1974 г. ОПТИМИЗАЦИЯ Выпуск 14(31)
Численные методы

УДК 512.25/26

# МЕТОД ДРОБНЫХ ШАГОВ ДЛЯ РЕШЕНИЯ ОПТИМИЗАЦИОННЫХ ЗАДАЧ

И.А.Ицкович

В статье предлагается итеративный метод решения конечномерных задач оптимизации линейной или выпуклой (в нужную сторону) функции на многогранном множестве, определяемом системой линейных неравенств. Одновременно решается прямая и двойственная задачи. Идея метода состоит в замене исходной задачи последовательностью задач о безусловном экстремуме функции, зависящей от переменных как прямой, так и двойственной задач. Движение происходит поочередно в пространстве переменных только одной из пары задач - прием, который (по другому поводу) был назван методом дробных шагов.

Для выбора направления движения используется вспомогательное пространство переменных, которые появляются, если в условиях дополняющей нежесткости нули в правой части заменить произвольными положительными числами.

Статья состоит из двух параграфов. В § I рассматривается линейная целевая функция, в § 2 - нелинейная целевая функция.

## § 1. Линейная целевая функция

### I. Введение

Пусть $A$ - матрица, имеющая $m$ строк и $n$ столбцов, $A'$ - ее транспонированная матрица, $b$ - $m$-мерный вектор, $c$ - $n$-мерный вектор.

Рассмотрим прямую задачу линейного программирования

$$\max \{(c,x): Ax \le b, x \ge 0\}$$

и двойственную к ней

$$\min \{(b,u): A'u \ge c, u \ge 0\}$$

при следующих допущениях:

а) в каждой строке и в каждом столбце матрицы $A$ имеются ненулевые элементы;

б) векторы $b$ и $c$ отличны от нуля;

в) обе задачи разрешимы;

г) множества оптимальных векторов в прямой и двойственной задачах ограничены;

д) существуют векторы $x > 0$, $u > 0$ такие, что $Ax < b$, $A'u > c$.

Пусть $X$ - $n$-мерное векторное пространство переменных $x_1, x_2, \dots, x_n$; $\bar{G} \subset X$ - множество допустимых векторов прямой задачи; $\bar{G} = \{x \in X : x \geq 0, Ax \le b\}$; $\mathcal{U}$ - $m$-мерное векторное пространство переменных $u_1, \dots, u_m$; $\bar{H} \subset \mathcal{U}$ - множество допустимых векторов двойственной задачи; $\bar{H} = \{u \in \mathcal{U} : u \geq 0, A'u \geq c\}$.

Из допущения д) следует существование внутренних точек во множествах $\bar{G}$ и $\bar{H}$; обозначим символами $G$ и $H$ множества внутренних точек соответственно множеств $\bar{G}$ и $\bar{H}$.

Введем $n+m$-мерное векторное пространство $E$ следующим образом: определим векторы $y = (y_1, y_2, \dots, y_m)$ и $v = (v_1, v_2, \dots, v_n)$ равенствами:

$$y = b - Ax, \quad v = A'u - c \tag{1}$$

и определим числа $p_k$ $(1 \le k \le n)$ и $q_j$ $(1 \le j \le m)$ равенствами:

$$p_k = x_k v_k, \quad q_j = u_j y_j; \tag{2}$$

пространство $E$ определим теперь как векторное пространство переменных $p_k$ и $q_j$. Пусть $Q$ - множество положительных векторов пространства $E$, $\bar{Q}$ - его замыкание.

Равенства (1), (2) описывают однозначное и непрерывное отображение $\varphi: X \times \mathcal{U} \to E$; легко видеть, что $\varphi: G \times H \to Q$ и

$\varphi : \bar{G} \times \bar{H} \to \bar{Q}$. В работе [I] доказано, что при допущениях а) – д) отображение $\varphi^{-1} : Q \to G \times H$ является однозначным и непрерывным, так что $\varphi$ отображает $G \times H$ на $Q$ взаимно-однозначно и непрерывно. Вообще, отображение $\varphi^{-1} : E \to X \times \mathcal{U}$ не является однозначным, и точка $O \in E$ является точкой ветвления. Точке $O$ соответствуют не только все оптимальные векторы прямой и двойственной задач, но и другие точки из $X \times \mathcal{U}$.

