в частности, следует, что во всех внутренних точках $x \in \operatorname{int}\Pi$ матрица $J(x)$ не вырождена, она становится особой только на границе множества Π.

Отметим также, что согласно условию 1 ортогональное к $S(x|\Pi)$ подпространство $S^\perp(x|\Pi)$ совпадает с пространством столбцов матрицы $J(x)$, и поскольку вектор $-J(x)$ $J^\top(x)L_x(x, u(x))$ принадлежит этому пространству, то вектор скорости $\dot{x}$ всегда лежит в ортогональном подпространстве $S^\perp(x|\Pi)$. Таким образом, если x — граничная точка Π, то в силу вырожденности матрицы $J^\top(x)$ вектор $\dot{x}$ будет принадлежать собственному подпространству пространства $\mathbb{E}^n$, которое совпадает с пространством $M(x) - x$, где $M(x)$ — пересечение всех опорных плоскостей к множеству Π в точке x. Это подпространство вырождается в единственную точку (начало координат), если конус $K^*(x|\Pi)$ имеет непустую внутренность.

Матрица $G(x)$ играет в (2.5), (2.6) роль барьера, не позволяющего траекториям покидать множество Π. Действительно, для того чтобы траектория $x(x_0, t)$, начинающаяся внутри множества Π, в некоторый момент $t_1 > 0$ покидала его, необходимо, чтобы нашелся такой вектор $p \in K^*(x(x_0, t_1)|\Pi)$, что $\langle \dot{x}(x_0, t_1), p\rangle < 0$. Но вектор $\dot{x}$, как было отмечено, всегда лежит в подпространстве, ортогональном подпространству $S(x|\Pi)$, которому принадлежит вектор p.

Лемма 2. *Пусть преобразование* $\xi(y)$ *удовлетворяет условию* 1. *Тогда если в точке* x *выполнено у.р.о., то матрица* $\Gamma(x)$ *положительно определена.*

Доказательство. Покажем, что ранг матрицы $B(x) = J^\top(x) g_x^\top(x)$ равен m. Тогда из того, что $\Gamma(x) = B^\top(x)B(x)$, будет следовать, что матрица $\Gamma(x)$ неособая, положительно-полуопределенная. Если $x \in \operatorname{int}\Pi$, то это утверждение очевидно, поскольку матрица $J(x)$ неособая, а ранг матрицы $g_x^\top(x)$ в силу условия регулярности равен m.

Пусть теперь $x \in \operatorname{fr}\Pi$. Если ранг матрицы $B(x)$ меньше, чем m, то существует такой ненулевой вектор $z \in \mathbb{E}^n$, что $B(x)z = J^\top(x)g_x^\top(x)z = 0$. Но тогда согласно условию 1 вектор $p = g_x^\top(x)z$, не равный нулю, принадлежит пространству $S(x|\Pi)$. Поэтому векторы $g_x^i(x)$, $1 \le i \le m$, и вектор p оказываются линейно зависимы, что противоречит у.р.о. Таким образом, матрица $B(x)$ имеет полный ранг, равный m. $\square$

На основании утверждения леммы 2 получаем, что при выполнении у.р.о. во всех точках $x \in \Pi$ матрица $\Gamma(x)$ будет неособой. Поэтому зависимость $u(x)$ однозначным образом определяется из решения уравнения (2.6):

$$u(x) = \Gamma^{-1}(x)[\tau g(x) - g_x(x)G(x)f_x(x)].$$

После подстановки $u(x)$ в правую часть (2.5) метод (2.5), (2.6) можно переписать в проективной форме, аналогичной (1.9):

$$\frac{dx}{dt} = -J(x)[\pi(g_x(x)J(x))J^\top(x)f_x(x) + \tau(g_x(x)J(x))^+ g(x)]. \tag{2.10}$$

Определение 3. *Точка* $[x_*, u_*] \in \Pi \times \mathbb{E}^m$ *называется* **точкой Куна–Таккера** *для задачи* (2.1), *если*

$$L_x(x_*, u_*) \in K^*(x_*|\Pi), \qquad g(x_*) = 0. \tag{2.11}$$

В случае когда $L_x(x_*, u_*) \in S(x_*|\Pi)$ *и имеет место второе равенство* (2.11), *то* $[x_*, u_*]$ *называется* **слабой точкой Куна–Таккера**.

Лемма 3. *Пусть преобразование* $\xi(y)$ *удовлетворяет условию* 1. *Тогда точка* $x_* \in \Pi$, *в которой выполнено у.р.о., является стационарной для системы* (2.10) *в том и только том случае, если пара* $[x_*, u_*]$, *где* $u_* = u(x_*)$, *есть слабая точка Куна–Таккера для задачи* (2.1).

