Выпуск 29(46) ОПТИМИЗАЦИЯ 1982 г.
Численные методы

УДК 519.3 + 512.25/26

ОЦЕНКИ ПО АРГУМЕНТУ СКОРОСТИ СХОДИМОСТИ МЕТОДА ШТРАФОВ

А.А.Каплан

В качестве исходной рассматривается экстремальная задача:

$$f(x) \to \min ! \tag{1}$$

на множестве

$$\Omega = \{ x \in R^n : g^j(x) \leq 0, \quad j = \overline{1,m} \}. \tag{2}$$

Здесь $f, g^j \in C^1(R^n)$ - выпуклые функции, причем выполнено условие регулярности Слейтера.

Используя для учета ограничений (2) выпуклые функции штрафа Φ_k вида

$$\Phi_k(x) = \sum_{j=1}^{m} \psi_k(g^j(x)), \tag{3}$$

дифференцируемые в $dom\, \Phi_k$, определяем приближенные решения $x_\varepsilon^k \in dom\, \Phi_k$ вспомогательных задач по критерию

$$\| \nabla f(x_\varepsilon^k) + \nabla \Phi_k(x_\varepsilon^k) \| \leq \varepsilon_k , \tag{4}$$

где $\{\varepsilon_k\}$ - последовательность положительных чисел с пределом в 0, $\|\cdot\|$ - евклидова норма вектора в R^n.

В предположении, что множество $\bar{X} = \{x \in \Omega : f(x) = \inf_{z \in \Omega} f(z) \equiv \bar{f}\}$ непусто и ограничено, в работах [2, 4-6] для ряда используемых на практике функций штрафа установлены неулучшаемые по порядку оценки близости $f(x_\varepsilon^k)$ к $\bar{f}$. Если f - сильно выпуклая функция, т.е.

$$f\left(\frac{x+z}{2}\right) \le \frac{1}{2}\left[f(x)+f(z)\right]-\mu|x-z|^2 \quad (\mu>0) \tag{5}$$

и $x_\varepsilon^k \in Q$, на их основе с помощью неравенства

$$|x_\varepsilon^k - \bar{x}|^2 \le \frac{1}{4\mu}\left[f(x_\varepsilon^k)-\bar{f}\right],$$

являющегося следствием (5), в свою очередь получаются неулучшаемые по порядку (в рассматриваемом классе задач) оценки близости x_ε^k к искомому решению $\bar{x}$.

Однако при более ограничительных предположениях относительно исходной задачи для $|x_\varepsilon^k - \bar{x}|$ могут быть получены оценки того же порядка, что и для величины $f(x_\varepsilon^k)-\bar{f}$.

Такого рода результат в случае функций*)

$$\psi_k(s) = \begin{cases} -\tau_k \ln(-s) & \text{при } s<0, \\ +\infty & \text{при } s \ge 0, \end{cases} \tag{6}$$

$$\psi_k(s) = \begin{cases} -\tau_k^2 s^{-1} & \text{при } s<0, \\ +\infty & \text{при } s \ge 0, \end{cases} \tag{7}$$

$$\psi_k(s) = \tau_k^{-1} \max{}^2 [0, s] \tag{8}$$

при условиях, гарантирующих дифференцируемость оптимальных траекторий в соответствующем непрерывном процессе, следует из теорем о гладкости, установленных А.Фиакко, Г.Мак-Кормиком [I] Ф.Лоотсма [2]. Р.Миффлин [3] для функций**) (6) и (8) получил оценки того же порядка , что и в [I,2], но при несколько более слабых предположениях гладкости функций f и g и при условии, что вспомогательные задачи решаются приближенно. Разработанная им техника дает возможность установить аналогичные оценки для существенно более широкого класса функций штрафа.

---

*) Везде $\tau_k > 0$, $\lim\limits_{k\to\infty} \tau_k = 0$. Говоря далее о непрерывном аналоге метода штрафов, имеем в виду, что τ_k заменено на $\tau > 0$ и $\tau \to 0$ непрерывно.

