ЖУРНАЛ
ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ ФИЗИКИ

Том 17 Июль 1977 Август № 4, с. 890–904

УДК 518:51:330.115

РЕЛАКСАЦИОННЫЙ МЕТОД РЕШЕНИЯ ЗАДАЧ НЕЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ

Ю.Г. ЕВТУШЕНКО, В.Г. ЖАДАН

117967 Москва, ул. Вавилова, 40, ВЦ РАН

(Пересмотрена 28 марта 2002 г.)

Предлагается релаксационный метод отыскания локальных экстремумов в общей задаче нелинейного программирования. Доказана сходимость, исследована скорость сходимости непрерывного и дискретного вариантов метода, дано обобщение на случай отыскания седловых точек. Приведены результаты численных расчетов.

Работа развивает подход, предложенный в [1, 2]. Иные варианты релаксационных методов изложены в [3, 4].

§1. Непрерывный вариант метода

Рассмотрим общую задачу нелинейного программирования. Требуется найти

$$\min_{x\in X} F(x), \ \ X = \{x \in \mathbb{E}^n : g(x) = 0_{e1}, \ \ h(x) \le 0_{c1}, \ \ 0_{n1} \le x\}. \tag{1.1}$$

Здесь $\mathbb{E}^i$ есть i-мерное евклидово пространство; непрерывно дифференцируемые функции $F(x)$, $g(x)$, $h(x)$ осуществляют отображения $F : \mathbb{E}^n \to \mathbb{E}^1$, $g : \mathbb{E}^n \to \mathbb{E}^e$, $h : \mathbb{E}^n \to \mathbb{E}^c$ соответственно. Символ 0_{ij} обозначает матрицу $i \times j$, все элементы которой равны нулю. Введем множество

$$X_0 = \{x \in \mathbb{E}^n : g(x) = 0_{e1}, \ \ h(x) < 0_{c1}, \ \ 0_{n1} < x\}.$$

Точки $x \in X_0$ условимся называть **внутренними точками** множества X, точки $x \in \in X\backslash X_0$ — **граничными**. В ряде мест удобно объединить ограничения типа равенств и неравенств одним символом $R = [g, h] = [R^1, \dots, R^m]$. Вектор-функция $R(x)$, таким образом, осуществляет отображение $\mathbb{E}^n \to \mathbb{E}^m$, где $m = e + c$. Пусть $R_x(x)$ — матрица $n \times m$, у которой (i, j)-й элемент есть $\partial R^j(x)/\partial x^i$. Символы $D(z)$, $D(z^{1/2})$ обозначают диагональные матрицы, у которых i-е диагональные элементы суть соответственно z^i и $(z^i)^{1/2}$; размеры этих матриц определяются размерностью вектора z. Индекс "$\top$" у векторов и матриц обозначает транспонирование.

Обозначим $\xi(R) = \{\xi(R^1), \dots, \xi(R^m)\}$, $\xi(x) = \{\xi(x^1), \dots, \xi(x^n)\}$, где функция скалярного аргумента $\xi(z)$ определена и непрерывна для всех значений $z \ge 0$, $\xi(z) > 0$ при $z > 0$ и $\xi(0) = 0$. Поэтому $\xi(-h^i(x)) > 0$, $\xi(x) > 0_{n_1}$ для любых $x \in X_0$. Можно положить, например, $\xi(z) = z$. Для численного решения задачи (1.1) предлагается искать предельные (при $t \to \infty$) точки решения задачи Коши для системы

$$\dot{x} = -D(\xi(x))[F_x(x) + R_x(x)v], \qquad x(0) = x_0 \in X_0. \tag{1.2}$$

Здесь точка над буквой обозначает дифференцирование по независимой переменной t, вектор $v \in \mathbb{E}^m$ определяется из решения следующей системы линейных уравнений:

$$G(x)v + R_x^\top(x)D(\xi(x))F_x(x) = 0_{m1}, \tag{1.3}$$

где $G(x) = R_x^\top(x)D(\xi(x))R_x(x) + D(\xi(-R(x)))$.

Если отсутствуют какие-либо ограничения помимо $x \geq 0_{n_1}$, метод (1.2) следует записать в виде

$$\dot{x} = -D(\xi(x))F_x(x). \tag{1.4}$$

Найдем v из (1.3) и подставим в правую часть системы (1.2); тогда (1.2) перепишется следующим образом:

$$\dot{x} = -M(x)F_x(x), \tag{1.5}$$

где $M(x) = D(\xi(x))[I_n - R_x[R_x^\top D(\xi(x))R_x + D(\xi(-R))]^{-1}R_x^\top D(\xi(x))]$, I_i — единичная матрица $i \times i$. Введем индексное множество

$$\sigma(x) = \{i : R^i(x) = 0, \ \ i = 1, 2, \dots, m\}.$$

Определение 1. *Ограничение $R(x) \leq 0_{m1}$ удовлетворяет условию регулярности в точке x, если вектор-функция $R(x)$ непрерывно дифференцируема в x и все векторы $R_x^i(x)$, где $i \in \sigma(x)$, линейно-независимы.*

Лемма 1. *Если в каждой точке $x \in X \backslash X_0$ выполнено условие регулярности ограничений $R(x) \leq 0_{m1}$, то матрица $Q(x) = R_x^\top(x)R_x(x) + D(\xi(-R(x)))$ неособая, положительно-полуопределенная для всех $x \in X$.*

Доказательство. Матрицу $Q(x)$ представим в виде произведения прямоугольной матрицы $B(x)$ размера $m \times (n+m)$ и транспонированной к ней $B^\top(x)$, причем матрица $B(x)$ состоит из двух матриц-блоков:

$$B(x) = ||R_x^\top(x) \ \ D(\xi^{1/2}(-R(x)))||, \qquad Q(x) = B(x)B^\top(x).$$

Лемма будет доказана, если удастся показать, что для любого $x \in X$ ранг матрицы $B(x)$ равен m, так как если ранг матрицы $B(x)$ максимален (равен m), то отсюда следует, что $Q(x)$ — неособая, положительно-полуопределенная матрица. Если отсутствуют ограничения типа равенства, то во всякой внутренней точке $x \in X_0$ ранг матрицы $B(x)$ равен m, так как в этом случае в качестве ненулевого минора матрицы $B(x)$ можно взять диагональную матрицу $D(\xi^{1/2}(-R(x)))$. Лемма также очевидна, если $R(x) = 0_{m1}$. В силу линейной независимости векторов $R_x^i(x)$ существует отличный от нуля минор порядка m матрицы $R_x^\top(x)$.