Доказательство. В стационарной точке выполнено

$$G(x_*)L_x(x_*, u_*) = 0,$$

т.е. $L_x(x_*, u_*) \in \ker G(x_*)$. Но матрицы $J^\top(x_*)$ и $G(x_*) = J(x_*)J^\top(x_*)$ имеют одинаковый ранг, равный m, и их нуль-пространства совпадают. Поэтому $L_x(x_*, u_*) \in \ker J^\top(x_*)$ и, следовательно, имеет место первое включение (2.11). Справедливость равенства $g(x_*) = 0$ вытекает из (2.6). Лемма доказана. $\square$

Будем полагать, что в точке Куна–Таккера $[x_*, u_*]$ выполнено *условие строгой дополняющей нежесткости* (у.с.д.н.), если

$$L_x(x_*, u_*) \in \mathrm{ri} K^*(x_*|\Pi), \tag{2.12}$$

где $\mathrm{ri}\, B$ — относительная внутренность множества B.

Предположим теперь, что все функции $f(x)$, $g^i(x)$, $1 \le i \le m$, дважды непрерывно дифференцируемы. Обозначим также $N(x) = \{h \in \mathbb{E}^n : g_x(x)J(x)h = 0\}$. Достаточные условия второго порядка из [17] для задачи (2.1) могут быть переформулированы следующим образом.

Теорема 1. *Пусть функция $\xi(y)$ удовлетворяет условию 1. Пусть, кроме того, в точке Куна–Таккера $[x_*, u_*]$ выполнено у.с.д.н. и*

$$\langle h, J^\top(x_*)L_{xx}(x_*, u_*)J(x_*)h\rangle > 0 \tag{2.13}$$

для любых таких $h \in N(x_*)$, что $J(x_*)h \ne 0_n$. Тогда x_* является точкой изолированного локального минимума в задаче (2.1).

Доказательство. Если x_* не есть точка изолированного локального минимума в задаче (2.1), то найдется последовательность допустимых точек $\{x_k\}$, сходящаяся к x_*, такая, что $f(x_k) \le f(x_*)$. Представим x_k в виде $x_k = x_* + \lambda_k s_k$, где $\|s_k\| = 1$, $\lambda_k > 0$, $\lambda_k \to 0$. Не умаляя общности, можно считать, что $s_k \to s_*$, $\|s_*\| = 1$. Так как $s_k \in K(x_*|\Pi)$ для всех $k > 0$, то $s_* \in \mathrm{cl}\, K(x_*|\Pi)$, где $\mathrm{cl}\, B$ — замыкание множества B. Имеют место соотношения

$$f(x_k) - f(x_*) = \lambda_k \langle f_x(x_* + \lambda_k\theta_k^0 s_k),\ s_k\rangle \le 0, \tag{2.14}$$

$$g^i(x_k) = \lambda_k\langle g_x^i(x_* + \lambda_k\theta_k^i s_k),\ s_k\rangle = 0, \qquad 1 \le i \le m. \tag{2.15}$$

Здесь $0 \le \theta_k^i \le 1$, $0 \le i \le m$. Умножая равенства (2.15) на u_*^i и складывая их с (2.14), получаем после деления на λ_k и перехода к пределу $\langle L_x(x_*, u_*), s_*\rangle \le 0$. Но согласно (2.12) $\langle L_x(x_*, u_*), s_*\rangle \ge 0$, причем $L_x(x_*, u_*) \ne 0_n$. Сопоставляя эти два неравенства, приходим к выводу, что $\langle L_x(x_*, u_*), s_*\rangle = 0$, т.е. вектор s_* ортогонален вектору $L_x(x_*, u_*) \in S(x_*|\Pi)$.

Покажем, что вектор s_* принадлежит ортогональному подпространству $S^\perp(x_*|\Pi)$. В самом деле, если это не так, то вектор s_* можно представить в виде $s_* = a + b$, где $a \in S(x_*|\Pi)$, $b \in S^\perp(x_*|\Pi)$, $a \ne 0$. Имеем $\langle L_x(x_*, u_*), s_*\rangle = \langle L_x(x_*, u_*), a\rangle = 0$. Таким образом, эти два ненулевых вектора расположены в одном и том же подпространстве и ортогональны друг другу. Но линейная оболочка $\mathrm{ri}\, K^*(x_*|\Pi)$ совпадает с линейной оболочкой самого конуса $\mathrm{K}^*(x_*|\Pi)$, равной $S(x_*|\Pi)$. Более того, если $p \in \mathrm{ri}\,\mathrm{K}^*(x_*|\Pi)$, то любой вектор из подпространства $S(x_*|\Pi)$, лежащий в некоторой окрестности вектора p, также будет принадлежать $\mathrm{ri}\,\mathrm{K}^*(x_*|\Pi)$. Поэтому в силу (2.12) можно указать вектор $q \in \mathrm{ri}\, K^*(x_*|\Pi)$, для которого $\langle s_*, q\rangle < 0$, что противоречит включению $s_* \in \mathrm{cl}\, K(x_*|\Pi)$. Таким образом, $s_* \in S^\perp(x_*|\Pi)$. Из условия 1 следует, что $S^\perp(x_*|\Pi)$ совпадает с пространством столбцов матрицы $J(x_*)$. Поэтому $s_* = J(x_*)h_*$ для некоторого ненулевого вектора $h_* \in \mathbb{E}^n$.