Нам удобно будет символом $[x, u]$ обозначать $n+m$-мерный вектор с компонентами $x_1, \ldots, x_n, u_1, \ldots, u_m$. Аналогичный смысл имеет символ $[p, q]$.

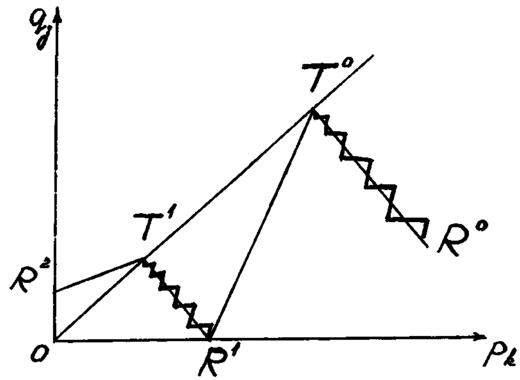


Рис. I.

Всякий итеративный метод, описанный в терминах пространства $E$, должен содержать процедуру перемещения из начальной точки $R^0 \in Q$ в начало координат $O$ (см. рис. I).

## 2. Описание процесса

Исходим из точки $[x^0, u^0] \in G \times H$. Такая точка существует по допущению д). Этой точке соответствует точка $R^0 = [p^0, q^0] \in Q$. Проведем в множестве $Q$ луч, каждая точка которого характеризуется тем, что все ее координаты одинаковы; будем называть этот луч осью множества $Q$. Зная точку $R^0$, найдем точку $T^0 = (t^0, t^0, \ldots, t^0)$ на оси такую, что $t^0 = (\sum_{k=1}^{n} p_k^0 + \sum_{j=1}^{m} q_j^0)/(n+m)$.

Первый шаг процесса состоит из четырех этапов:

Первый этап. Ищем точку $M^1=[x^1,u^1]=\varphi^{-1}(T^0)\in G\times H$, координаты точки $M^1$ обращают в максимум функцию

$$F(x,u,t^0)=\sum_{k=1}^{n}\ln p_k+\sum_{j=1}^{m}\ln q_j-(\sum_{k=1}^{n}p_k+\sum_{j=1}^{m}q_j)/t^0.$$

Второй этап. Исходя из точки $x^1\in G$, движемся по лучу $x^1+\lambda c\ (\lambda>0)$ до границы множества $\bar{G}$. Обозначим символом $x^2$ точку пересечения луча с границей множества $\bar{G}$. Координаты точки $u^1$ при этом не изменяются. Точке $[x^2,u^1]$ соответствует точка $R^1$ на границе ортанта $\bar{Q}$.

Третий этап. Определяем координаты $t^1$ точки $T^1$ на оси по известным координатам точки $R^1$ аналогично тому, как определили координаты точки $T^0$, исходя из точки $R^0$. Ищем координаты точки $M^2=[x^3,u^3]=\varphi^{-1}(T^1)\in G\times H$, максимизируя функцию $F(x,u,t^1)$.

Четвертый этап. Исходя из $u^3\in H$, движемся по лучу $u^3-\lambda b$ ($\lambda>0$) до границы множества $\bar{H}$, приходим в точку $[x^3,u^4]$, которой соответствует точка $R^2$ на границе ортанта $Q$.

Затем процесс продолжается, каждый шаг процесса состоит из четырёх этапов.

## 3. Сходимость процесса

Примем величину $\tau=\sum_{k=1}^{n}p_k+\sum_{j=1}^{m}q_j$ за меру отклонения точки из $Q$ от начала координат O. Нужно доказать, что $\tau\to 0$. Первый и третий этапы не изменяют величины $\tau$, на втором и четвертом этапах $\tau$ уменьшается. Рассмотрим подробно второй этап, четвертый ему аналогичен.

На втором этапе мы исходим из точки $T^0$ на оси, все координаты которой равны между собой и равны $t^0$ ($p_k=q_j=t^0$). Фиксируем $u$ и изменяем $x$ вдоль луча $x(\lambda)=x+\lambda c\ (\lambda>0)$ до пересечения с границей множества $Q$. Это приводит нас к задаче о нахождении максимального неотрицательного значения $\lambda$, удовлетворяющего $m+n$ неравенствам $p_k(\lambda)=t+\lambda c_k v_k\geq 0$, $q_j(\lambda)=t-\lambda u_j\sum_{k=1}^{n}a_{jk}c_k\geq 0$.