**) Точнее, в случае, когда часть ограничений учтена с помощью функции (6), а для остальных использована функция (8).

В данной работе предлагается иной подход, позволяющий при использовании некоторых функций штрафа избавиться от существенного в [I-3] требования единственности решения двойственной задачи. В случае линейных ограничений соответствующие оценки получаются на этом пути почти для всех известных в литературе гладких функций штрафа типа (3).

В дальнейшем нам потребуется следующее легко проверяемое утверждение.

ЛЕММА. Пусть линейная функция $\varphi(x) = (c, x)$ неположительна на множестве решений системы неравенств

$$(d^j, x) \geq 0, \quad j = \overline{1,m},$$

и существует положительный вектор $u^* = (u_1^*, \ldots, u_m^*)$ такой, что $\sum_{j=1}^{m} u_j^* d^j = -c$.

Тогда множество решений задачи линейного программирования

$$\varphi(x) - max!$$

при ограничениях

$$(d^j, x) \geq 0, \quad j = \overline{1,m},$$

является линейным многообразием.

Рассматривая задачу (I)-(2) при дополнительном условии сильной выпуклости функции f, будем далее предполагать, что в множестве Y векторов Лагранжа, отвечающих единственному решению $\bar{x}$ этой задачи, найдется вектор $\bar{y}$, для которого

$$\bar{y}_j > 0 \text{ при } j \in J_0(\bar{x}) \equiv \{j' : g^{j'}(\bar{x}) = 0\}; \; J_0(\bar{x}) \neq O. \qquad (9)$$

На основании признака оптимальности для задачи выпуклого программирования имеем

$$-\sum_{j \in J_0(\bar{x})} \bar{y}_j \nabla g^j(\bar{x}) = \nabla f(\bar{x}), \qquad (10)$$

причем для любого вектора $s \in R^n$, удовлетворяющего условиям

$$-\nabla g^j(\bar{x}) s \geq 0, \quad j \in J_0(\bar{x}),$$

должно выполняться неравенство

$$-\nabla f(\bar{x}) s \leq 0.$$

Пользуясь леммой , отсюда заключаем, что множество T решений

задачи

$$-\nabla f(\bar{x})s - max!,$$
$$-\nabla g^j(\bar{x})s \geq 0, \quad j \in J_0(\bar{x}), \tag{II}$$

является линейным многообразием.

Пусть $T^+ \perp T, \bar{x} \in T^+, \; T \oplus (T^+_{-\bar{x}}) = R^n$. Обозначим

$$z^k = \arg\min_{z \in T^+_{-\bar{x}}} |z - x^k_\varepsilon|,$$

где x^k_ε определяется по критерию (4), а выбор функций штрафа будет уточнен позднее.

ТЕОРЕМА. Пусть в дополнение к принятым выше выполнены следующие условия:

I) на выпуклом множестве G, содержащем Ω и последовательность $\{z^k\}$, при некотором L справедливо неравенство

$$|\nabla f(x) - \nabla f(z)| \leq L|x - z|;$$

2) функции g^j - аффинные, либо

2′) g^j - выпуклые функции; существуют константы L_j такие, что при $x, z \in \Omega$ имеет место

$$|\nabla g^j(x) - \nabla g^j(z)| \leq L_j |x - z|;$$

$$\psi_k(s) = \frac{a}{2}(s + \sqrt{s^2 + \tau_k}), \tag{I2}$$

причем[ж)] $a > 2 \max_{\bar{y} \in \bar{Y}} \max_{1 \leq j \leq m} \bar{y}_j$;

3) начиная с некоторого номера,

$$\sum_{j \notin J_0(\bar{x})} [\psi_k(g^j(x^k_\varepsilon)) - \psi_k(g^j(z^k))] \geq -\zeta(\tau_k),$$

где ζ - фиксированная функция, $\zeta(\tau_k) > 0$ при $\tau_k > 0$;

4) $\overline{\lim\limits_{k \to \infty}} \dfrac{\varepsilon_k^2}{\zeta(\tau_k)} \leq d < +\infty$;

5) существует K такое, что

ж) Как известно, при выполнении условия Слейтера, множество $\bar{Y}$ является ограниченным.