Разлагая теперь функции $f(x)$ и $g^i(x)$ в ряд Тейлора до второго члена включительно, получаем

$$f(x_k) - f(x_*) \;=\; \lambda_k\langle f_x(x_*), s_k\rangle + \frac{\lambda_k^2}{2}\langle s_k, f_{xx}(x_* + \lambda_k\theta_k^0 s_k)s_k\rangle \le 0, \tag{2.16}$$

$$g^i(x_k) \;=\; \lambda_k\langle g_x^i(x_*), s_k\rangle + \frac{\lambda_k^2}{2}\langle s_k, g_{xx}^i(x_* + \lambda_k\theta_k^i s_k)s_k\rangle = 0, \tag{2.17}$$

$0 \le \theta_k^i \le 1$, $0 \le i \le m$. Снова умножая равенства (2.17) на u_*^i и складывая их с (2.16), приходим к

$$\langle L_x(x_*, u_*), s_k\rangle + \frac{\lambda_k}{2}\left[\left\langle s_k, f_{xx}(x_* + \lambda_k\theta_k^0 s_k)s_k\right\rangle + \sum_{i=1}^m u_*^i\langle s_k, g_{xx}^i(x_* + \lambda_k\theta_k^i s_k)s_k\right\rangle\right] \le 0. \tag{2.18}$$

Из $s_k \in K(x_*|\Pi)$, $L_x(x_*, u_*) \in K^*(x_*|\Pi)$ следует, что $\langle L_x(x_*, u_*), s_k\rangle \geq 0$. Поэтому наряду с (2.18) имеют место неравенства

$$\langle s_k, f_{xx}(x_* + \lambda_k \theta_k^0 s_k) s_k\rangle + \sum_{i=1}^{m} u_*^i \langle s_k, g_{xx}^i(x_* + \lambda_k \theta_k^i s_k) s_k\rangle \leq 0.$$

Переходя в них к пределу, получаем $\langle s_*, L_{xx}(x_*, u_*) s_*\rangle \leq 0$, или

$$\langle h_*, J^\top(x_*) L_{xx}(x_*, u_*) J(x_*) h_*\rangle \leq 0, \tag{2.19}$$

причем $h_* \in N(x_*)$ и $||J(x_*)h_*|| \neq 0$. Неравенство (2.19) противоречит (2.10). Теорема доказана. □

В частном случае, когда множество Π совпадает со всем пространством $\mathbb{E}^n$, беря в качестве $\xi(y)$ тождественное преобразование $x = y$, получаем, что утверждение теоремы сводится к достаточным условиям изолированного локального минимума для задачи (1.1), приведенным в [17].

§ 3. Сходимость барьерно-проективных методов

Исследуем локальное поведение траекторий системы (2.10) в окрестности точки x_*. Предположим теперь, что функция $\xi(y)$ такова, что матрица $G(x)$ непрерывно дифференцируема. Пусть $p \in \mathbb{E}^n$. Обозначим через $G_x(x; p)$ квадратную матрицу порядка n, (i, j)-й элемент которой равен

$$G_x^{ij}(x; p) = \sum_{k=1}^{n} \frac{\partial G^{ik}(x)}{\partial x^j} p^k.$$

Наложим на преобразование $\xi(y)$ два дополнительных условия.

Условие 2. *В каждой точке $x \in \Pi$ для любого вектора $p \in \mathrm{ri} K^*(x|\Pi)$ матрица $G_x(x; p)$ является симметрической и ее нуль-пространство совпадает с $S^\perp(x|\Pi)$.*

Условие 3. *Если $x \in \Pi$, то $h^\top G_x(x; p) h > 0$ для любого ненулевого вектора $h \in \in S(x|\Pi)$ и для любого вектора $p \in \mathrm{ri} K^*(x|\Pi)$.*

При выполнении условий 1 и 2 матрицы $G(x)$ и $G_x(x; p)$ в каждой точке $x \in \Pi$ коммутируют между собой. Действительно, так как $J(x)a \in S^\perp(x|\Pi)$ для любого вектора $a \in \mathbb{E}^n$, то $G_x(x; p) J(x) a \equiv 0_n$. Но это означает, что матрица $G_x(x; p) J(x)$ нулевая, поэтому $G_x(x; p) G(x) = G_x(x; p) J(x) J^\top(x) = 0_{nn}$. С другой стороны в силу симметричности матрицы $G_x(x; p)$

$$G(x) G_x(x; p) = J(x) J^\top(x) G_x(x; p) = J(x) [G_x(x; p) J(x)]^\top = 0_{nn}.$$

Отметим, что в точках $x \in \mathrm{int}\Pi$ из-за того, что $\mathrm{K}^*(x|\Pi) = \{0\}$, сама матрица $G_x(x; p)$ всегда будет нулевой.