Пусть k компонент, $e < k < m$, вектор-функции $R(x)$ в точке $x \in X$ обращаются в нуль. Не нарушая общности, можно считать, что эти компоненты есть $R^1(x), \dots, R^k(x)$. Тогда значения функций $R^{k+1}(x), \dots, R^m(x)$ строго меньше нуля. Выделим в матрице

$$V_1(x) = \left\| \begin{array}{c} (R_x^1(x))^\top \\ \dots\dots\dots \\ (R_x^k(x))^\top \end{array} \right\|$$

размера $k \times n$ матрицу $C(x)$ размера $k \times k$ такую, что ее определитель, являющийся минором матрицы $V_1(x)$ порядка k, не равен нулю. Такой минор существует в силу условий регулярности. Определитель матрицы

$$V_2(x) = \left\| \begin{array}{c|cc} C(x) & \multicolumn{2}{c}{0_{k\times(m-k)}} \\ \hline & \xi^{1/2}(-R^{k+1}(x)) & 0 \\ 0_{(m-k)\times k} & \dots & \dots \\ & 0 & \xi^{1/2}(-R^m(x)) \end{array} \right\|$$

размером $m \times m$ не равен нулю. Но определитель матрицы $V_2(x)$ является одновременно минором m-го порядка матрицы $B(x)$. Поэтому ранг матрицы $B(x)$ максимальный, т.е. равен m. Лемма доказана. $\square$

Обозначим через e^i единичный вектор, у которого равна единице i-я координата.

Определение 2. *Ограничения $R(x) \le 0_{m1}$ и $0_{n_1} \le x$ удовлетворяют условию регулярности в точке x, если вектор-функция $R(x)$ непрерывно дифференцируема в x и все векторы $R^j_x(x)$, где $j \in \sigma(x)$, и векторы e^i такие, что $x^i = 0$, линейно-независимы.*

Из этого определения следует, что в случае, если ограничения удовлетворяют условию регулярности в точке x, число координат векторов $R(x)$ и x, которые одновременно обращаются в нуль, не превышает n.

Лемма 2. *Если в каждой точке $x \in X \backslash X_0$ выполнено условие регулярности ограничений $R(x) \le 0_{m1}$, $0_{n1} \le x$, то матрица $G(x)$ неособая, положительно-полуопределенная для всех $x \in X$.*

Доказательство. Покажем, что для любого $x \in X$ ранг матрицы

$$\Psi(x) = \| R^\top_x(x) D(\xi^{1/2}(x)) \quad D(\xi^{1/2}(-R(x))) \|$$

размера $m \times (n+m)$ равен m. Тогда из того, что $G(x) = \Psi(x)\Psi^\top(x)$, будет следовать, что матрица $G(x)$ неособая, положительно-полуопределенная. Утверждение леммы очевидно, если $x = 0_{n1}$, так как при этом обязательно $R(x) < 0_{m1}$ и $G(x) = D(\xi(-R))$. Пусть теперь $x \ne 0_{n1}$. Совокупность функций $R(x)$, $-x$ можно рассмотреть как одну вектор-функцию $\{R(x), -x\}$ и применить к ней лемму 1. Матрице B в этом случае будет соответствовать матрица B_1 размера $(m+n) \times (m+2n)$. Представим ее в блочном виде:

$$B_1 = \left\| \begin{array}{ccc} R^\top_x & D(\xi^{1/2}(-R)) & 0_{mn} \\ -I_n & 0_{nm} & D(\xi^{1/2}(x)) \end{array} \right\|.$$

Пусть среди координат вектора x есть s нулевых, $s > 0$. Будем считать, что вектор x можно представить как совокупность двух векторов $x = \{y, z\}$, где $y \ne 0_{k1}$, $y \in \mathbb{E}^k$, $z = 0_{s1}$, $k = n - s$. Аналогично, $R^\top_x = \{R^\top_y, R^\top_z\}$, причем $R^\top_y$, $R^\top_z$ — матрицы размеров, соответственно, $m \times k$ и $m \times s$. Матрицу B_1 запишем в виде

$$B_1 = \left\| \begin{array}{ccccc} R^\top_y & R^\top_z & D(\xi^{1/2}(-R)) & 0_{mk} & 0_{ms} \\ -I_k & 0_{ks} & 0_{km} & D(\xi^{1/2}(y)) & 0_{ks} \\ 0_{sk} & -I_s & 0_{sm} & 0_{sk} & 0_{ss} \end{array} \right\|.$$

Здесь $\xi^{1/2}(y) = [\xi^{1/2}(y^1), \dots, \xi^{1/2}(y^k)]$. Согласно лемме 1 ранг матрицы B_1 максимальный и равен $m+n$, поэтому в матрице B_1 можно выделить квадратную матрицу B_2 порядка

$n+m$, определитель которой есть ненулевой минор матрицы B_1. Так как в нижней строке-матрице матрицы B_1 стоят все нулевые матрицы, кроме $-I_s$, в матрице B_2 обязательно должен быть столбец

$$T = \left\| \begin{array}{c} R_z^\top \\ 0_{ks} \\ -I_s \end{array} \right\| \qquad \text{и} \qquad B_2 = \left\| \begin{array}{c|c} B_3 & \\ \hline 0_{sk} & T \end{array} \right\|.$$

Матрица B_3 размера $(m+k)\times(m+k)$ неособая. Это следует из формулы Фробениуса, так как $|B_2| = |-I_s|\,|B_3|$. Опуская в матрице B_1 столбец T, нижнюю матрицу-строку и правую нулевую матрицу-столбец, получим матрицу

$$B_4 = \left\| \begin{array}{ccc} R_y^\top & D(\xi^{1/2}(-R)) & 0_{mk} \\ -I_k & 0_{km} & D(\xi^{1/2}(y)) \end{array} \right\|$$

размера $(m+k)\times(m+2k)$, из элементов которой можно образовать матрицу B_3. Ранг B_4 поэтому максимальный. Если в B_4 опустить произвольные k строк, то ранг полученной матрицы будет также максимальным. Вычеркивая в B_4 нижнюю матрицу-строку и правую нулевую матрицу-столбец, получим матрицу ранга m

$$B_5 = \|R_y^\top \;\; D(\xi^{1/2}(-R))\|.$$

Умножим B_5 справа на неособую диагональную матрицу

$$B_6 = \left\| \begin{array}{cc} D(\xi^{1/2}(y)) & 0_{km} \\ 0_{mk} & I_m \end{array} \right\|$$