$x_\varepsilon^k \in \operatorname{int}\Omega$ при $k \geq K$.

Тогда при некоторых $c_1 > 0$, $c_2 > 0$ для больших K имеет место

$$\|x_\varepsilon^k - \bar{x}\| \leq \max\{c_1(f(x_\varepsilon^k) - f(\bar{x})),\ c_2\sqrt{\zeta(\tau_k)}\}. \tag{13}$$

ДОКАЗАТЕЛЬСТВО. Проверим вначале, что

$$\varliminf_{k\to\infty} \frac{\nabla f(\bar{x})(z^k - \bar{x})}{\|z^k - \bar{x}\|} > 0. \tag{14}$$

Не ограничивая общности, можно считать, что $\|\bar{x}\| \neq 1$. Ввиду (10) справедливо соотношение

$$\frac{\nabla f(\bar{x})(z^k - \bar{x})}{\|z^k - \bar{x}\|} = -\sum_{j \in J_0(\bar{x})} \bar{y}_j \frac{\nabla g^j(\bar{x})(z^k - \bar{x})}{\|z^k - \bar{x}\|}, \tag{15}$$

причем, так как $z^k - x_\varepsilon^k \in T$, то

$$\nabla f(\bar{x})(z^k - \bar{x}) = \nabla f(\bar{x})(x_\varepsilon^k - \bar{x}), \tag{16}$$

$$\nabla g^j(\bar{x})(z^k - \bar{x}) = \nabla g^j(\bar{x})(x_\varepsilon^k - \bar{x}),\quad j \in J_0(\bar{x}), \tag{17}$$

и с учетом условия 5) при больших k должно быть

$$\nabla f(\bar{x})(z^k - \bar{x}) > 0, \tag{18}$$

$$\nabla g^j(\bar{x})(z^k - \bar{x}) < 0,\quad j \in J_0(\bar{x}). \tag{19}$$

Если (14) не выполняется, то из (18) следует

$$\varliminf_{k\to\infty} \frac{\nabla f(\bar{x})(z^k - \bar{x})}{\|z^k - \bar{x}\|} = 0, \tag{20}$$

а из (9),(15) и (19) -

$$\varlimsup_{k\to\infty} \frac{\nabla g^j(\bar{x})(z^k - \bar{x})}{\|z^k - \bar{x}\|} = 0,\quad j \in J_0(\bar{x}). \tag{21}$$

Но последовательность $\left\{\frac{z^k - \bar{x}}{\|z^k - \bar{x}\|}\right\}$ имеет предельную точку $\bar{s}$, $\|\bar{s}\| = 1 \neq \|\bar{x}\|$, которая ввиду (15), (20) и (21) удовлетворяет соотношениям

$$\nabla f(\bar{x})\bar{s} = 0,\quad \nabla g^j(\bar{x})\bar{s} = 0 \quad (j \in J_0(\bar{x}))$$

и, значит, принадлежит T. С другой стороны, так как $\frac{z^k - \bar{x}}{\|z^k - \bar{x}\|} \in T^+$ при любом k, то $\bar{s} \in T^+$, что противоречит определению T^+.

СЛУЧАЙ I: выполнено условие 2).

С учетом условия I) имеем

$$f(z^k) \le f(\bar{x}) + \nabla f(z^k)(z^k - \bar{x}) \le$$
$$\le f(\bar{x}) + \nabla f(\bar{x})(z^k - \bar{x}) + L\, \| z^k - \bar{x} \|^2 , \tag{22}$$

так что при $F_k = f + \Phi_k$

$$F_k(z^k) \le f(\bar{x}) + \nabla f(\bar{x})(z^k - \bar{x}) +$$
$$+ L\, \| z^k - \bar{x} \|^2 + \sum_{j=1}^{m} \Psi_k (g^j(z^k)) \equiv V_k , \tag{23}$$