Теорема 2. *Пусть функция $\xi(y)$ удовлетворяет условиям $1 - 3$. Пусть, кроме того, в точке x_*, являющейся решением задачи (2.1), выполнено у.р.о. и достаточные условия второго порядка теоремы 1. Тогда для любого $\tau > 0$ точка x_* — экспоненциально устойчивое положение равновесия для системы (2.10).*

Доказательство. Уравнение в вариациях для системы (2.10) имеет вид

$$\delta \dot{x} = -Q(x_*, u_*) \delta x, \tag{3.1}$$

где

$$Q(x,u) = \tilde{M}(x)[G(x)L_{xx}(x,u) + G_x(x; L_x(x,u))] + \tau G(x)\tilde{P}(x), \tag{3.2}$$

$$\tilde{M}(x) = I_n - G(x)\tilde{P}(x), \qquad \tilde{P}(x) = g_x^\top(x)[g_x(x)G(x)g_x^\top(x)]^{-1}g_x(x).$$

Предположим для определенности, что точка x_* такова, что ранг матрицы $G(x_*)$ равен s, причем $s < n$. Так как $G(x_*)$ — симметрическая матрица, то можно указать такую ортогональную матрицу U, что $G(x_*) = UHU^\top$ и матрица H имеет вид

$$H = \begin{bmatrix} H^B & 0_{s,n-s} \\ 0_{n-s,s} & 0_{n-s,n-s} \end{bmatrix}.$$

Здесь H^B — диагональная матрица порядка s, диагональные элементы которой представляют собой отличные от нуля собственные значения матрицы $G(x_*)$. Так как $G(x_*)$ является неотрицательно-определенной матрицей, то все они строго положительны. Более того, учитывая симметричность матрицы $G_x(x_*; L_x(x_*, u_*))$ и коммутативность матриц $G(x_*)$ и $G_x(x_*; L_x(x_*, u_*))$, матрицу U можно подобрать таким образом, чтобы матрица $Y = U^\top G_x U$ также была диагональной. Поэтому матрица $Q(x_*, u_*)$ представима в виде

$$Q(x_*, u_*) = URU^\top, \qquad R = (I_n - HU^\top\tilde{P}U)(HU^\top L_{xx}U + Y) + \tau HU^\top\tilde{P}U$$

и, следовательно, ее собственные значения совпадают с собственными значениями матрицы R.

Пусть V и $V^\perp$ — соответственно нуль-пространство матрицы H и его ортогональное дополнение:

$$\begin{aligned} V &= \ker H = \{y \in \mathbb{E}^n : y^1 = \ldots = y^s = 0\}, \\ V^\perp &= \{y \in \mathbb{E}^n : y^{s+1} = \ldots = y^n = 0\}. \end{aligned}$$

Имеет место связь между подпространствами $S(x_*|\Pi)$, $S^\perp(x_*|\Pi)$ и V, $V^\perp$, а именно:

$$S(x_*|\Pi) = UV, \qquad S^\perp(x_*|\Pi) = UV^\perp.$$

Согласно условию 2 $G_x z = 0$, если $z \in S^\perp(x_*|\Pi)$. Поэтому $G_x U y = 0$ для всех $y \in V^\perp$. Отсюда следует, что матрица Y имеет вид

$$Y = [0_{n,s}, B], \qquad B^\top = [0_{n-s,s}, C],$$

где C — диагональная невырожденная матрица порядка $n - s$.

Обозначим через U^B и U^N подматрицы матрицы U, состоящие, соответственно, из первых s и последних $n - s$ ее столбцов. Обозначим также через H^B левую верхнюю квадратную подматрицу порядка s матрицы H. Положим

$$P^B = (U^B)^\top g_{x_B}^\top [g_{x_B} U^B H^B (U^B)^\top g_{x_B}^\top]^{-1} g_{x_B} U^B, \qquad L_{xx}^B = (U^B)^\top L_{xx} U^B.$$

Тогда матрицу R можно записать в следующем блочно-диагональном виде:

$$R = \begin{bmatrix} R_1 & R_3 \\ 0_{n-s,s} & R_2 \end{bmatrix}.$$

Здесь

$$R_1 = (I_s - H^B P^B) H^B L_{xx}^B + \tau H^B P^B, \qquad R_2 = (U^N)^\top G_x U^N.$$

