КРАТКИЕ СООБЩЕНИЯ

Ю.П. ЛАПТИН

УДК 519.8

ОДИН ПОДХОД К РЕШЕНИЮ НЕЛИНЕЙНЫХ ЗАДАЧ ОПТИМИЗАЦИИ С ОГРАНИЧЕНИЯМИ

Ключевые слова: *выпуклое программирование, безусловная оптимизация, условный ε-субдифференциал, почти дифференцируемые функции.*

При решении нелинейных задач оптимизации с ограничениями один из часто используемых приемов — формирование вспомогательной задачи безусловной оптимизации, решение которой совпадает с решением исходной задачи. Такими являются методы штрафных функций, в частности метод негладких штрафных функций, широко применяемый в выпуклом программировании [1–4]. При использовании этих методов возникают определенные трудности, связанные с подбором значений штрафных коэффициентов. Дополнительные проблемы порождает случай, когда функции, описывающие задачу, определены на ограниченных множествах [5].

В настоящей работе рассматривается один из возможных подходов к сведению задачи выпуклого программирования к задаче безусловной оптимизации. Предполагается заданной начальная точка, принадлежащая внутренности допустимого множества. Эквивалентная задача безусловной оптимизации формируется таким образом, что градиенты (субградиенты) и значения функций исходной задачи вычисляются только в точках допустимого множества. Сформулированная задача оказывается близко связанной с понятием условного ε-субдифференциала [3, 4].

Исследуются свойства введенных функций. Формулируются условия, при которых задача безусловной оптимизации оказывается выпуклой. Полученные результаты могут быть полезны при разработке алгоритмов решения оптимизационных задач с ограничениями.

1. Рассматривается следующая задача выпуклого программирования: найти

$$f^* = \min\ f(x) \tag{1}$$

при ограничениях

$$h(x) \le 0, \tag{2}$$

где $x \in R^n$ $f, h : R^n \to R$— выпуклые функции, принимающие конечные значения при любых x.

Обозначим $S = \{x \in R^n :\ h(x) \le 0\}$. Предполагается заданной допустимая точка $x^0 \in S$ такая, что $h(x^0) < 0$.

Для произвольной точки $x \in R^n$ определим отображение на множество S:

$$\pi_S(x) = (1-\alpha)x^0 + \alpha x,\ \text{где}\ \alpha = \max\{\tilde{\alpha} : (1-\tilde{\alpha})x^0 + \tilde{\alpha}x \in S,\ 0 \le \tilde{\alpha} \le 1\}.$$

Если $x \in S$, то $x = \pi_S(x)$, если $x \notin S$, то $\pi_S(x)$ принадлежит границе множества S.

Такого рода отображения рассматривались Н.Г. Журбенко при разработке модификаций методов проекции градиентов, в которых трудоемкая операция проектирования на допустимое множество заменяется приведенной операцией отображения. Одномерный поиск для определения значения α может быть реализован достаточно эффективно. Близкий подход использовался в [6] для регуляризации задач в схемах декомпозиции по переменным.

Положим $\varphi(x)=f(\pi_S(x))$, $d(x)=\|x-\pi_S(x)\|$ и определим функцию

$$\psi_\varepsilon(x)=\varphi(x)+\gamma_\varepsilon(x)d(x), \tag{3}$$

где $\gamma_\varepsilon(x)=\dfrac{f(\pi_S(x))-f(x^0)+\varepsilon}{\|\pi_S(x)-x^0\|}$, если $x\neq x^0$, и $\gamma_\varepsilon(x)=0$, если $x=x^0$, $\varepsilon>0$ — некоторое число. Заметим, что $\psi_\varepsilon(x)=f(x)$, если $x\in S$. Рассмотрим задачу

$$\psi_\varepsilon^*=\inf\{\psi_\varepsilon(x): x\in R^n\}. \tag{4}$$

Нетрудно видеть, что если $\gamma_\varepsilon(x)\geq 0$ для всех $x\notin S$, то $\psi_\varepsilon^*=f^*$. Если найдется точка $x\notin S$ такая, что $\gamma_\varepsilon(x)<0$, то $\psi_\varepsilon^*=-\infty$.

В дальнейшем будет показано, что если множество S ограничено, то при достаточно больших ε функция $\psi_\varepsilon(x)$ выпукла.