Пусть $\lambda^* = \max \lambda$. Ясно, что

$$t/\lambda^* = \max_{j,k} \{-c_k v_k, u_j \sum_{k=1}^{n} a_{jk} c_k\} = f.$$

Определяем $f$ как максимальное из $n+m$ чисел; сумма этих $n+m$ чисел равна

$$\sum_{k=1}^{n} (-c_k v_k) + \sum_{j=1}^{m} u_j \sum_{k=1}^{n} a_{jk} c_k =$$
$$= \sum_{k=1}^{n} c_k [-v_k + \sum_{j=1}^{m} a_{jk} u_j] = \sum_{k=1}^{n} c_k^2 > 0,$$

так что $f \geq \sum_{k=1}^{n} c_k^2 / (n+m)$ - максимальное число, не меньшее среднего арифметического.

Далее, вместе с $x$ изменяются $p_k$ и $q_j$, а также $t = (\sum_{k=1}^{n} p_k + \sum_{j=1}^{m} q_j)/(n+m)$, причем $t(\lambda) = t - \lambda \sum_{k=1}^{n} c_k^2 / (n+m)$. Легко видеть, что $t(\lambda^*) = t - \lambda^* \sum_{k=1}^{n} c_k^2 / (n+m) = t[1 - \frac{1}{f(n+m)} \sum_{k=1}^{n} c_k^2] > 0$

по определению числа $f$ ; случаи $t(\lambda^*) = 0$ является исключительным: это возможно, если все $n+m$ чисел, определяющих число $f$, равны между собой; если это так, то на конечном шаге мы достигаем оптимума.

Рассмотрим $\Omega_0 \in Q$, $\Omega_0 = \{[p,q] \in Q : \sum_{k=1}^{n} p_k + \sum_{j=1}^{m} q_j \leq t^0\}$, где $t^0$ соответствует начальной точке $R^0$, и пусть $\varphi^{-1}(\Omega_0) = \omega_0 \subset G \times H$. Множество $\omega_0$ ограничено, следовательно, ограничена и величина $f_0 = \max\{f : [x,u] \in \omega_0\}$, причем $f_0$ зависит только от параметров задачи и начального приближения и не зависит от текущих значений $x$ и $u$. Обозначим $\varepsilon_1 = \frac{1}{f_0(n+m)} \sum_{k=1}^{n} c_k^2$. Тогда $\frac{t(\lambda^*)}{t} \leq \varepsilon_1 < 1$ и $\varepsilon_1$ не зависит

от шага процесса; на четвертом этапе аналогичные рассуждения приводят нас к числу $\varepsilon_2 < 1$. Пусть $\varepsilon = \min(\varepsilon_1, \varepsilon_2)$, тогда на каждом шаге величина $\tau$ уменьшается на $\varepsilon^2$ и после $s$ шагов $\tau \leq \tau_0 \varepsilon^{2s}$. При $s \to \infty$ $\tau \to 0$, чем и оканчивается доказательство сходимости процесса.

### Замечания

1) На первом и третьем этапах максимизируется функция, содержащая логарифмы, поэтому на втором и четвертом этапах движение по градиенту целевой функции не должно продолжаться вплоть до границы ($\ln 0 = -\infty$). Тогда следует взять число $\alpha$, близкое к единице (например, 0,9 или 0,99), и прекращать движение к границе при $\lambda = \alpha\lambda^*$.

2) Исходя из точки $[p^\circ, q^\circ]$ на первом и третьем этапах не обязательно нужно двигаться к точке на оси с координатами $t^\circ = (\sum_{k=1}^{n} p_k^\circ + \sum_{j=1}^{n} q_j^\circ)/(n+m)$, можно выбрать другую точку с координатами $\beta t^\circ$, где $0<\beta<1$, однако не следует брать $\beta$ слишком малым.