Характеристическое уравнение для матрицы R расщепляется на два уравнения:

$$|R_1 - \lambda_i I_s| = 0, \qquad |R_2 - \lambda_j I_{n-s}| = 0, \qquad 1 \le i \le s, \qquad s+1 \le j \le n.$$

Найдем сначала решения второго уравнения. Если λ_j — собственное значение, соответствующее собственному вектору $z_j \in \mathbb{E}^{n-s}$, то имеют место равенства

$$(U^N)^\top G_x U^N z_j = \lambda_j z_j, \qquad s+1 \le j \le n,$$

или после умножения их справа на $z_j^\top$

$$z_j^\top (U^N)^\top G_x U^N z_j = \lambda_j ||z_j||^2, \qquad s+1 \le j \le n.$$

Так как векторы $h_j = U^N z_j \in S(x_*|\Pi)$, то согласно условию 2 числа

$$\lambda_j = \frac{z_j^\top (U^N)^\top G_x U^N z_j}{||z_j||^2} = \frac{h_j^\top G_x h_j}{||z_j||^2} \tag{3.3}$$

должны быть действительными и строго положительными. Обозначим

$$\hat{\lambda}_1 = \min_{s+1 \le j \le n} \lambda_j > 0.$$

Пусть Λ — квадратный корень из матрицы H, Λ^B — верхняя левая квадратная подматрица порядка s. Отыскание корней первого уравнения можно заменить нахождением собственных значений матрицы $W_1 = (\Lambda^B)^{-1} R_1 \Lambda^B$, подобной матрице R_1. После элементарных преобразований получаем

$$|W_1 - \lambda_i I_s| = |\hat{M} \hat{L}^B_{xx} + \tau \hat{P} - \lambda_i I_s| = 0, \qquad 1 \le i \le s,$$

где $\hat{M} = I_s - \hat{P}$, $\hat{P} = \Lambda^B P^B \Lambda^B$, $\hat{L}^B_{xx} = \Lambda^B L^B_{xx} \Lambda^B$.

Так как введенные матрицы $\hat{M}$ и $\hat{P}$ являются идемпотентными, для них $\hat{M} \times \hat{M} = \hat{M}$, $\hat{P} \times \hat{P} = \hat{P}$, $\hat{M} \times \hat{P} = 0$. Матрица $\hat{M}$ осуществляет проектирование любого s-мерного вектора на касательное многообразие:

$$\hat{K}(x_*) = \{\bar{x} \in \mathbb{E}^s : g_{x_B}(x_*) U^B \Lambda^B \bar{x} = 0_m\}.$$

Матрица $\hat{P}$ проектирует s-мерные векторы на ортогональное дополнение $\hat{K}^\perp(x_*)$ к этому пространству.

Пусть z_i — собственный вектор и λ_i — соответствующее собственное значение матрицы W_1; тогда

$$(\hat{M} \hat{L}^B_{xx} + \tau \hat{P}) z_i = \lambda_i z_i, \qquad z_i \in \mathbb{E}^s. \tag{3.4}$$

Если ненулевой собственный вектор z_i таков, что $||\hat{P} z_i|| \ne 0$, то, умножая (3.4) слева на матрицу $\hat{P}$, получаем $\lambda_i = \tau$. Если теперь $||\hat{P} z_i|| = 0$, т.е. $z_i \in \hat{K}(x_*)$, то, умножая (3.4) слева на $z_i^\top$, находим

$$\lambda_i = \frac{z_i^\top \Lambda^B L^B_{xx} \Lambda^B z_i}{||z_i||^2}. \tag{3.5}$$

Учтем теперь, что $U^B \Lambda^B z_i = U \Lambda h_i = J h_i$ для некоторого вектора $h_i \in \mathbb{E}^n$, у которого первые s компонент совпадают с соответствующими компонентами вектора z_i. Тогда (3.5) перепишется в виде

$$\lambda_i = \frac{h_i^\top J^\top L_{xx} J h_i}{||z_i||^2}. \tag{3.6}$$

Из $z_i \in \hat{K}(x_*)$, $z_i \neq 0_s$, следует, что $h_i \in N(x_*)$, $Jh_i \neq 0_n$. Поэтому на основании неравенства (2.13) заключаем, что всякому собственному вектору z_i матрицы W_1 из касательного многообразия $\hat{K}(x_*)$ соответствует положительное собственное значение λ_i и

$$\hat{\lambda}_2 = \min_{i \in \Delta(x_*)} \lambda_i > 0, \qquad \Delta(x_*) = \{i : z_i \in \hat{K}(x_*)\}.$$