Обозначим Φ надграфик функции f на множестве S:

$$\Phi=\{(\lambda,x)\in R\times S:\ \lambda\geq f(x)\}.$$

Положим $z=(\lambda,x)$. Пусть задано число $\varepsilon>0$. Рассмотрим в пространстве $R\times R^n$ выпуклую коническую оболочку $\mathrm{K}(\varepsilon)$ надграфика Φ относительно точки $z^\varepsilon=(f(x^0)-\varepsilon,x^0)$,

$$\mathrm{K}(\varepsilon)=\{v=z^\varepsilon+\alpha(z-z^\varepsilon)\,|\,\alpha\geq 0,\ z\in\Phi\}. \tag{5}$$

Нетрудно видеть, что множество $\mathrm{K}(\varepsilon)$ — надграфик некоторой выпуклой функции. Обозначим эту функцию $\chi_\varepsilon(x)$.

Пусть $\chi'_\varepsilon(x^0,p)$ — производная по направлению p функции $\chi_\varepsilon(x)$ в точке x^0.

Очевидно, что $\chi_\varepsilon(x^0)=f(x^0)-\varepsilon$, эта функция линейна по любому направлению p при смещении из точки x^0, т.е. $\chi_\varepsilon(x^0+pt)=f(x^0)-\varepsilon+\chi'_\varepsilon(x^0,p)\cdot t$, $t>0$. Заметим также, что

$$\chi_\varepsilon(x)\leq f(x)\quad \forall x\in S. \tag{6}$$

Рассмотрим точку $\bar{x}$, принадлежащую границе множества S, обозначим $p(\bar{x})=\dfrac{\bar{x}-x^0}{\|\bar{x}-x^0\|}$, $f'(\bar{x},p)$ — производная функции f в точке $\bar{x}$ по направлению $p=p(\bar{x})$.

Лемма 1. Пусть выполняется $f(\bar{x})-f'(\bar{x},p)\cdot\|\bar{x}-x^0\|\geq f(x^0)-\varepsilon$. Тогда $f'(\bar{x},p)\leq\chi'_\varepsilon(x^0,p)$ и $\chi_\varepsilon(\bar{x})=f(\bar{x})$.

Доказательство. Для заданого направления $p=p(\bar{x})$ рассмотрим функции $f_p(t)=f(x^0+pt)$ и $\chi_p^\varepsilon(t)=\chi_\varepsilon(x^0+pt)$. При $t=\bar{t}=\|\bar{x}-x^0\|$ имеем $\bar{x}=x^0+p\bar{t}$. Покажем, что $f_p(\bar{t})=\chi_p^\varepsilon(\bar{t})$.

Предположим противное, т.е. $f_p(\bar{t})>\chi_p^\varepsilon(\bar{t})$. Нетрудно видеть (по построению), что существует t', $0\leq t'<\bar{t}$, такое, что $f_p(t')=\chi_p^\varepsilon(t')$. Отсюда получаем

$$f(\bar{x})-f'(\bar{x},p)\cdot\|\bar{x}-x^0\|\leq f_p(\bar{t})-\frac{f_p(\bar{t})-f_p(t')}{\bar{t}-t'}(\bar{t}-t'+t')=$$

$$= \chi_p^\varepsilon(t') - \frac{f_p(\bar t) - \chi_p^\varepsilon(t')}{\bar t - t'} t' <$$

$$< \chi_p^\varepsilon(t') - \frac{\chi_p^\varepsilon(\bar t) - \chi_p^\varepsilon(t')}{\bar t - t'} t' = \chi_p^\varepsilon(t') - \chi'_\varepsilon(x^0, p) \cdot t' = f(x^0) - \varepsilon,$$

что противоречит условию леммы, т.е. $\chi_\varepsilon(\bar x) = f(\bar x)$.

Покажем, что $f'(\bar x, p) \le g^\varepsilon(p)$. Из условия леммы следует $f'(\bar x, p) \cdot \|\bar x - x^0\| \le \le f(\bar x) - f(x^0) - \varepsilon = \chi_\varepsilon(\bar x) - f(x^0) - \varepsilon = \chi'_\varepsilon(x^0, p) \cdot \|\bar x - x^0\|$. Откуда получаем утверждение леммы.