и добавим матрицу $R_z^\top D(\xi^{1/2}(z))$; обе эти операции не изменяют ее ранга. В результате получим матрицу

$$\|R_y^\top D(\xi^{1/2}(y)) \qquad R_z^\top D(\xi^{1/2}(z)) \qquad D(\xi^{1/2}(-R))\| = \Psi(x).$$

Если $x > 0$, т.е. $k = n$, $s = 0$, то в матрице B_1 исчезают вторая и последняя матрицы-столбцы и последняя матрица-строка. Тогда для того, чтобы получить $\Psi(x)$, следует в B_1 опустить нижнюю матрицу-строку, правую нулевую матрицу-столбец и оставшуюся матрицу B_5 умножить на B_6. Ранг матрицы $\Psi(x)$ равен m, отсюда следует утверждение леммы. $\square$

Из леммы 2 получаем, что если условия регулярности ограничений выполнены в каждой граничной точке X, то матрица $G(x)$ имеет обратную в любой точке множества X и правые части системы (1.2) определены всюду на X.

В дальнейшем будем говорить, что ограничения удовлетворяют условию регулярности всюду на X, если они выполнены в каждой граничной точке множества X.

Лемма 3. *Если выполнены условия леммы 2, тогда симметрическая матрица $M(x)$ положительно полуопределена для всех $x \in X$.*

Введем матрицы

$$\begin{aligned} K_1 &= -[R_x^\top D(\xi(x))R_x + D(\xi(-R))]^{-1} R_x^\top D(\xi(x)), \\ K &= \left\| \begin{array}{c} D(\xi^{1/2}(x))(I_n + R_x K_1) \\ D(\xi^{1/2}(-R))K_1 \end{array} \right\| \end{aligned}$$

размеров $m\times n$ и $(m+n)\times n$ соответственно. Доказательство леммы следует из представления $M = K^\top K$, справедливость которого проверяем непосредственными вычислениями.

Если x_* — локальное решение задачи (1.1), то при выполнении условия регулярности ограничений $R(x) \le 0_{m1}$ $0_{n1} \le x$ необходимо существование вектора $v_* \in \mathbb{E}^m$ такого, что

$$\begin{array}{ll} D(x_*)(F_x(x_*) + R_x(x_*)v_*) = 0_{n1}, & v_*^i \ge 0, \qquad 1+e \le i \le m, \\ D(R(x_*))v_* = 0_{m1}, & F_x(x_*) + R_x(x_*)v_* \ge 0_{n1}. \end{array} \tag{1.6}$$

В случае задач выпуклого программирования (ограничения $g(x)$ при этом линейные, функции $h(x)$ и $F(x)$ выпуклые) условия (1.6) являются достаточными условиями минимума в задаче (1.1).

Введем множество $Z = \{x : F(x) \le F(x_0),\ x \in X\}$. Всюду в дальнейшем будем считать, что для каждого $x_0 \in X$ система (1.2) определяет единственное решение $x(x_0, t)$.

Лемма 4. *Пусть всюду на компактном множестве X функции $F(x)$, $R(x)$ непрерывно дифференцируемы, выполнены условия регулярности ограничений $R(x) \le 0_{m1}$ и $0_{n1} \le$ $\le x$. Тогда решения $x(x_0,t)$ системы (1.2) продолжимы при $t \to \infty$, множества X, Z положительно-инвариантные (т.е. из $x_0 \in X$ следует $x(x_0,t) \in X$ и из $x_0 \in Z$ следует $x(x_0,t) \in Z$ для всех $t \ge 0$).*

Доказательство. Вычислим первые производные функций $R(x)$ и $F(x)$ в силу системы (1.2), после ряда преобразований получим

$$\dot R = D(v)\xi(-R), \qquad \dot F = -\|D(\xi^{1/2}(x))H_x(x,v)\|_n^2 - \|D(\xi^{1/2}(-R))v\|_m^2. \tag{1.7}$$

Здесь $H_x(x,v) = F_x(x) + R_x(x)v$, $\|z\|_i^2 = z^\top z$, i — размерность вектора z.

Решение (1.2) существует по крайней мере для тех значений t, для которых $x(x_0,t) \in X$. Покажем, что решение $x(x_0,t)$ не выйдет из множества X ни при каких $t \ge 0$. Допустим, что $R^j(x(x_0,t)) > 0$ для некоторого $t > 0$. Тогда найдется момент t_1 такой, что $R^j(x(x_0,t_1)) = 0$ и $\dot R^j(x(x_0,t_1)) > 0$, что противоречит (1.7), так как $\|v(t_1)\|_m < \infty$. Аналогично, из вида системы (1.2) следует, что $x(x_0,t) \ge 0_{n1}$. Введенные выше функции $\xi(x)$, $\xi(-R)$ играют, таким образом, роль “барьеров”, не давая возможности траектории $x(x_0,t)$ пройти через поверхности $x = 0_{n1}$, $R(x) = 0_{m1}$. Траектория $x(x_0,t)$ может подойти к граничным точкам X лишь при $t \to \infty$. Если начальная точка x_0 является граничной, то вся траектория системы (1.2) принадлежит границе. Функции $g^i(x)$, задающие ограничения типа равенства, являются интегралами системы (1.2). Поэтому в силу ограниченности множества X решения системы (1.2) продолжимы при $t \to \infty$, множество X положительно-инвариантное относительно (1.2). Отсюда и из (1.7) следует положительная инвариантность множества Z. $\square$

Обозначим через $\bar x$ точки из X, в которых правые части системы (1.2) обращаются в нуль. Такие точки назовем *стационарными*. Соответствующие значения $v(\bar x)$ обозначим через $\bar v$. В точках $\bar x$ выполнены условия

$$\begin{array}{l} D(\xi(\bar x))H_x(\bar x,\bar v) = 0_{n1}, \qquad D(\xi(-R(\bar x)))\bar v = 0_{m1}, \\ R_x^\top(\bar x)D(\xi(\bar x))R_x(\bar x)\bar v + R_x^\top(\bar x)D(\xi(\bar x))F_x(\bar x) = 0_{n1}. \end{array} \tag{1.8}$$

Всякая точка x_*, являющаяся локальным решением задачи (1.1), будет стационарной. Действительно, если это не так, то, взяв x_* в качестве начальной точки для системы (1.2), мы получили бы, что решение $x(x_*,t) \in X$ и $F(x(x_*,t)) < F(x_*)$ для $t > 0$, так как $dF(x_*)/dt < 0$. Но это противоречит условию локального минимума функции $F(x)$.