Таким образом, собственные значения матрицы Q расщепляются на три группы:
1) $n - s$ корней (3.3),
2) k корней $\lambda = \tau$,
3) $s - k$ корней (3.6).
Если $\tau > 0$, то все собственные значения матрицы Q строго положительны и согласно теореме Ляпунова об асимптотической устойчивости по первому приближению положение равновесия $x = x_*$ локально экспоненциально устойчиво. $\square$

Из доказательства теоремы 2 вытекает следующая оценка скорости сходимости решений системы (2.10):

$$\lim_{t \to \infty} \frac{\ln |x^i(x_0, t) - x^i_*|}{t} \leq -\lambda_*, \qquad \lambda_* = \min[\tau, \hat{\lambda}_1, \hat{\lambda}_2], \qquad 1 \leq i \leq n. \tag{3.7}$$

Отметим, что если $x_0 \in X$, то траектории системы (2.5), (2.6) совпадают с траекториями системы (2.8), (2.9). Поэтому на основании теоремы 2 можно утверждать, что и метод (2.8), (2.9) локально экспоненциально сходится к точке x_* на допустимом множестве X. Для него сохраняется оценка (3.7), однако при этом λ_* должно быть заменено на $\lambda_* = \min[\hat{\lambda}_1, \hat{\lambda}_2]$. Заметим также, что если множество Π есть все пространство $\mathbb{E}^n$, то, беря в качестве $\xi(y)$ преобразование $x = y$, приходим из (2.5), (2.6) к методу (1.2). Для него также становятся справедливыми условия сходимости, приведенные в теореме 2.

Рассмотрим дискретный вариант метода (2.10); он может быть записан в виде

$$\begin{aligned} x_{k+1} &= x_k - \alpha_k G(x_k) L_x(x_k, u_k), & (3.8)\\ u_k &= \Gamma^{-1}(x_k)[\tau g(x_k) - g_x(x_k) G(x_k) f_x(x_k)], & (3.9) \end{aligned}$$

где $\alpha_k > 0$ — шаг интегрирования системы (2.10) по схеме Эйлера.

Теорема 3. *Пусть в задаче* (2.1) *выполнены условия теоремы* 2. *Тогда метод* (3.8), (3.9) *локально сходится к точке x_* со скоростью геометрической прогрессии, если шаг α_k постоянный, равный α, где*

$$0 < \alpha < 2/\lambda^*, \tag{3.10}$$

λ^* — максимальное собственное значение матрицы Q, задаваемой формулой (3.2).

Доказательство. Представим (3.8), (3.9) как метод простой итерации:

$$x_{k+1} = \Phi(x_k), \qquad \Phi(x) = x - \alpha G(x) L_x(x, u(x)). \tag{3.11}$$

Точка $x = x_*$ является неподвижной точкой оператора $\Phi(x)$. Согласно теореме Островского (см. [18]) достаточным условием линейной локальной сходимости метода (3.11) является требование, чтобы спектральный радиус ρ матрицы $\Phi_x(x_*)$ был меньше единицы. Рассмотрим характеристическое уравнение $|\Phi_x(x_*) - \chi I_n| = 0$. Несложно видеть, что $\Phi_x(x_*) = I_n - \alpha Q$, где матрица Q определяется формулой (3.2). Пусть λ — произвольное собственное значение матрицы Q; тогда соответствующее ему значение χ равно $1 - \alpha\lambda$. При доказательстве теоремы 2 было установлено, что все собственные значения матрицы Q

являются действительными положительными числами. Поэтому если шаг α удовлетворяет условию (3.10), то $|\chi| < 1$ и, следовательно, $\rho < 1$. $\square$

Теорема 3 дает достаточные условия локальной сходимости метода (3.8), (3.9). Скорость сходимости линейная, т.е. для любого $\varepsilon > 0$ и достаточно больших k справедлива оценка

$$||x_k - x_*|| \le C(\rho + \varepsilon)^k,$$

где $\rho = \max[|1 - \alpha\lambda_*|, |1 - \alpha\lambda^*|]$, λ_* — минимальное собственное значение матрицы Q, C — некоторая положительная константа.

Рассмотрим важный частный случай задачи (2.1), когда множество Π есть положительный ортант $\mathbb{E}^n_+$ пространства $\mathbb{E}^n$. Тогда преобразование координат можно строить в следующем сепарабельном виде: $\mathrm{x}^i = \xi^i(y^i)$, $1 \le i \le n$. Матрицы $J(x)$ и $G(x)$ для таких преобразований будут диагональными:

$$\begin{array}{rcl} J(x) &=& D(\gamma(x)) = \mathrm{diag}(\gamma^1(x^1), \ldots, \gamma^n(x^n)), \quad \gamma^i(t) = \xi^i((\xi^i)^{-1}(t)), \\ G(x) &=& D(\theta(x)) = \mathrm{diag}(\theta^1(x^1), \ldots, \theta^n(x^n)), \quad \theta^i(t) = [\gamma^i(t)]^2. \end{array}$$