а ввиду сильной выпуклости f

$$F_k(x_\varepsilon^k) - \varepsilon_k^2/8\mu \ge f(\bar{x}) + \nabla f(\bar{x})(x_\varepsilon^k - \bar{x}) +$$
$$+ 4\mu\, \| x_\varepsilon^k - \bar{x} \|^2 + \sum_{j=1}^{m} \Psi_k (g^j(x_\varepsilon^k)) - \varepsilon_k^2/8\mu \equiv V_k' . \tag{24}$$

Так как $z^k - x_\varepsilon^k \in T$, а функции g^j аффинные, то $g^j(z^k) = g^j(x_\varepsilon^k)$ при $j \in J_0(\bar{x})$ и из (23),(24) и условия 3) следует, что

$$V_k' - V_k \ge 4\mu\, \| x_\varepsilon^k - \bar{x} \|^2 - L\, \| z^k - \bar{x} \|^2 - \zeta(z_k) - \frac{\varepsilon_k^2}{8\mu} . \tag{25}$$

СЛУЧАЙ II: выполнено условие 2').

На основании теоремы 3 из [7] легко показать, что из неравенства $a > 2 \max_{\bar{y} \in Y} \max_{1 \le j \le m} \bar{y}_j$ следует условие 5).

При $j \in J_0(\bar{x})$ с учетом выпуклости функций g^j справедливо

$$g^j(x_\varepsilon^k) \ge \nabla g^j(\bar{x})(x_\varepsilon^k - \bar{x}),$$
$$g^j(z^k) \le \nabla g^j(\bar{x})(z^k - \bar{x}) + L_j\, \| z^k - \bar{x} \|^2 .$$

Обозначим для краткости

$$d_k^j = \nabla g^j(\bar{x})(x_\varepsilon^k - \bar{x}) = \nabla g^j(\bar{x})(z^k - \bar{x}),$$
$$\tilde{d}_k^j = L_j\, \| z^k - \bar{x} \|^2 .$$

при $j \in J_0(\bar{x})$ и $k \ge K$ (K определено в 5)) имеем $d_k^j < 0$, $\tilde{d}_k^j \ge 0$, откуда на основании монотонности функций $\Psi_k(s) = \frac{a}{2}(s + \sqrt{s^2 + z_k})$ следует, что

$$\Psi_k(g^j(z^k)) - \Psi_k(g^j(x_\varepsilon^k)) \le$$

$$\leqslant \frac{1}{2} a\,(\tilde{d}_k^j + \sqrt{(d_k^j+\tilde{d}_k^j)^2+\tau_k} - \sqrt{(d_k^j)^2+\tau_k}\,) =$$

$$= \frac{1}{2} a\,(\tilde{d}_k^j + \frac{2d_k^j\tilde{d}_k^j + (\tilde{d}_k^j)^2}{\sqrt{(d_k^j+\tilde{d}_k^j)^2+\tau_k} + \sqrt{(d_k^j)^2+\tau_k}}\,)\,.$$

Из последнего неравенства в случае $2d_k^j + \tilde{d}_k^j \leqslant 0$ получаем

$$\psi_k(g^j(z^k)) - \psi_k(g^j(x_\varepsilon^k)) \leqslant \frac{1}{2} a \tilde{d}_k^j,$$

а при $2d_k^j + \tilde{d}_k^j > 0$ -

$$\psi_k(g^j(z^k)) - \psi_k(g^j(x_\varepsilon^k)) \leqslant$$

$$\leqslant \frac{1}{2} a \left[\tilde{d}_k^j + \frac{(\tilde{d}_k^j)^2}{\tilde{d}_k^j + d_k^j + |d_k^j|}\right] = a\tilde{d}_k^j.$$

Тем самым, при $k \geqslant K$ справедливо неравенство

$$\sum_{j\in J_0(\bar{x})} [\psi_k(g^j(z^k)) - \psi_k(g^j(x_\varepsilon^k))] \leqslant a\,|z^k-\bar{x}|^2 \sum_{j\in J_0(\bar{x})} L_j\,. \qquad (26)$$