Пусть $\operatorname{fr} S$ — граница множества S, $\bar x \in \operatorname{fr} S$, $p(\bar x) = \dfrac{\bar x - x^0}{\|\bar x - x^0\|}$. Обозначим

$$\varepsilon(\bar x) = f(x^0) - f(\bar x) + f'(\bar x, p(\bar x)) \cdot \|\bar x - x^0\|, \tag{7}$$

$$\varepsilon^* = \sup\{\varepsilon(\bar x) : \bar x \in \operatorname{fr} S\}. \tag{8}$$

Теорема 1. Пусть множество S ограничено, f, h — выпуклые функции, принимающие конечные значения при любых $x \in R^n$. Тогда ε^* конечно и для всех $\varepsilon > \varepsilon^*$ функция $\psi_\varepsilon(x)$ выпуклая.

Доказательство. Можно показать, что из конечности функций f, h при любых x следует, что множество S замкнуто, а функция f удовлетворяет условию Липшица на S. Отсюда заключаем, что величина ε^* в (8) конечна.

Положим $\xi(x) = \begin{cases} \chi_\varepsilon(x), & x \notin S, \\ f(x), & x \in S. \end{cases}$

Пусть $\varepsilon > \varepsilon^*$, тогда из леммы 1 вытекает $\chi_\varepsilon(\bar x) = f(\bar x)$ для $\bar x \in \operatorname{fr} S$, т.е. функция $\xi(x)$ непрерывна при любых x.

Будем говорить, что линейная функция $l(x) = \langle g, x\rangle - \mu$ является опорной относительно множества S к функции $f(x)$, если $l(x) \le f(x)$ для любых $x \in S$ и существует точка $x' \in S$ такая, что $l(x') = f(x')$ и $g \in \partial f(x')$, где $\partial f(x')$ — субдифференциал функции $f(x)$ в точке x'. Точку x' будем называть опорной точкой функции $l(x)$.

Рассмотрим совокупность всех опорных относительно множества S функций $l(x) = \langle g, x\rangle - \mu$, обозначим H множество всех пар (g, μ), порождающих такие функции.

Положим

$$\breve f(x) = \max\{\langle g, x\rangle - \mu : (g, \mu) \in \mathrm{H}\}, \ x \in R^n.$$

Ясно, что функция $\breve f(x)$ выпуклая и $\breve f(x) = f(x) \ \forall x \in S$.

Можно показать, что при $\varepsilon > \varepsilon^*$ из леммы 1 следует $\chi_\varepsilon(x) \ge \breve f(x)$ для любых $x \notin S$, откуда, учитывая (6), получаем $\xi(x) = \max\{\breve f(x), \chi_\varepsilon(x)\}$, т.е. функция $\xi(x)$ выпуклая.

Пусть $x \notin S$. Положим $p = \dfrac{x - x^0}{\|x - x^0\|}$, $\bar x = \pi_S(x)$. Учитывая, что $\chi_\varepsilon(x^0) = = f(x^0) - \varepsilon$ и то, что функция $\chi_\varepsilon(x)$ линейна по любому направлению p при смещении из точки x^0, получаем

$$\xi(x) = \chi_\varepsilon(x) = \chi_\varepsilon(\bar x) + \frac{\chi_\varepsilon(\bar x) - \chi_\varepsilon(x^0)}{\|x - x^0\|} \|x - \bar x\| = f(\bar x) + \frac{f(\bar x) - f(x^0) + \varepsilon}{\|x - x^0\|} \|x - \bar x\| = \psi_\varepsilon(x).$$

Таким образом, $\psi_\varepsilon(x)$ — выпуклая функция.

Теорема доказана. ■

Рассмотрим $\partial_\varepsilon^S f(x^0) = \{v \in E^n \mid f(z) - f(x^0) \ge \langle v, z - x^0 \rangle - \varepsilon \ \ \forall z \in S\}$ — условный ε-субдифференциал функции $f(x)$ в точке x^0, $\varepsilon > 0$. Можно показать, что для введенной ранее функции $\chi_\varepsilon(x)$ имеет место

$$\chi_\varepsilon(x) = \sup\{f(x^0) + \langle v, x - x^0 \rangle - \varepsilon \mid v \in \partial_\varepsilon^S f(x^0)\}, \ x \in R^n. \tag{9}$$

Теорема 2. Пусть выполняются условия теоремы 1 и $\varepsilon > \varepsilon^*$. Задача (4) имеет конечное значение $\psi_\varepsilon^* = f^*$ тогда и только тогда, когда $0 \in \partial_\varepsilon^S f(x^0)$.