Теорема 1. *Пусть выполнены условия леммы 4, все стационарные точки из множества X изолированы. Тогда для любых нестационарных точек $x_0 \in X_0$ решение $x(x_0, t)$ системы (1.2) сходится к допустимой стационарной точке, в которой выполнены необходимые (в случае задач выпуклого программирования — и достаточные) условия минимума (1.6).*

Доказательство. Пусть x_0 — произвольная точка из X_0, $x(x_0, t)$ — решение задачи Коши (1.2). В силу компактности X множество ω-предельных точек ω для решения $x(x_0, t)$ не пусто. Покажем, что ω принадлежит множеству допустимых стационарных точек. Так как $F(x)$ ограничена снизу на X и $F(x(x_0, t))$ — монотонно убывающая функция, то, согласно [5, 6], все точки из ω лежат на одной и той же поверхности уровня функции F. Пусть $\tilde{x} \in \omega$. Проведем через $\tilde{x}$ траекторию $x(\tilde{x}, t)$. Любая ее точка также принадлежит ω, поэтому $\dot{F}(x(\tilde{x}, t)) \equiv 0$ и, значит, $\dot{F}(\tilde{x}) = 0$. Но из (1.7) видно, что это возможно только тогда, когда $\tilde{x}$ — стационарная точка для системы (1.2). Отсюда, поскольку все стационарные точки из X изолированы, получим, что ω состоит из единственной допустимой стационарной точки $\tilde{x}$, к которой $x(x_0, t)$ приближается при $t \to \infty$.

При каждом $x = x(x_0, t)$ из (1.3) можно определить $\varphi^j(t) = v^j(x(x_0, t))$ и вычислить $\psi^i(t) = H_x^i(x(x_0), t), v(x(x_0, t)))$, $j = 1, 2, \ldots, m$, $i = 1, 2, \ldots, n$. Эти функции непрерывны, поэтому из существования предела $x_1 = \lim\limits_{t\to\infty} x(x_0, t)$ следует существование пределов

$$\bar{\varphi}^j = \lim_{t\to\infty} \varphi^j(t), \qquad \bar{\psi}^i = \lim_{t\to\infty} \psi^i(t).$$

Покажем, что при $j = e+1, \ldots, m$, $i = 1, \ldots, n$ они неотрицательны. В предельной точке x_1 выполнены условия (1.8) и $\bar{\psi}^i = 0$, если $x_1^i > 0$. Рассмотрим теперь случай, когда $x_1^i = 0$. Из системы (1.2) получаем

$$x^i(x_0, t) = x_0^i \exp(-\Phi^i(t)), \qquad \Phi_i(t) = \int_0^t \frac{\xi(x^i(x_0, \tau))}{x^i(x_0, \tau)} \psi^i(\tau) d\tau.$$

Предположим, что $\bar{\psi}^i < 0$; тогда можно указать такое $\bar{t} > 0$, что $\psi^i(t) < 0$ для всех $t > \bar{t}$ и, следовательно, $\Phi^i(t) < \Phi^i(\bar{t})$ для тех же значений t. Но это противоречит условию $x_1^i = 0$. Итак, $\bar{\psi}^i \geq 0$, $i = 1, 2, \ldots, n$. Аналогично с помощью (1.7), (1.8) устанавливается, что $\bar{\varphi}^j \geq$ ≥ 0, $j = e+1, e+2, \ldots, m$. Отсюда и из (1.8) следует, что в x_1 выполнены необходимые условия (1.6). Теорема доказана. $\square$

Легко видеть, что условия теоремы можно ослабить, потребовав, чтобы все они выполнялись на множестве Z (а не на X).

В случае задач выпуклого программирования требование единственности решения задачи (1.1), по-видимому, несущественно. Строгое доказательство этого факта удалось получить лишь в одном частном случае. Имеет место следующая теорема.

Теорема 2. *Пусть $F(x)$ — выпуклая, непрерывно дифференцируемая функция, $g(x)$, $h(x)$ линейные, всюду на компактном множестве X выполнены условия регулярности ограничений $R(x) \leq 0_{m1}$, $0_{n1} \leq x$. Тогда решения системы (1.2) сходятся к множеству решений задачи (1.1) при любых $x_0 \in X_0$.*

Метод (1.2) допускает большой произвол в выборе функции $\xi(z)$. Каждое из ограничений может быть учтено с помощью специальной функции. В тех случаях, когда в задаче (1.1) отсутствует ограничение $x \geq 0_{n1}$, в формулах (1.2) – (1.8) вместо матрицы $D(\xi(x))$ следует писать единичную матрицу. Если вместо условия $x \geq 0_{n1}$ наложено ограничение $x \geq a$, то в упомянутых формулах вместо $D(\xi(x))$ следует писать $D(\xi(x-a))$.

Если ограничения имеют вид $a^i \le x^i \le b^i$ или $c^j \le h^j(x) \le d^j$, тогда вводятся две вектор-функции $\xi_1(x)$, $\xi_2(h(x))$, у которых i-я и j-я координаты суть, например, $\xi_1^i(x) = (x^i - a^i)(b^i - x^i)$, $\xi_2^j(h(x)) = (h^j(x) - c^j)(d^j - h^j(x))$. Системы (1.2) и (1.3) имеют вид

$$\begin{aligned}&\dot{x} = -D(\xi_1(x))(F_x + R_x v),\\ &(R_x^\top D(\xi_1(x))R_x + D(\xi_2(R(x))))v + R_x^\top D(\xi_1(x))F_x = 0_{n1}.\end{aligned}$$

Таким образом, ограничения такого типа не повышают порядок линейной системы (1.3).

Если в задаче (1.1) отсутствуют ограничения типа неравенств ($m = e$, $R(x) = g(x)$, $R_x(x) = g_x(x)$, $X_0 = \emptyset$), то (1.2) совпадает с методом, описанным в [7, гл.II, §3]. Матрица M в этом случае имеет вид

$$M(x) = I_n - g_x(x)(g_x^\top(x)g(x))^{-1}g_x^\top(x).$$

Легко видеть, что $M(x)g_x(x) = 0_{n1}$, $M^\top M = M$, следовательно, $M(x)$ проектирует любой вектор $x \in \mathbb{E}^n$ на касательное многообразие ко множеству X в точке x, т.е. на ортогональное дополнение подпространства, порожденного векторами $g_x^1(x), \ldots, g_x^e(x)$. Условие стационарности $M(x)F_x(x) = 0_{n1}$ означает, что проекция вектора $F_x(x)$ на касательное многообразие равна нулю (необходимое условие экстремума). Вектор $\dot{x}$ лежит в касательном многообразии, благодаря чему $g^i(x)$ являются интегралами системы (1.2).