## § 2. Нелинейная целевая функция

### I. Постановка задачи

Рассмотрим задачу о минимизации строго вогнутой функции $\varphi(x)+(c,z)$ при ограничениях $Ax+Bz \geq b$, $x \geq 0$, $z \geq 0$. Здесь $b$ - $m$-мерный вектор, $c$, $z$ - $k$-мерные векторы, $x$ - $n$-мерный вектор, $A=\{\alpha_{ij}\}$, $B=\{\beta_{il}\}$. Двойственной [2] является задача о максимизации функции $\theta(y)+(b,t)$ при ограничениях $A't - y \leq 0$, $B't \leq c$, $t \geq 0$; здесь $y$ - $n$-мерный вектор, $\theta(y)=\varphi((\nabla\varphi)^{-1}y)-(y,(\nabla\varphi)^{-1}y)$, где $\nabla\varphi$ - градиент функции $\varphi$, а $(\nabla\varphi)^{-1}$ - отображение, обратное к $\nabla\varphi$. Переход от функции $\varphi$ к функции $\theta$ называется преобразованием Лежандра.

Примем два допущения:

1) существуют векторы $x^\circ>0$, $t^\circ>0$, $z^\circ>0$, $y^\circ$ такие, что $Ax^\circ+Bz^\circ>b$, $A't^\circ-y^\circ<0$, $B't^\circ<c$ (условия Слейтера для прямой и двойственной задач);

2) если область изменения переменной $x$ не ограничена, то $\varphi(x)$ задана во всей области и $\varphi(x)\to+\infty$ при любом стремлении $x$ к бесконечности; аналогично, функция $\theta(y)\to-\infty$, когда $y$ неограниченно удаляется.

Приведем некоторые свойства этой пары задач [2].

1. Для того чтобы векторы $x, z$ были оптимальны, необходимо и достаточно, чтобы существовали векторы $t \geq 0$, $u \geq 0$, $v \geq 0$, $w \geq 0$ и $y$ такие, что выполняются равенства $Ax + Bz - u = b$, $A't - y + v = 0$, $B't + w = c$ (векторы $x$, $z$, $t$, $y$ допустимые).

$y = \nabla\varphi(x)$ – связь между переменными прямой и двойственной задач;

$x_j v_j = 0 (1 \leq j \leq n)$, $t_i u_i = 0 (1 \leq i \leq m)$, $z_\ell w_\ell = 0 (1 \leq \ell \leq k)$ – условия дополняющей нежесткости.

2. Для допустимых векторов $x$, $y$, $z$, $t$ $(y = \nabla\varphi(x))$ имеет место равенство

$$\varphi(x) + (c, x) - \theta(y) - (b, t) = \sum_{j=1}^{n} x_j v_j + \sum_{i=1}^{m} t_i u_i + \sum_{\ell=1}^{k} z_\ell w_\ell.$$

Пусть $x$, $y$, $z$, $t$ – допустимые векторы и $y = \nabla\varphi(x)$. Обозначим $p_j = x_j v_j$ $(1 \leq j \leq n)$, $q_i = t_i u_i$ $(1 \leq i \leq m)$, $r_\ell = z_\ell w_\ell$ $(1 \leq \ell \leq k)$ или $p_j = x_j\,(y_j - \sum_i \alpha_{ij} t_j)$, $q_i = t_i (\sum_j \alpha_{ij} x_j + \sum_\ell \beta_{i\ell} z_\ell - b_i)$, $r_\ell = z_\ell\,(c_\ell - \sum_i \beta_{i\ell} t_i)$. Эти равенства осуществляют отображение множества допустимых $[x, z, t]$ во множестве неотрицательных $[p, q, r]$. В [1] доказано, что это отображение взаимно-однозначно, причем внутренним точкам допустимых множеств соответствуют положительные векторы $[p, q, r]$. Пространство переменных $p$, $q$, $r$ названо в [1] пространством возмущений, так как числами $p_j$, $q_i$, $r_\ell$ "возмущаются" условия дополняющей нежесткости.

## 2. Описание алгорифма

Итеративный метод на каждом шаге состоит из четырех этапов:

Первый этап. Исходя из начальной точки $[x^\circ, z^\circ, t^\circ]$, определяем точку $[p^\circ, q^\circ, r^\circ]$ в пространстве возмущений и число

$$\tau = \Big(\sum_j p_j^\circ + \sum_i q_i^\circ + \sum_\ell r_\ell^\circ\Big)/(n + m + k).$$