Доказательство. Для точек $x \notin S$ имеем $\psi_\varepsilon(x) = \chi_\varepsilon(x)$. Поскольку функция $\chi_\varepsilon(x)$ линейна по любому направлению p при смещении из точки x^0, для конечности ψ_ε^* необходимо и достаточно, чтобы производная по направлению p была неотрицательна, что эквивалентно $0 \in \partial_\varepsilon^S f(x^0)$. ■

Пусть задано выпуклое множество Ω, $\mathrm{fr}\, S$ — граница множества S. Положим

$$\varepsilon^*(\Omega) = \sup\{\varepsilon(x) : x \in \mathrm{fr}\, S \cap \Omega\}. \tag{10}$$

Теорема 3. Пусть выполняются условия теоремы 1 и дополнительно $x^0 \in \Omega$. Тогда для всех $\varepsilon > \varepsilon^*(\Omega)$ функция $\psi_\varepsilon(x)$ выпукла на множестве Ω.

Доказательство аналогично доказательству теоремы 1.

2. Задача (4) есть задача без ограничений, для ее решения необходимо исследовать свойства функций $\varphi(x)$, $d(x)$.

Рассмотрим отображение π_S. Пусть заданы точка $x \notin S$ и некоторое направление $p \in R^n$, $\|p\| = 1$. Введем следующие обозначения:

- $\pi'_S(x, p)$ — производная по направлению p отображения π_S в точке x, если такая производная существует;
- $\bar{x} = \pi_S(x)$;
- Γ — конус возможных направлений множества S в точке $\bar{x}$;
- $v = x - \bar{x}$.

Определим отображение $\theta(p, v, \bar{x})$ вектора p на границу конуса Γ:

$$\theta(p, v, \bar{x}) = p + \beta v, \ \text{где} \ \beta = \max\{\tilde{\beta} : p + \tilde{\beta} v \in \Gamma, \ \tilde{\beta} \in \mathrm{R}\}.$$

Отображение $\theta(p, v, \bar{x})$ перемещает точку p по направлению v до пересечения с границей конуса Γ.

Лемма 2. Для любого $p \in R^n$ отображение π_S дифференцируемо по направлению p в точке x и $\pi'_S(x, p) = r \cdot \theta(p, v, \bar{x})$, где $r = \|\bar{x} - x^0\| \cdot \|x - x^0\|^{-1}$.

Доказательство. Обозначим $\bar{\Gamma} = \bar{x} + \Gamma$. Рассмотрим отображение $\pi_{\bar{\Gamma}}$. Понятно, что $\bar{x} = \pi_S(x) = \pi_{\bar{\Gamma}}(x)$. Пусть $x' = x + p \cdot t$, $t > 0$. Положим $\bar{x}' = \pi_{\bar{\Gamma}}(x + pt)$. Для иллюстрации доказательства рассмотрим рис. 1, на котором изображено сечение множества $\bar{\Gamma}$ (заштриховано) плоскостью, проходящей через точки x^0, x, x'. Через точку x проведена вспомогательная прямая, параллельная отрезку $(\bar{x}, \bar{x}')$, через точку x' — прямая, параллельная отрезку $(x, \bar{x})$. Таким образом, отрезки $(\bar{x}', \bar{x})$, (z, x) и $(x, \bar{x})$, (y, x') параллельны. Нетрудно видеть, что $y - x = \theta(p \cdot t, v, \bar{x})$, откуда $\bar{x}' - \bar{x} = r \cdot (z - x) = r \cdot (z - y + y - x) = r \cdot (z - y) + r \cdot \theta(p \cdot t, v, \bar{x})$. Можно показать, что $z - y = o(t)$, а $\theta(p \cdot t, v, \bar{x}) = t \cdot \theta(p, v, \bar{x})$. Откуда $\pi_{\bar{\Gamma}}(x + pt) = \pi_{\bar{\Gamma}}(x) + t \cdot r \cdot \theta(p, v, \bar{x}) + o(t)$, т.е.