В методе (1.2) матрица $M(x)$ в граничных точках X проектирует вектор $F_x(x)$ на касательное многообразие ко множеству

$$Y(x) = \{z \in \mathbb{E}^n : R^i(z) = 0,\ \ i \in \sigma(x),\ \ z^j = 0,\ \text{ если }\ x^j = 0\},$$

не давая возможности траектории $x(x_0, t)$ пересечь границу множества X. В тех точках x, где некоторые $R^i(x) < 0$, проекция $M(x)F_x(x)$ на $R_x^i(x)$ согласно (1.7) равна $\xi(-R^i(x))v^i$. В численных расчетах обычно берется $\xi(-R^i(x)) = -R^i(x)$. Таким образом, скорость движения траекторий в направлении границы $R^i(x) = 0$ стремится к нулю при подходе к ней. Различные "барьерные" функции $\xi(-R)$ приводят к разным характерам изменения этой скорости. Введение функций $\xi(x)$, $\xi(-R)$ существенно упрощает численные расчеты по сравнению с методом проекции градиента [3], автоматически изменяя направление вектора $\dot{x}(x_0, t)$ вблизи границы.

Вдали от гиперповерхности $h^i(x) = 0$, когда $h^i(x) \ll 0$, можно не опасаться, что траектория $x(x_0, t)$ на небольшом интервале $(t, t + \delta)$ пересечет ее, и в формуле для определения $M(x)$ функцию $h^i(x)$ и ее производную можно опустить, вводя их в рассмотрение, делая ограничение активным лишь тогда, когда $-\varepsilon < h^i(x(x_0, t)) < 0$, где $\varepsilon > 0$ выбирается в зависимости от шага интегрирования системы (1.2). Такой прием позволяет понизить порядок системы (1.3). Вместе с тем введение барьерных функций $\xi(-R)$ может привести к тому, что $h^i(x(x_0, t)) \equiv 0$ при $h^i(x_0) = 0$, либо к тому, что $|h^i(x(x_0, t))|$ становится очень малой, несмотря на то, что $\dot{h}^i < 0$. Чтобы устранить этот недостаток, опустим в формулах для M функции h^i, h_x^i и вычислим в силу новой системы производную $\dot{h}^i = (h_x^i)^\top \dot{x}$. Если оказалось, что производная отрицательная, то продолжаем движение вдоль траектории этой системы. Другими словами, устраняя "барьер", мы проверяем, можно ли это сделать, не нарушив ограничений. Особенно просто эта процедура выглядит для случая, когда $m = 0$ и в (1.1) есть только ограничения $x > 0_{n1}$; система (1.4) заменяется следующей:

$$\dot{x} = \begin{cases} -F_{x^i}, & \text{если }\ x^i > \varepsilon > 0 \text{ или } 0 \le x^i \le \varepsilon \text{ и } F_{x^i} = 0,\\ -x^i F_{x^i}/\varepsilon, & \text{если }\ 0 \le x^i \le \varepsilon \text{ и } F_{x^i} > 0.\end{cases}$$

Здесь $i = 1, 2, \ldots, n$. Если F_x непрерывна, то и правые части системы также непрерывны и в ней отсутствуют скользящие режимы.

В некоторых частных случаях (1.2) переходит в методы, предложенные в [1, 2].

Рассмотрим задачу линейного программирования, заданную в стандартной форме: найти

$$\min_{x \in X} c^\top x, \qquad X = \{x : Ax = b, \;\; x \geq 0_{n1}\}, \tag{1.9}$$

где $x, c \in \mathbb{E}^n$, $b \in \mathbb{E}^m$, A — матрица $m \times n$. Двойственная к (1.9) задача состоит в отыскании

$$\max_{p \in P} b^\top p, \qquad P = \{p \in \mathbb{E}^m : A^\top p \leq c\}.$$

Положив $\xi(z) = z$, получим, что метод (1.2) для решения прямой задачи приводит к системе

$$\dot{x} = D(x)(A^\top p - c), \;\text{ где }\; AD(x)A^\top p = AD(x)c, \quad x(0) = x_0. \tag{1.10}$$

В этом случае $c^\top \dot{x} = -\|D(x^{1/2})(c - A^\top p)\|_n^2 \leq 0$, если $x_0 > 0_{n1}$, $Ax_0 = b$. Аналогично, для двойственной задачи

$$\begin{aligned} &\dot{p} = b - Ax, \;\text{ где }\; [A^\top A - D(x)(A^\top p - c)]x = A^\top b, \quad p(0) = p_0, \\ &b^\top \dot{p} = ||b - Ax||_m^2 + x^\top D(x)(c - A^\top p) \geq 0, \end{aligned} \tag{1.11}$$

если $A^\top p_0 \leq c$. Релаксационный метод решения прямой задачи (1.9) оказался весьма эффективным при решении задач большой размерности, но с малым числом ограничений ($n \gg m$). Аналогично, метод (1.11) удобен, если $n \ll m$. Методы (1.10) и (1.11) несущественно изменяются в случае задач квадратичного программирования.

Для работы алгоритма важно, чтобы начальная точка $x_0 \in X_0$. Если такая точка неизвестна, ее можно найти, используя предложенный алгоритм следующим образом. Пусть

$$\begin{array}{rclcrcl} W^1(x) & = & \{i : h^i(x) < 0, \quad 1 \leq i \leq c\}, & \qquad & W^2(x) & = & \{i : h^i(x) \geq 0, \quad 1 \leq i \leq c\}, \\ W^3(x) & = & \{i : |g^i(x)| \leq \varepsilon, \;\; 1 \leq i \leq e\}, & \qquad & W^4(x) & = & \{i : |g^i(x)| \geq \varepsilon, \;\; 1 \leq i \leq e\}. \end{array}$$

Здесь ε — допустимая точность выполнения условия типа равенства. Решаем задачу о минимуме

$$f(x) = \sum_{i \in W^2(x)} h^i(x) + \sum_{j \in W^4(x)} [g^i(x)]^2$$

при наличии ограничений $x \in W^5$, где

$$W^5 = \{x : x \geq 0_{n1}, \;\; h^i(x) \leq 0, \;\; i \in W^1(x), \;\; g^j(x) = \text{const}, \;\; j \in W^3(x)\}.$$

Если в процессе расчетов получено, что какие-нибудь индексы i при некоторых значениях x не принадлежат множеству $W_2(x)$, тогда рассматриваем функции $h^i(x)$ как ограничения типа неравенства и исключаем их из формулы для $f(x)$. Аналогично поступаем с ограничениями типа равенств, включая в выражение для $f(x)$ те из них, у которых $|g^i(x)| \geq \varepsilon$, и исключая их, как только нарушится это условие. Процесс продолжаем до тех пор, пока не получим допустимую точку.

