УДК 512.25/26

К АНАЛИЗУ НЕКОТОРЫХ ОЦЕНОК СКОРОСТИ СХОДИМОСТИ МЕТОДА ШТРАФОВ

А.А.Каплан

Пусть f, $g^j (j \in J = \{1, 2, \dots, m\})$ - выпуклые непрерывно дифференцируемые в R^n функции, и для решения задачи

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

при ограничениях

$$g^j(x) \le 0, \; j \in J, \tag{2}$$

применяется метод штрафов. В предположении, что ограничения удовлетворяют условию Слейтера, а множество решений исходной задачи непусто и ограничено, исследуется вопрос об оценке быстроты сходимости метода при использовании следующих функций штрафа:

$$\varphi_k^1(x) = \sum_{j \in J} e^{t_k^{-1} g^j(x)}, \tag{3}$$

$$\varphi_k^2(x) = -t_k \sum_{j \in J} \frac{1}{g^j(x)}, \tag{4}$$

$$\varphi_k^3(x) = t_k^{-1} \sum_{j \in J} \left(g^j(x) + \sqrt{g^{j^2}(x) + t_k^{2+\theta}} \right). \tag{5}$$

Здесь $\{t_k\}$ – последовательность положительных чисел, $\lim t_k = 0$, $\theta \geq 0$ – фиксированное число.

Определим $\Omega = \{x \in R^n: g^j(x) \leq 0, j \in J\}$, $F_k^i(x) = f(x) + \varphi_k^i(x)$, $x^{k,i} = \arg\min_{x \in R^n} F_k^i(x)$ при $i = 1,3$ и $x^{k,2} = \arg\min_{x \in \operatorname{int}\Omega} F_k^2(x)$. Системе функций $\{\varphi_k^i\}$, $k = 1,2,\ldots$, сопоставим функцию V^i, получающуюся при замене t_k непрерывно изменяющимся параметром $t \in R^1_+$. Соответственно $F^i(x,t) = f(x) + V^i(x,t)$, $x^i(t) = \arg\min_{x \in R^n(\operatorname{int}\Omega)} F^i(x,t)$.

В [1,2] для некоторых классов функций штрафа исследован вопрос о дифференцируемости траектории безусловных минимумов в полуокрестности точки $t = 0$. При этом предполагалось, что относительно седловой точки $(\bar{x}, \bar{y})$ исходной задачи выполнены следующие условия:

1) множители $\bar{y}_j$, $j \in J(\bar{x}) = \{l \in J: g^l(\bar{x}) = 0\}$, положительны;

2) векторы $\nabla g^j(\bar{x})$, $j \in J(\bar{x})$, линейно-независимы;

3) для любого $z \in R^n$, $z \neq 0$, такого, что $\nabla g^j(\bar{x}) z = 0$, имеет место $(L''_{x^2}(\bar{x},\bar{y})z, z) > 0$, где L – функция Лагранжа.

Отметим, что из 1)–3) следует единственность седловой точки.

В частности, в [1,2] устанавливается, что при использовании функции

$$V^4(x,t) = -t \sum_{j \in J} \ln(-g^j(x))$$

и

$$V^5(x,t) = -t^2 \sum_{j \in J} \frac{1}{g^j(x)}$$

траектории $(x^i(t), y^i(t))$, $i = 4,5$, где $y^i(t) = (y_1^i(t), \ldots, y_m^i(t))$, а $y_j^4(t) = \frac{t}{g^j(x^4(t))}$, $y_j^5(t) = \frac{t^2}{g^j(x^5(t))}$, единственным образом определены и дифференцируемы в полуокрестности точки $t = 0$. Отсюда легко следует, что

$$f(x^i(t)) - f(\bar{x}) = c_i t + O_i(t), \tag{6}$$

причем $c_4 = |J(\bar{x})|$, $c_5 = \sum \bar{y}_j^{1/2}$, $|J(\bar{x})|$ – число элементов множества $J(\bar{x})$.