Пусть $\sigma(x) = \{i : x^i = 0\}$ — множество активных индексов в точке $x \in \Pi$. Конус $K^*(x|\Pi)$ и подпространство $S(x|\Pi)$ для $\Pi = \mathbb{E}^n_+$ имеют вид

$$\begin{array}{rcl} K^*(x|\Pi) &=& \{z \in \mathbb{E}^n_+ : z^i = 0, \quad i \notin \sigma(x)\}, \\ S(x|\Pi) &=& \{z \in \mathbb{E}^n : z^i = 0, \quad i \notin \sigma(x)\}. \end{array}$$

Поэтому условие 1 в данном случае сводится к следующему: $\gamma^i(0) = 0$ и $\gamma^i(t) > 0$, если $t > 0$. Для того чтобы выполнялись условия 2 и 3, достаточно потребовать, чтобы функции $\theta^i(t)$, $1 \le i \le n$, были дифференцируемы и

$$\dot{\theta}^i(0) > 0. \tag{3.12}$$

В качестве простейших примеров преобразований данного вида можно указать следующие два:

$$x^i = (y^i)^2/4, \qquad x^i = e^{-y^i}, \qquad 1 \le i \le n.$$

Для них получаем соответственно $G(x) = D(x)$ и $G(x) = D^2(x)$. Условие (3.12) выполняется только для первого преобразования, для второго оно нарушено.

Рассмотрим еще один случай задачи (2.1), когда множество Π есть “параллелепипед”:

$$\Pi = \{x \in \mathbb{E}^n : a \le x \le b\}, \qquad a \in \mathbb{E}^n, \qquad b \in \mathbb{E}^n.$$

Для него также удобно строить преобразование $\xi(y)$ в сепарабельном виде, приводящем к диагональным матрицам $G(x) = D(\theta(x))$. Если, например, воспользоваться преобразованиями

$$x = \frac{1}{2}[a + b + (b - a)\sin y], \qquad x = \frac{1}{2}\left[a + b + \frac{2(b-a)}{\pi}\mathrm{arctg}\, y\right], \tag{3.13}$$

то получаем соответственно

$$\theta(x) = (b - x)(x - a), \qquad \theta(x) = \frac{(b-a)^2}{\pi^2}\cos^4\frac{\pi(2x - a - b)}{2(b-a)}.$$

Условие 3 сводится к требованиям, чтобы $\dot{\theta}(a^i) > 0$, $\dot{\theta}(b^i) < 0$. Оно выполнено только для первого преобразования (3.13).

§ 4. Задача с ограничениями типа неравенства

Рассмотрим задачу (2.1), в которой ограничения типа равенства заменены на неравенства:

$$\min_{x \in X} f(x), \qquad X = \{x \in \mathbb{E}^n : g(x) \leq 0_m, \quad x \in \Pi\}. \tag{4.1}$$

Введение дополнительных неотрицательных переменных позволяет свести эту задачу к виду (2.1) и применить для ее решения метод (2.5), (2.6), причем в окончательных численных схемах удается избежать уравнений для дополнительных искусственных переменных. Данный прием для случая, когда $\Pi = \mathbb{E}^n_+$, был исследован в [10, 12].

Возможен также другой подход к построению барьерно-проективных методов решения задачи (4.1), аналогичный рассмотренному в [19]. Составим модифицированную функцию Лагранжа:

$$M(x, u, \tau) = L(x, u) - \frac{1}{2\tau} \langle L_x(x, u), G(x) L_x(x, u) \rangle.$$

Зависимость $u(x)$, определяемая из (2.6), является решением параметрической задачи безусловной максимизации

$$\max_{u \in \mathbb{E}^m} M(x, u, \tau),$$

а метод (3.8), (3.9) может интерпретироваться как метод простой итерации для решения системы уравнений

$$G(x) L_x(x, u(x)) = 0_n. \tag{4.2}$$

Его обобщением на случай задачи (4.1) является следующий итеративный процесс:

$$u_k = \arg \max_{u \in \mathbb{E}^m_+} M(x_k, u, \tau), \qquad x_{k+1} = x_k - \alpha G(x_k) L_x(x_k, u_k), \tag{4.3}$$

где $\alpha > 0$ — шаг интегрирования системы.

Обозначим $\sigma_0(\text{x}) = \{1 \leq i \leq m : g^i(x) = 0\}$, $N_1(x) = \{h \in \mathbb{E}^n : g^i_x(x) J(x) h = 0, i \in \sigma_0(x)\}$. Для задачи (4.1) определения у.р.о. и у.с.д.н. могут быть переформулированы следующим образом.