Для решения задачи (1.1) можно применить метод Ньютона. Тогда для случая $\xi(z) = z$ получим следующую систему из $n + m$ уравнений:

$$\begin{array}{rcl} \dot{x} & = & -U^{-1}D(x)(H_x + R_x\dot{v}), \qquad U = D(H_x) + D(x)H_{xx}, \\ \dot{v} & = & -[R_x^\top U^{-1}D(x)R_x]^{-1}[R_x^\top U^{-1}D(x)H_x - D(R)v]. \end{array}$$

Эта система составлена таким образом, что

$$D(x(t))H_x(x(t), v(t)) = D(x_0)H_x(x_0, v(x_0))e^{-t}, \qquad \dot{R} = -D(v)R.$$

Система довольно громоздкая, и мы не будем здесь останавливаться на ее исследовании.

§2. Оценка скорости сходимости

Метод (1.2) можно использовать для нахождения седловых точек. В этом случае, однако, схема доказательства сходимости из § 1 оказывается неприменимой, так как она существенно опирается па монотонное изменение функции $F(x(x_0, t))$. В задачах отыскания седловых точек этого свойства нет. Поэтому обоснование приходится строить исходя из других рассуждений. Ниже приведем другое доказательство сходимости (1.2), которое обобщается на случай отыскания седловых точек и позволяет оценить скорость сходимости метода. Описание подхода проведем конспективно, приводя теоремы и давая лишь идею их доказательства. Для простоты изложения будем считать, что ограничения $x \geq 0_{n1}$ в задаче (1.1) отсутствуют и $\xi(z) = z$. Говоря о задаче (1.1) и о формулах § 1, мы имеем в виду этот случай. Как указывалось выше, всюду следует в формулах § 1 положить $D(\xi(x)) = I_n, D(\xi(-h)) = -D(h)$.

Для точек $x \in X$, $v \in \mathbb{E}^m$ можно определить условие **S**: для любого ненулевого вектора $z \in \mathbb{E}^{n+c}$ такого, что

$$N(x)z = 0_{m1}, \tag{2.1}$$

справедливо неравенство

$$z^\top H_{zz}(x,v)z > 0.$$

Здесь

$$N(x) = \left\| \begin{array}{cc} g_x^\top(x) & 0_{mc} \\ h_x^\top(x) & D([-2h(x)]^{1/2}) \end{array} \right\|,$$

$$H_{zz}(x,v) = \left\| \begin{array}{cc} H_{xx}(x,v) & 0_{nc} \\ 0_{cn} & D(\tilde{v}) \end{array} \right\|$$

суть матрицы, соответственно, размера $m \times (n+c)$ и порядка $(n+c) \times (n+c)$,

$$H_{xx}(x,v) = F_{xx}(x) + \sum_{i=1}^{m} v^i R_{xx}^i(x)$$

есть матрица порядка $n \times n$, $\tilde{v} = [v^{e+1}, v^{e+2}, \ldots, v^m]$.

Сформулируем достаточные условия минимума в задаче (1.1).

Теорема 3. *Для того чтобы допустимая точка x_* была локальным, изолированным минимумом задачи* (1.1), *где $F(x)$, $R(x)$ — дважды непрерывно дифференцируемые функции, достаточно, чтобы существовал вектор $v_* \in \mathbb{E}^m$ такой, что в точке x_* выполняются условия стационарности*

$$H_x(x_*, v_*) = F_x(x_*) + R_x(v_*)v_* = 0_{n1}, \qquad D(x_*)R(x_*) = 0_{m1}$$

и в точке (x_*, v_*) имеет место условие **S**.

Заметим, что из условий теоремы следует, что $\tilde{v}_* \geq 0_{m-e,1}$ и в точке x_* выполнено условие строгой дополняющей нежесткости. Введем вектор $y \in \mathbb{E}^c$ соотношением

$$2h^i(x) + (y^i)^2 = 0, \qquad i = 1, 2, \ldots, c. \tag{2.2}$$

Задачу (1.1) заменим следующей: найти $\min F(z)$ по $z = [x, y]$, $x \in \mathbb{E}^n$, $y \in \mathbb{E}^c$, при наличии ограничений типа равенств $g(x) = 0_{e1}$ и (2.2). Такой прием широко применялся в ряде работ (см., например, [8, 9]). Дальнейшие рассуждения проводятся аналогично доказательству теоремы 4 из [10].

Определение 3. *Решение $x(t)$ системы (1.2) называется условно асимптотически устойчивым в окрестности точки x_*, если для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что для всякого решения $x(x_0, t)$, подчиненного условию $\|x_0 - x_*\|_n < \delta$ для всех $t > 0$, выполнены условия*

$$||x(x_0,t) - x_*||_n < \varepsilon, \qquad \lim_{t\to\infty} x(x_0,t) = x_*.$$

Теорема 4. *Пусть в точке x_* выполнены достаточные условия минимума теоремы 3, ограничения на X удовлетворяют условию регулярности; тогда решение $x(x_0, t)$ системы (1.2) условно асимптотически устойчиво в окрестности точки x_*; если, кроме того,*

$$z^\top H_{zz}(x_*, v_*)z \geq \gamma||z||^2_{n+c}\gamma > 0 \quad \textit{для всех } \ z, \tag{2.3}$$

$N(x_*)z = 0_{m1}$, *то имеет место следующая оценка скорости сходимости метода (1.2):*

$$\begin{array}{rcl} 0 & \leq & -\dot{F}(x(x_0,t)) = \|H_x(x(x_0,t), v(t))\|^2_n \\ & + & ||D([-R(x(x_0,t))]^{1/2})v(t)||^2_m \leq [||H_x(x_0, v(x_0))||^2_n \\ & + & ||D([-R(x_0)]^{1/2})v(x_0)||^2_m]e^{-\gamma t}. \end{array} \tag{2.4}$$

Доказательство. Составим неотрицательную функцию

$$\varphi(x,v) = \frac{1}{2}\, H_z^\top(x,v)H_z(x,v).$$

Здесь $H_z^\top(x,v) = [H_x^\top(x,v), \tilde{v}^\top D([-h(x)]^{1/2})]$ — матрица-строка $1 \times (n+c)$.