Максимизация строго вогнутой функции

$$F = \sum_{j=1}^{n} \ln p_j + \sum_{i=1}^{m} \ln q_i + \sum_{\ell=1}^{k} \ln r_\ell - \Big(\sum_{j=1}^{n} p_j + \sum_{i=1}^{m} q_i + \sum_{\ell=1}^{k} r_\ell\Big)/\tau =$$

$$= \sum_j \ln x_j + \sum_i \ln t_i + \sum_\ell \ln z_\ell + \sum_j \ln (y_j - \sum_i \alpha_{ij} t_i) +$$

$$+ \sum_i \ln (\sum_j \alpha_{ij} x_j + \sum_\ell \beta_{i\ell} z_\ell - b_i) + \sum_\ell \ln (c_\ell - \sum_i \beta_{i\ell} t_i) -$$

$$- [\varphi(x) + (c, z) - \theta(y) - (b, t)] / \tau$$

приводит к точке $[x^1, z^1, t^1]$, которой соответствует вектор $[p^1, q^1, r^1]$, все компоненты которого равны $\tau$.

Второй этап. Исходя из точки $[x^1, z^1, t^1]$, движемся в пространстве переменных прямой задачи в направлении градиента целевой функции, т.е. по прямой $x^1 - \lambda \nabla \varphi(x^1)$, $z^1 - \lambda c$, до пересечения с границей области допустимых векторов в точке $[x^2, z^2, t^1]$.

Третий этап. Исходя из точки $[x^2, z^2, t^1]$, максимизируя функцию $F$, приходим к точке $[x^3, z^3, t^3]$, которой соответствует точка на оси положительного ортанта в пространстве возмущений.

Четвертый этап. Исходя из точки $[x^3, z^3, t^3]$, на оси движемся в пространстве переменных двойственной задачи в направлении градиента целевой функции, т.е. по прямой $t^3 + \lambda b$, $y^3 - \lambda x^3$, до пересечения с границей области допустимых векторов двойственной задачи.

Затем процесс повторяется.

## 3. Сходимость процесса

Докажем сходимость процесса. Мерой отклонения допустимых векторов от оптимальных будем считать величину $\tau = \sum_j p_j + \sum_i q_i + \sum_\ell r_\ell$.

Первый и третий этапы не изменяют величины $\tau$. Рассмотрим подробно второй и четвертый этапы и покажем, что процесс не замедляется при приближении к оптимуму.

Второй этап. На втором этапе величины $p$, $q$, $r$ вычисляются по формулам:

$$p_j(\lambda) = v_j x_j(\lambda) = v_j [x_j - \lambda(\nabla \varphi(x))_j] = \tau - \lambda v_j [\nabla \varphi(x)]_j \geqslant 0,$$

$$q_i(\lambda) = t_i u_i(\lambda) = t_i [(Ax)_i - \lambda(A \nabla \varphi(x))_i + (Bz)_i - \lambda(Bc)_i - b_i] =$$

$$= \tau - \lambda [t_i (A \nabla \varphi(x))_i + t_i (Bc)_i] \geq 0,$$

$$r_\ell(\lambda) = w_\ell z_\ell(\lambda) = w_\ell [z_\ell - \lambda c_\ell] = \tau - \lambda w_\ell c_\ell \geq 0.$$

Мы ищем максимальное значение $\lambda$, при котором сохраняются написанные выше неравенства. Пусть $\lambda^*$ - максимальное значение $\lambda$ при этих ограничениях. Тогда

$$f = \tau/\lambda^* = \max_{i,j,\ell} \{t_i (A \nabla \varphi(x))_i + t_i (Bc)_i ; v_j (\nabla \varphi(x))_j ; w_\ell c_\ell\}.$$

Найдем сумму $\Sigma$ тех величин, среди которых ищем максимум:

$$\Sigma = (t, A \nabla \varphi(x)) + (t, Bc) + (v, \nabla \varphi(x)) + (w, \ell) = (t, Ay) + (t, Bc) +$$

$$+ (y - A't, y) + (c - B't, c) = (t, Ay) + (t, Bc) + (y, y) - (A't, y) +$$

$$+ (c, c) - (B't, c) = (y, y) + (c, c) > 0.$$

Поскольку максимум из ряда чисел не меньше среднего арифметического этих же чисел, то $f = \tau/\lambda^* \geq \dfrac{(y,y)+(c,c)}{n+m+k} > 0$. Пусть теперь

$$\tau(\lambda) = \frac{\sum_j p_j(\lambda) + \sum_i q_i(\lambda) + \sum_\ell r_\ell(\lambda)}{n + m + k} = \tau - \lambda \frac{(y,y)+(c,c)}{n+m+k}.$$

Тогда при $\lambda = \lambda^*$ имеем

$$\tau(\lambda^*) = \tau \left[1 - \frac{\lambda^*}{\tau} \frac{(y,y)+(c,c)}{n+m+k}\right] = \tau \left[1 - \frac{1}{f} \frac{(y,y)+(c,c)}{n+m+k}\right] \geq 0.$$

в силу неравенства для $f$.