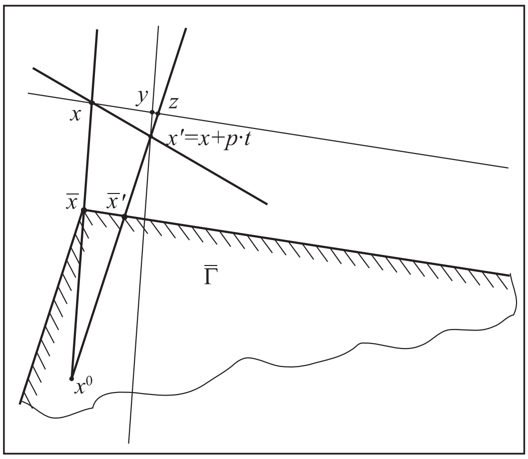

Рис. 1. Построения к доказательству леммы 2

отображение $\pi_{\bar{\Gamma}}$ дифференцируемо по направлению p в точке x.

Рассмотрим разность $\Delta(t)=\pi_{\bar{\Gamma}}(x+p\cdot t)-\pi_S(x+p\cdot t)$. Можно показать, что из выпуклости множества S следует монотонное убывание величины $\frac{\|\Delta(t)\|}{t}$ при $t\to+0$. Более того, поскольку Γ — конус возможных направлений множества S в точке $\bar{x}$, то $\lim\limits_{t\to+0}\frac{\|\Delta(t)\|}{t}=0$. Далее получаем $\pi_S(x+p\cdot t)-\pi_S(x)=\pi_{\bar{\Gamma}}(x+p\cdot t)-$

$$-\pi_{\bar{\Gamma}}(x)-\Delta(t)=t\cdot r\cdot\theta(p,v,\bar{x})-\Delta(t)+o(t).$$

Лемма доказана. ■

Лемма 3. Пусть функция h дифференцируема в точке $\bar{x}=\pi_S(x)$ и вектор $\nabla h(\bar{x})$ — ее градиент в точке $\bar{x}$. Тогда отображение π_S дифференцируемо в точке x и

$$\pi'_S(x,p)=r\left(p-\frac{(p,\nabla h(\bar{x}))}{(x-\bar{x},\nabla h(\bar{x}))}(x-\bar{x})\right), \tag{11}$$

где r определено в лемме 2.

Доказательство. Понятно, что конус Γ — полупространство, определяемое нормальным вектором $\nabla h(\bar{x})$. Результат отображения $\theta(p,v,\bar{x})$ вектора p на границу множества Γ имеет вид $y=p+\beta v$. Граница конуса $\Gamma(\bar{z})$ определяется соотношением $(\nabla h(\bar{x}),y)=0$. Отсюда находим β и получаем утверждение леммы. ■

Рассмотрим функцию $d(x)=\|x-\pi_S(x)\|$.

Лемма 4. В точках $x\notin S$ функция d дифференцируема по направлениям. Если функция h дифференцируема в точке $\bar{x}=\pi_{S(x)}$, то d дифференцируема в точке x и

$$\nabla d(x)=\frac{1-r}{d(x)}(x-\bar{x})+\frac{r}{(x-\bar{x},\nabla h(\bar{x}))}\nabla h(\bar{x}). \tag{12}$$

Доказательство. Пусть задано некоторое направление p. Нетрудно проверить, что при малых $t>0$ имеет место

$$d(x+tp)=\|x+tp-\bar{x}-t\pi'_S(x,p)+o(t)\|=\sqrt{\|x-\bar{x}\|^2+2t(p-\pi'_S(x,p))+o_1(t)}=$$

$$=\|x-\bar{x}\|+t\frac{(x-\bar{x},p-\pi'_S(x,p))}{\|x-\bar{x}\|}+o_2(t).$$

Отсюда производная функции d по направлению p в точке x равна $d'(x,p)=\frac{(x-\bar{x},\ p-\pi'_S(x,p))}{d(x)}$. Если функция h дифференцируема в точке $\bar{x}=\pi_S(x)$, то, подставляя для $\pi'_S(x,p)$ выражение из (11), получаем $d'(x,p)=\left(\frac{1-r}{d(x)}(x-\bar{x})+\frac{r}{(x-\bar{x},g)}g,\ p\right)$, где $g=\nabla h(\bar{x})$. Отсюда следует (12). ■

Рассмотрим функцию $\varphi(x)=f(\pi_S(x))$.