Функция $v(t)$, определяемая из (1.3), дифференцируема. Поэтому ее можно продифференцировать в силу системы (1.2). Учитывая, что $\dot{R} = -R_x^\top H_x = -D(v)R$, получим $\dot{\varphi} \leq$ $\leq H_z^\top H_{zz} H_z$.

Воспользуемся свойством $\mathbf{S}$, положим $z = [H_x(x,v), D([-h(x)]^{1/2}\tilde{v}] \in \mathbb{E}^{n+c}$; тогда условие (2.1) будет выполнено при любых $t \geq 0$. Квадратичная форма, стоящая в формуле для $\dot{\varphi}$, строго отрицательна, если элементы матриц H_{xx} и $D(\tilde{v})$ вычислены в точке (x_*, v_*), но в силу непрерывности H_{xx} и $D(\tilde{v})$ это свойство будет иметь место, если (x,v) лежат в некоторой окрестности точки (x_*, v_*). Поэтому $\dot{\varphi} < 0$ и $\dot{\varphi} = 0$ лишь в стационарной точке (x_*, v_*), причем x_* — локальное изолированное решение задачи (1.1). На основании (2.3) получаем $\varphi(t) \leq \varphi(0)e^{-\gamma t}$. Тогда из (1.7) следует (2.4). Теорема 4 доказана. □

§3. Дискретный вариант метода

Интегрируя (1.5) по схеме Эйлера, получим

$$x_{s+1} = x_s - \alpha_s M(x_s) F_x(x_s), \tag{3.1}$$

где $0 < \alpha_s$ — шаги интегрирования.

Если выполнены условия леммы 2, то на X можно определить максимальное значение нормы матрицы M соотношением

$$\lambda = \max_{x\,\in\, X}\ \max_{y \in \mathbb{E}^n} \frac{y^\top M(x) y}{||y||^2} < \infty.$$

Обозначим

$$\mu = \max_i \max_{x\in X} \frac{\partial H(x,v)}{\partial x^i}, \qquad \nu = \max_j \max_{x\in X} v^j(x);$$

здесь $i = 1,2,\dots,n$, $j = 1,2,\dots,m$, $v^j(x)$ определяется из (1.3). В дальнейшем будем писать $v_s^j = v^j(x_s)$.

Функция $F_x(x)$ удовлетворяет условию Липшица на X с константой L, если для любых $x_1 \in X$, $x_2 \in X$ справедливо

$$\|F_x(x_1) - F_x(x_2)\|_n \le L\|x_1 - x_2\|_n. \tag{3.2}$$

Считаем для простоты, что $\xi(z) = z$, α_s постоянно.

Теорема 5. *Если выполнены условия теоремы* 1, *условие* (3.2), *функции* $R(x)$ *линейные, тогда при* $0 < \alpha_s < \min(1/\mu, 1/\nu, 2/\lambda L)$ *и при любых нестационарных начальных точках* $x_0 \in X_0$ *последовательность* $\{x_s\}$ *сходится к допустимой стационарной точке, в которой выполнены необходимые (в случае задач выпуклого программирования — и достаточные) условия минимума* (1.6), *причем* $x_s \in X_0$, $F(x_{s+1}) \le F(x_s)$ *для* $s = 0,1,\dots$. *Если, кроме того,* $\gamma\|z\|_{n+c}^2 \le z^\top H_{zz}(x,v(x))z \le \Gamma\|z\|_{n+c}^2$ *при всех* $x \in X$, $z \in \mathbb{E}^{n+c}$, *тогда для любой начальной точки* $x_0 \in X_0$ *будет*

$$\begin{aligned} &\|H_x(x_s,v_s)\|_n^2 + \|D([-R(x_s)]^{1/2})v_s\|_m^2 \le \\ \le\quad &[\|H_x(x_0,v_0)\|_n^2 + \|D([-R(x_0)]^{1/2})v_0\|_m^2][1-\alpha_s\gamma+\alpha_s^2\Gamma^2]^s, \end{aligned} \tag{3.3}$$

последовательность $x_s \to x_*$, *где* x_* *— единственное решение задачи* (1.1).

Доказательство. Из линейности функций $R(x)$ следует формула

$$R^i(x_{s+1}) = R^i(x_s)[1-\alpha_s v_s^i], \qquad i = 1,2,\dots,m. \tag{3.4}$$

Учитывая, что $R^i(x_0) \le 0$, $|v_s^i| < \nu$, получим, что $R^i(x_{s+1}) \le R^i(x_0)(1-\alpha_s\nu)^s$, если $\alpha_s < 1/\nu$. Причем для ограничений $R^i(x)$ типа равенства, $i = 1,2,\dots,e$, из $R^i(x_s) = 0$ следует $R^i(x_{s+1}) = 0$ при любых α_s. Аналогично показывается, что при $\alpha_s < 1/\mu$ координаты вектора x_s не меняют знак для всех $s \ge 0$. Итак, множество X оказывается положительно-инвариантно относительно (3.1), как и в непрерывном случае.

Из формулы Ньютона–Лейбница получаем

$$F(x_{s+1}) \le F(x_s) - \alpha_s F_x(x_s)M(x_s)F_x(x_s) + \frac{L\alpha_s^2}{2}\|M(x_s)F_x(x_s)\|_n^2. \tag{3.5}$$

Симметрическая матрица $M(x_s)$ неотрицательно-определенная, поэтому существует арифметический квадратный корень из матрицы M. Обозначим его через $M^{1/2}$; тогда $M = = M^{1/2}M^{1/2}$. Введя вектор $y = M^{1/2}F_x$, преобразуем неравенство (3.5) к виду

$$F(x_{s+1}) - F(x_s) \le \alpha_s\|y\|^2\left[-1 + \frac{\alpha_s L}{2}\frac{y^\top M(x_s)y}{\|y\|_n^2}\right] \le \alpha_s\|y_s\|_n^2\left[-1+\frac{\alpha_s L\lambda}{2}\right].$$

Таким образом, при $\alpha_s < 2/\lambda L$ последовательность $F(x_s)$ монотонно убывает. Из ограниченности $F(x)$ снизу на X следует, что существует предел $F(x_s)$. Поэтому

$$\lim_{s\to\infty}[F(x_{s+1}) - F(x_s)] = 0, \qquad 0 \le F_x^\top(x_s)M(x_s)F_x(x_s) \le \frac{2[F(x_s)-F(x_{s+1})]}{\alpha_s[2-\alpha_s\lambda L]}$$