На основании соотношений (23) и (24), используя (26) и условие 3), имеем

$$\nu_k' - \nu_k \geqslant 4\mu\,|x_\varepsilon^k - \bar{x}|^2 - (L + a\sum_{j\in J_0(\bar{x})} L_j)|z^k-\bar{x}|^2 -$$
$$- \zeta(\tau_k) - \varepsilon_k^2/8\mu\,. \qquad (25')$$

Пусть $\hat{c}$ - достаточно большое число и

$$\mathcal{R} = \left\{k : \frac{|x_\varepsilon^k - \bar{x}|^2}{\zeta(\tau_k)} \geqslant \hat{c}\right\}. \qquad (27)$$

Оценка (I3) очевидна, если $\mathcal{R}$ состоит из конечного числа элементов. Если же $\mathcal{R}$ - бесконечное множество, то должно выполняться неравенство

$$\varliminf_{k\in\mathcal{R}} \frac{\nabla f(\bar{x})(x_\varepsilon^k - \bar{x})}{|x_\varepsilon^k - \bar{x}|} > 0\,. \qquad (28)$$

Действительно, в случае $\varliminf\limits_{k\in\mathcal{R}} \frac{\nabla f(\bar{x})(x_\varepsilon^k-\bar{x})}{|x_\varepsilon^k-\bar{x}|} = 0$ на основании неравенства (I4) имели бы $\varliminf\limits_{k\in\mathcal{R}} \frac{|z^k-\bar{x}|}{|x_\varepsilon^k-\bar{x}|} = 0$.
Используя последнее соотношение, с учетом определения $\mathcal{R}$ и условия 4) из (25) (либо из (25′)) получаем, что при некото-

ром $k_0 \in \tilde{K}$ должно выполняться $v'_{k_0} - v_{k_0} > 0$. А значит,

$$F_{k_0}(x_\varepsilon^{k_0}) - \frac{\varepsilon_{k_0}^2}{8\mu} > F_{k_0}(z^{k_0}). \qquad (29)$$

Но, как известно (см., например,[5]), из (4) и сильной выпуклости f при любом k следует неравенство

$$F_k(x_\varepsilon^k) \leq \min_{x \in R^n} F_k(x) + \frac{\varepsilon_k^2}{8\mu},$$

т.е. (29) противоречит выбору $x_\varepsilon^{k_0}$. Ввиду соотношения (28) при некотором $c_1 > 0$ и больших $k \in \tilde{K}$ должно быть

$$\nabla f(\bar{x})(x_\varepsilon^k - \bar{x}) \geq c_1^{-1} \| x_\varepsilon^k - \bar{x} \|,$$

откуда

$$f(x_\varepsilon^k) - f(\bar{x}) \geq c_1^{-1} | x_\varepsilon^k - \bar{x} |. \qquad (30)$$

Если $k \notin \tilde{K}$, то

$$\| x_\varepsilon^k - \bar{x} \| < c_2 \sqrt{\zeta(\tau_k)}, \qquad (31)$$

где $c_2 = \sqrt{\bar{c}}$, и остается объединить неравенства (30),(31).∎

Оценка величины

$$\delta_k = \sum_{j \notin J_0(\bar{x})} [\Psi_k(g^j(z^k)) - \Psi_k(g^j(x_\varepsilon^k))]$$

для различных функций Ψ_k дает следующие результаты.

В случае, когда Ψ_k задана соотношением (6), то для $j \notin J_0(\bar{x})$

$$\Psi_k(g^j(z^k)) - \Psi_k(g^j(x_\varepsilon^k)) = \tau_k \ln\left(\frac{-g^j(x_\varepsilon^k)}{-g^j(z^k)}\right) =$$

$$= \tau_k \ln\left(1 + \frac{g^j(z^k) - g^j(x_\varepsilon^k)}{-g^j(z^k)}\right) \leq$$

$$\leq \tau_k \ln\left(1 + \frac{\|\nabla g^j(z^k)\| \| x_\varepsilon^k - z^k \|}{-g^j(z^k)}\right) \leq$$

$$\leq \tau_k \ln\left(1 + \frac{\|\nabla g^j(z^k)\| | x_\varepsilon^k - \bar{x} |}{-g^j(z^k)}\right)$$