Таким образом, на втором этапе мы уменьшаем $\tau$.

Если нам удастся доказать, что $f$ ограничена сверху независимо от положения точки $[x, z, t]$ в нашем процессе, то мы получим оценку сверху для множителя при $\tau$:

$$1 - \frac{1}{f} \frac{(y,y)+(c,c)}{n+m+k} \leq 1 - \frac{1}{f_{max}} \frac{(c,c)}{n+m+k} = \varepsilon_1 < 1,$$

и $\varepsilon_1$ будет единым для всего процесса.

Мы начинаем процесс с точки $[x^\circ, z^\circ, t^\circ]$, которой соответствует точка $[p^\circ, q^\circ, r^\circ]$. Рассмотрим множество

$$\Omega_0 = \{[p, q, r] \geq 0 : \sum_j p_j + \sum_i q_i + \sum_\ell r_\ell \leq \sum_j p_j^\circ + \sum_i q_i^\circ + \sum_\ell r_\ell^\circ\}.$$

Множество $\Omega_0$ ограничено, и ему соответствуют ограниченные множества допустимых векторов в прямой и двойственной задачах, поэтому и $f$ ограничено. Пусть $f_{max}$ - его верхняя граница.

Четвертый этап. $p_j(\lambda)$, $q_i(\lambda)$, $r_\ell(\lambda)$ имеют вид:

$$p_j(\lambda) = x_j v_j(\lambda) = x_j[y_j - \lambda x_j - (A't)_j - \lambda(A'b)_j] =$$
$$= \tau - \lambda[x_j^2 - x_j(A'b)_j] \geq 0.$$

$$q_i(\lambda) = u_i t_i(\lambda) = u_i[t_i + \lambda b_i] = \tau + \lambda u_i b_i \geq 0.$$

$$r_\ell(\lambda) = z_\ell w_\ell(\lambda) = z_\ell[c_\ell - (B't)_\ell - \lambda(B'b)_\ell] = \tau - \lambda z_\ell(B'b)_\ell \geq 0.$$

Здесь $\tau/\lambda^* = \max_{j,i,\ell}\{x_j^2 - x_j(A'b)_j; -u_i b_i; z_\ell(B'b)_\ell\} = f$. Сумма $\Sigma$ тех величин, среди которых имеется максимум, равна

$$\Sigma = (x,x) + (x, A'b) - (u, b) + (z, B'b) = (x,x) + (x, A'b) -$$
$$-(Ax + Bz - b, b) + (z, B'b) = (x,x) + (b,b) > 0,$$

так что $f = \tau/\lambda^* \geq \dfrac{(x,x)+(b,b)}{n+m+k} > 0$.

Далее, $\tau(\lambda) = \tau - \lambda\dfrac{(x,x)+(b,b)}{n+m+k}$, и при $\lambda = \lambda^*$ имеем

$$\tau(\lambda^*) = \tau\left[1 - \frac{\lambda^*}{\tau}\frac{(x,x)+(b,b)}{n+m+k}\right] \leq \tau\left[1 - \frac{1}{f_{max}}\frac{(b,b)}{n+m+k}\right] \leq \varepsilon_2 < 1,$$

причем $\varepsilon_2$ не зависит от шага процесса. Выбирая $\varepsilon = \max(\varepsilon_1, \varepsilon_2)$, мы имеем $0 < \varepsilon < 1$, причем $\varepsilon$ не зависит от шага процесса.

Этим доказательство сходимости оканчивается.

## Литература


I. ИЦКОВИЧ И.А., АНДЕРСОН М.В. Об устойчивости условий дополняющей нежесткости. - В кн.: Математические методы решения экономических задач. Новосибирск, "Наука", 1971, с.153-166.

2. ДЕННИС Дж. Математическое программирование и электрические цепи. М., ИЛ, 1961 (Приложения Д и Е).

Поступила в ред.-изд. отд.
30.YI.1974 г.