и, значит,

$$\lim_{s\to\infty} F_x^\top(x_s)M(x_s)F_x(x_s) = 0.$$

Используя представление (1.7), получаем, что

$$\lim_{s\to\infty}\|D(x_s)(F_x(x_s)+R_x(x_s)v_s)\|_n=\lim_{s\to\infty}\|D([R(x_s)]^{1/2})v_s\|_m=0,$$

т.е. в каждой предельной точке последовательности x_s выполнены условия стационарности (1.8). В силу изолированности стационарных точек существуют пределы

$$\bar{x}=\lim_{s\to\infty}x_s,\qquad \bar{R}^j=\lim_{s\to\infty}R^j(x),\qquad \bar{v}^j=\lim_{s\to\infty}v_s^j,\qquad \bar{H}_x^i=\lim_{s\to\infty}H_x^i(x_s,v_s).$$

Согласно (3.4) при $\bar{R}^j=0$ бесконечное произведение

$$\prod_{s=0}^{\infty}[1-\alpha_s v_s^j]$$

должно иметь нулевое значение. Для этого необходимо [11], чтобы ряд

$$\sum_{s=0}^{\infty}\ln[1-\alpha_s v_s^j]=-\infty,$$

но это возможно только при условии, что $\bar{v}^j\ge 0$. При $R^j>0$ должно обязательно выполняться $\bar{v}^j=0$. Точно так же можно убедиться, что $\bar{H}_x^i\ge 0$.

Итак, в предельной точке последовательности $\{x_s\}$ выполнены условия Куна–Таккера (1.6).

Оценка скорости сходимости (3.3) получается аналогично (2.4). В данном случае следует вычислить первую разность функции φ в силу системы (3.1). Теорема доказана. □

Метод (3.1) оказывается особенно эффективным, если в (1.1) ограничения типа равенства линейные, так как тогда можно брать сравнительно большие шаги α_s, строить различные модификации метода. Успешно использовался в этом случае вариант метода наискорейшего спуска. Нелинейность ограничений типа неравенства не вносит существенных усложнений в алгоритм.

Метод (3.1) можно использовать при достаточно малом α_s, если функции $g(x)$ нелинейные, однако в этом случае следует предусмотреть проверку условий $|g^i(x_s)|<\varepsilon$ и, если они нарушаются, уточнить текущее значение x_s.

На основе предложенных методов в ВЦ АН СССР разработаны три стандартные программы решения задач (1.1). В первой программе для случая, когда $g(x)$ нелинейные, производится интегрирование системы (1.2) по схеме Эйлера с постоянным шагом; во второй реализуется метод с переменным шагом α_s, который дробится в том случае, если нарушается условие релаксационности процесса или точка x_s выходит из допустимого множества. В третьей программе вдали от границы опускаются несущественные ограничения типа неравенства. Приведем некоторые результаты численных расчетов по первой программе. Пусть $n=3$, $e=c=1$, $F(x)=[x^1+3x^2+x^3]^2+4[x^1-x^2]^2$, $g(x)=x^1+$ $+x^2+x^3-1$, $h(x)=3-4x^3-6x^2+[x^1]^3$, $x\ge 0$. В качестве x_0 брался вектор $(0.1, 0.7, 0.2)$. Интегрирование системы проводилось до тех пор, пока уменьшение функции на каждом шаге было больше 10^{-5}. При $\alpha_s=0.1$ было сделано 83 шага, получено $F=1.8310951$, $x^1=0.2937327$, $x^2=0.1001667$, $x^3=0.6061005$. При $\alpha_s=0.5$ задача была решена за 13 шагов, причем $F=1.831030$, $x^1=0.2937386$, $x^2=0.1001510$, $x^3=0.6061104$.

§4. Нахождение седловых точек

Все приведенные выше методы очевидным образом обобщаются для решения задач отыскания седловых точек функций на несвязанных множествах. Пусть отыскивается

$$\min_{x\in X}\max_{y\in Y}F(x,y),\qquad X=\{x\in\mathbb{E}^n: h(x)\le 0_{p1}\},\quad Y=\{y\in\mathbb{E}^m: f(y)\le 0_{q1}\}.$$

Введем векторы $v \in \mathbb{E}^p$, $w \in \mathbb{E}^q$ и обозначим

$$\Phi(x, y, v, w) = F(x, y) + \sum_{i=1}^{p} v^i h^i(x) + \sum_{i=1}^{q} w^i f^i(y).$$

Метод (1.2) в этом случае приводит к системам $\dot{x} = -\Phi_x$, $\dot{y} = \Phi_y$, где

$$[h_x^\top h_x + D(-h)]v + h_x^\top F_x = 0, \qquad [f_y^\top f_y + D(-f)]w + f_y^\top F_y = 0.$$

Условия сходимости формулируются и доказываются аналогично теоремам 4, 5.

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

1. *Евтушенко Ю.Г., Жадан В.Г.* Численные методы решения некоторых задач исследования операций // ЖВМ и МФ, 1973, т.13, No.3, с. 583–598.

2. *Евтушенко Ю.Г.* Два численных метода решения задач нелинейного программирования // ДАН СССР, 1974, т.215, No.1, с. 38–40.

3. *Rosen J.B.* The gradient projection method for nonlinear programming // Pt. I. J. Soc. Industr. and Appl. Math.. Part I: 1960, v.8, No.1, p. 181–217; Part II: 1961, v.9, No.4, p. 514–532.

4. *Карманов В.Г.* Математическое программирование. М.: Наука, 1975.

5. *Барбашин Е.А.* Введение в теорию устойчивости. М.: Наука, 1967.

6. *Евтушенко Ю.Г., Жадан В.Г.* Применение метода функций Ляпунова для исследования сходимости численных методов // ЖВМ и МФ, 1975, т.15, No.1, с. 101–112.

7. *Эрроу К.Дж., Гурвиц Л., Удзава Х.* Исследования по линейному и нелинейному программированию. М.: Изд-во ин. лит., 1962.

8. *Taylor J.G.* A squared-variable transformation approach to nonlinear programming optimality conditions // Naval Res. Logist. Quart., 1973, v.20, No.1, p. 25–29.

9. *Taha Hamdy, Curry Guy.* Classical derivation of the necessary and sufficient conditions for optimal linear programs // Operat. Res., 1971, v.19, No.4, p. 1045–1050.

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

11. *Фихтенгольц Г.М.* Курс дифференциального и интегрального исчисления. Т.II. М.: Физматгиз, 1959.