Определение 4. *В точке $x \in \Pi$ выполнено у.р.о. для задачи* (4.1)*, если все векторы $g^i_x(x)$, $i \in \sigma_0(x)$, и произвольный ненулевой вектор $p \in S(x|\Pi)$ линейно независимы.*

Определение 5. *Точка $[x_*, u_*] \in \Pi \times \mathbb{E}^m_+$ является точкой Куна–Таккера для задачи* (4.1) *и в ней выполнено у.с.д.н., если*

$$L_x(x_*, u_*) \in \text{ri} K^*(x_*|\Pi), \qquad g(x_*) = L_u(x_*, u_*) \in -\text{ri} K^*(u_*|\mathbb{E}^m_+).$$

Аналогом теорем 1 и 2 являются два утверждения.

Теорема 4. *Пусть функция $\xi(y)$ удовлетворяет условию 1. Пусть, кроме того, в точке Куна–Таккера $[x_*, u_*]$ выполнено у.с.д.н. для задачи* (4.1) *и для любых $h \in N_1(x_*)$ таких, что $J(x_*) h \neq 0_n$, имеет место неравенство* (2.13)*. Тогда x_* является точкой изолированного локального минимума в задаче* (4.1).

Теорема 5. *Пусть функция $\xi(y)$ удовлетворяет условиям $1 - 3$. Пусть, кроме того, в точке x_*, являющейся решением задачи* (4.1)*, выполнено у.р.о. и достаточные условия второго порядка теоремы 4. Тогда для любого $\tau > 0$ и любого α, удовлетворяющего* *неравенству* (3.10), *итеративный процесс* (4.3) *локально сходится к* x_* *с линейной скоростью.*

Доказательство теоремы 5 основано на применении теоремы Островского и почти дословно повторяет доказательство теоремы 2. При этом матрица $\tilde{P}$, входящая в определение матрицы (3.2), составляется только из градиентов активных ограничений $g_x^i(x_*)$, $i \in \sigma_0(x_*)$.

СПИСОК ЛИТЕРАТУРЫ

1. *Rozen J.B.* The gradient projection method for nonlinear programming. Part I: Linear constraints //SIAM J. Appl. Math., 1960, v.8, N° 1, pp. 181–217; Part II: Nonlinear constraints // 1961, v.9. N° 4, pp. 514–532.

2. *Левитин Е.С., Поляк Б.Т.* Методы минимизации при наличии ограничений // ЖВМ и МФ, 1966, т.6, N° 5, с. 787–823.

3. *Демьянов В.Ф., Рубинов А.М.* Приближенные методы решения экстремальных задач. Л.: Изд-во ЛГУ, 1968.

4. *Антипин А.С.* Непрерывные и итеративные процессы с операторами проектирования // Вопр. кибернетики. Вычисл. вопр. анализа больших систем. М.: Наука, 1989, с. 1–43.

5. *Tanabe K.* A geometric method in nonlinear programming // J. Optimizat. Theory and Appl., 1980, v.30, N° 2, pp. 181–210.

6. *Дикин И.И.* Итеративное решение задач линейного и квадратичного программирования // Докл. АН СССР, 1967, т.174, N° 4, с. 745–747.

7. *Дикин И.И.* О сходимости одного итерационного процесса // Управляемые системы. Новосибирск, 1974, вып.12, с. 54–60.

8. *Евтушенко Ю.Г., Жадан В.Г.* Численные методы решения некоторых задач исследования операций // ЖВМ и МФ, 1973, т.13, N° 3, с. 583–597.

9. *Евтушенко Ю.Г.* Два численных метода решения задач нелинейного программирования // Докл. АН СССР, 1974, т.215, NN° 1, с. 38–40.

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

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

12. *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.

13. *Karmarkar N.* A new polinominal-time algorithm for linear programming // Combinatorica, 1984, N° 4, pp. 373–395.

14. *Barnes E.R.* A variation on Karmarkar's algorithm for solving linear programming problems // Math. Program., 1986, v.36, pp. 174–182.

15. *Vanderbei R.J., Meceton M.S., Freeman B.A.* On a modification of Karmarkar's linear programming algorithm // Algorithmica, 1986, v.1, pp. 395–407.

16. *Wei Zi-Luan.* An interior point method for linear programming // J. Comput. Math., 1987, Oct., pp. 342–350.

17. *Фиакко А, Мак-Кормик Г.* Нелинейное программирование. Методы последовательной безусловной минимизации. М.: Мир, 1972.

18. *Ортега Дж., Рейнболдт В.* Итерационные методы решения нелинейных систем уравнений со многими неизвестными. М.: Мир, 1975.

19. *Жадан В.Г.* Модифицированные функции Лагранжа в нелинейном программировании // ЖВМ и МФ, 1982, т.22, N° 2, с. 296–308.