Из сопоставления функций $F^2(x,t)$ и $F^5(x,t)$ сразу получаем $x^2(t) = x^5(\sqrt{t})$ и на основании (6) при малых t, t_k

$$f(x^2(t))-f(\bar{x})=\sum_j \bar{y}_j^{1/2}\sqrt{t}+o(\sqrt{t}),$$

$$f(x^{k,2})-f(\bar{x})=\sum_j \bar{y}_j^{1/2}\sqrt{t_k}+o(\sqrt{t_k}). \qquad (7)$$

Отметим, что в отличие от $x^5(t)$ траектория $x^2(t)$, как правило, не дифференцируема в точке $t=0$.

Формула (7) показывает, что при использовании функций штрафа (4) порядок сходимости $1/2$ не может быть улучшен ни при каком усилении требований относительно исходной задачи (исключая случай, когда $\bar{x}\in int\Omega$). Покажем, однако, что тот же порядок $1/2$ имеет место и для существенно более широкого класса задач, именно: условия I-3 могут быть отброшены.

ТЕОРЕМА I. Пусть $\tilde{x}\in int\Omega$ и $\tau\in(0,1)$ фиксированы, $\sigma=\min_{j\in J}|g^j(\tilde{x})|$. Тогда, если $t_k^{1-\tau}/\sigma<1$, имеем

$$f(x^{k,2})-f(\bar{x})<\frac{t_k^{1-\tau}}{\sigma}(f(\tilde{x})-f(\bar{x}))+mt_k^{\tau}. \qquad (8)$$

ДОКАЗАТЕЛЬСТВО. Положим $\lambda_k=\frac{t_k^{1-\tau}}{\sigma}$. При $\lambda_k<1$, $x(\lambda_k)=$ $=\bar{x}+\lambda_k(\tilde{x}-\bar{x})$ на основании выпуклости функций g^j имеем

$$g^j(x(\lambda_k))\le\lambda_k g^j(\tilde{x})+(1-\lambda_k)g^j(\bar{x})\le-\lambda_k|g^j(\tilde{x})|\le-t_k^{1-\tau}.$$

Отсюда

$$\varphi_k^2(x(\lambda_k))=-t_k\sum_{j\in J}\frac{1}{g^j(x(\lambda_k))}\le mt_k^{\tau}$$

и, следовательно,

$$f(x(\lambda_k))+\varphi_k^2(x(\lambda_k))\le\frac{t_k^{1-\tau}}{\sigma}f(\tilde{x})+\left(1-\frac{t_k^{1-\tau}}{\sigma}\right)f(\bar{x})+mt_k^{\tau}. \qquad (9)$$

Так как $f(x(\lambda_k))+\varphi_k^2(x(\lambda_k))\ge f(x^{k,2})+\varphi_k^2(x^{k,2})$ и $\varphi_k^2(x^{k,2})>$ >0, из (9) получаем, что

$$f(x^{k,2})<\frac{t_k^{1-\tau}}{\sigma}f(\tilde{x})+\left(1-\frac{t_k^{1-\tau}}{\sigma}\right)f(\bar{x})+mt_k^{\tau},$$

т.е. (8) выполнено.

Ясно, что, имея в виду малые значения t_k, выгодно брат

$\tau = 1/2$, и при этом оценка (8) дает порядок сходимости $1/2$. Для определения констант σ и $f(\tilde{x})$ в качестве $\tilde{x}$ можно взять начальную точку итерационного процесса, которая при использовании функции штрафа (4) должна принадлежать $int\Omega$. Таким образом, чтобы воспользоваться неравенством (8), нужно лишь уметь с удовлетворительной точностью оценивать снизу величину $f(\bar{x})$. Как правило, это существенно проще, чем оценка константы $\sum_j \bar{y}_j^{1/2}$ и функции $O(\sqrt{t_k})$ в (7).