и при $c = 2 \max\limits_{j \notin J_0(\bar{x})} \frac{\|\nabla g^j(\bar{x})\|}{-g^j(x)}$ для больших k имеем

$$\delta_k \leq \bar{c} m | x_\varepsilon^k - \bar{x} | \tau_k,$$

т.е. условие 3) выполнено, если $\zeta(\tau_k) = \bar{c} m | x_\varepsilon^k - \bar{x} | \tau_k$.

Учитывая, что (см., например,[5, §3.6]) для $f(x_\varepsilon^k) - f(\bar{x})$ имеет место оценка вида

$$f(x_\varepsilon^k) - f(\bar{x}) \leq c z_k + \frac{\varepsilon_k^2}{8\mu},$$

и выбирая $\{\varepsilon_k\}$ с соблюдением условия 4), согласно (I3) получаем, что при некотором $\tilde{c}$ и больших k должно быть

$$\| x_\varepsilon^k - \bar{x} \| \leq \tilde{c} z_k.$$

Совсем просто проверяется, что в случае функции $\varphi(s) = \exp(+z_k^{-1} s)$ можно взять $\zeta(z_k) = m \exp(-z_k^{-1}\bar{c})$, где $\bar{c} = \frac{1}{2}\min_{j \notin J_0(\bar{x})} |g^j(\bar{x})|$, а в случае функции (7) годится $\zeta(z_k) = m\bar{c} z_k^2$, где $\bar{c} = 2\max_{j \notin J_0(\bar{x})} [-g^j(\bar{x})]^{-1}$, т.е. и здесь оценки близости x_ε^k к $\bar{x}$ имеют тот же порядок, что и соответствующие оценки для $f(x_\varepsilon^k) - f(\bar{x})$ в [5, §4.3].

Наконец, если в качестве φ выбрана функция (I2), то $\zeta(z_k) = \frac{1}{2}\alpha\bar{c} m z_k$, где $\bar{c} = \max_{j \notin J_0(\bar{x})} [-g^j(\bar{x})]^{-1}$.

Используя рассмотренный в §4.3 работы [5] подход, нетрудно установить, что в этом случае имеем место оценка вида

$$f(x_\varepsilon^k) - f(\bar{x}) \leq c_1 z_k^{1/2} + \frac{\varepsilon_k^2}{8\mu}.$$

На основании неравенства (I3) заключаем, что величина $|x_\varepsilon^k - x|$ также имеет порядок $z_k^{1/2}$.

ЛИТЕРАТУРА

I. FIACCO A.V., MCCORMICK G.P. Nonlinear programming: sequential unconstrained minimization techniques.- N.Y: Wiley, 1968. /Русск. перевод: Нелинейное программирование: методы последовательной безусловной минимизации. - М.: Мир, I972.

2. LOOTSMA F.A. Boundary propreties of penalty functions for constrained minimization.- Philips Res.Rept., Suppl.3, 1970.

3. MIFFLIN R. Convergence bounds for nonlinear programming algorithms. - Math. Programming, 1975, v.8, p.251-271 (см. также Administrative Sci.Tech.Rept., N 57, Yale Univ., 1972).

4. ЕРЕМИН И.И. Метод штрафов в выпуклом программировании. - Докл. АН СССР, I967, т.I73, №4, с.748-756.

5. ГРОССМАН К., КАПЛАН А.А. Нелинейное программирование на основе безусловной минимизации. - Новосибирск: Наука, I98I.

6. КАПЛАН А.А. О скорости сходимости метода штрафов. - Докл. АН СССР, 1976, т.229, №2, с.288-291.
7. КАПЛАН А.А. Об одном подходе к решению задач выпуклого программирования. - Докл. АН СССР, 1981, т.258, № 4, с.785-788.

Поступила в ред.-изд. отдел
25.10.1981 г.