Лемма 5. Пусть $x \notin S$, в точке $\bar{x} = \pi_S(x)$ функция f дифференцируема. Тогда в точке x $\varphi(x)$ дифференцируема по направлениям. Если дополнительно в точке $\bar{x} = \pi_S(x)$ дифференцируема функция h, то $\varphi(x)$ дифференцируема и

$$\nabla\varphi(x) = r\left[\nabla f(\bar{x}) - \frac{(x-\bar{x}, \nabla f(\bar{x}))}{(x-\bar{x}, \nabla h(\bar{x}))}\nabla h(\bar{x})\right]. \qquad (13)$$

Доказательство. Пусть задано некоторое направление p, число $t > 0$. Нетрудно видеть, что $\varphi(x+tp) = f(\pi_S(x+t\cdot p)) = f(\pi_S(x) + t\pi'_S(x,p) + o(t))) = f(\bar{x}) + t(\nabla f(\bar{x}), \pi'_S(x,p)) + o(t)$, т.е. существует производная направлению p функции $\varphi(x)$ и она равна $\varphi'(x,p) = (\nabla f(\bar{x}), \pi'_S(x,p))$.

Если дополнительно в точке $\bar{x} = \pi_S(x)$ дифференцируема функция h, то, подставляя выражение для $\pi'_S(x,p)$, получаем $\varphi'(x,p) = r(\nabla f(\bar{x}) - \frac{(x-\bar{x}, \nabla f(\bar{x}))}{(x-\bar{x}, g)} g, p)$, где $g = \nabla h(\bar{z})$. Откуда следует (13). ■

3. Вернемся к задаче (4). Градиенты функции ψ_ε в точках, где она дифференцируема, вычисляются по правилам дифференцирования сложных функций в соответствии с соотношениями (11)–(13).

Нетрудно видеть, что поскольку функции f и h выпуклы, множество точек, в которых ψ_ε недифференцируема, имеет меру 0. Следуя [7], определим почти градиент функции ψ_ε, в точке x как предел некоторой последовательности градиентов $\{\nabla\psi_\varepsilon(x^k)\}_{k=1}^{\infty}$, где $\{x^k\}_{k=1}^{\infty}$ — последовательность точек, сходящаяся к x и такая, что во всех точках этой последовательности функция ψ_ε дифференцируема. В качестве приближения к почти градиенту в точке x можно взять градиент $\nabla\psi_\varepsilon(x^k)$ в точке x^k, достаточно близкой к x.

Для решения задачи (4) может применяться любой алгоритм минимизации выпуклых функций. Конкретная реализации алгоритмов требует решения многих вопросов. Необходимо определять и уточнять итерационно значение ε, чтобы обеспечивалась выпуклость функции, и соотношение $\gamma_\varepsilon(x) > 0$. Понятно, что точку x^0, относительно которой формируются функции φ, ψ_ε, можно изменять в ходе решения, но это требует дополнительного оценивания параметра ε.

Автор выражает глубокую благодарность Н.Г. Журбенко и Т.А. Бардадым за полезные обсуждения и ценные советы.

СПИСОК ЛИТЕРАТУРЫ

1. Шор Н.З. Методы минимизации недифференцируемых функций и их приложения. — Киев: Наук. думка, 1979. — 199 с.
2. Shor N.Z. Nondifferentiable Optimization and Polynomial Problems. — London: Kluwer Academ. Publ., 1998. — 381 p.
3. Демьянов В.Ф., Васильев Л.В. Недифференцируемая оптимизация. – М.: Наука, 1981. — 384 с.
4. Ржевский С.В. Монотонные методы выпуклого программирования. — Киев: Наук. думка, 1993. — 324 с.
5. Лаптин Ю.П., Журбенко Н.Г., Левин М.М., Волковицкая П.И. Использование средств оптимизации в системе автоматизированного проектирования энергетических котлоагрегатов КРОКУС // Энергетика и электрификация. — 2003. — № 7. — С. 41–51.
6. Лаптин Ю.П., Журбенко Н.Г. Некоторые вопросы решения блочных нелинейных задач оптимизации со связывающими переменными // Кибернетика и системный анализ. — 2006. — № 2. — С. 47–55.
7. Шор Н.З. О классе почти дифференцируемых функций и одном методе минимизации функций этого класса // Кибернетика. — 1972. — № 4. — С. 9–17.

Поступила 20.11.2008