Простой и, на первый взгляд, грубый способ построения априорных оценок, избранный при доказательстве теоремы I, на самом деле позволяет получить неулучшаемые по порядку оценки и для некоторых других используемых на практике функций штрафа, в частности для функций (3) и (5).

ТЕОРЕМА 2. Пусть $\tilde{x} \in int\Omega$, $\sigma = \min_{i\in J} |g^i(\tilde{x})|$.

I. Если при фиксированных $\tau \in (-\theta, 1)$ и k имеет место $x^{k,3} \in \Omega$ и $\frac{t_k^{1-\tau}}{\sigma} < 1$, то

$$f(x^{k,3}) - f(\bar{x}) < \frac{t_k^{1-\tau}}{\sigma}(f(\tilde{x}) - f(\bar{x})) + \frac{m}{2} t_k^{\tau+\theta}. \quad \text{(I0)}$$

2. Если при фиксированном k выполнено $x^{k,1} \in \Omega$ и $-\frac{t_k \ln t_k}{\sigma} < 1$, то

$$f(x^{k,1}) - f(\bar{x}) < -\frac{t_k \ln t_k}{\sigma}(f(\tilde{x}) - f(\bar{x})) + m t_k. \quad \text{(II)}$$

3. Включение $x^{k,3} \in \Omega$ ($x^{k,1} \in \Omega$) имеет место, начиная с некоторого номера K_3 (K_1).

Отметим, что $\max_{\tau} \min\{1-\tau, \tau+\theta\} = \frac{1+\theta}{2}$, т.е. (I0) гарантирует порядок сходимости $\frac{1+\theta}{2}$.

Утверждения I и 3 теоремы 2 доказаны в [3]. Утверждение 2 нетрудно проверить по той же схеме, если в качестве λ_k взять $-\frac{t_k \ln t_k}{\sigma}$. Для использования неравенств (I0) и (II) в качестве априорных оценок требуется предварительно определить числа K_3 и K_1 соответственно. Теорема 4° из [3] дает значение K_3 в случае $\theta = 0$. Аналогичный результат нетрудно получить и для определения K_1.

Приведем пример, показывающий, что порядок сходимости в

(10) и (11) не может быть улучшен.

ПРИМЕР. Минимизировать $f(x)=\frac{1}{4}(x-1)^2$ при ограничении $x \le 0$.

Очевидно, что $\bar{x}=0$, $\bar{y}=1/2$ - единственная седловая точка соответствующей функции Лагранжа.

При использовании функции штрафа (3) имеем

$$F_k^1(x)=\frac{1}{4}(x-1)^2+e^{t_k^2}x.$$

Если k достаточно велико и $z^k=t_k \ln t_k$, то $\nabla F_k^1(z_k)=-\frac{1}{2}(1-t_k \ln t_k)+1>0$, следовательно, $x^k<t_k \ln t_k$, и так как функция f на участке $(-\infty,0]$ строго убывает, то

$$f(x^k)-f(\bar{x})>f(z^k)-f(\bar{x})=\frac{1}{4}(t_k \ln t_k-1)^2-1/4=$$

$$=-\frac{1}{2}t_k \ln t_k+\frac{1}{4}(t_k \ln t_k)^2.$$

Аналогично, при использовании функции штрафа (5), если k велико, в точке $z^k=-\frac{1}{2}t_k^{\frac{1+\theta}{2}}$ имеем

$$\nabla F_k^3(z^k)\geq-\frac{1}{4}t_k^{\frac{1+\theta}{2}}-\frac{1}{2}+\frac{2}{1+4t_k}>0,$$

откуда $x^k<-\frac{1}{2}t_k^{\frac{1+\theta}{2}}$ и

$$f(x^k)-f(\bar{x})>\frac{1}{16}t_k^{1+\theta}+\frac{1}{4}t_k^{\frac{1+\theta}{2}}.$$

Отметим, что для приведенного примера условия 1-4 выполнимы. Однако траектория $x^1(t)$ и $x^3(t)$ не дифференцируема справа в 0. Это можно показать и в общем случае (если только $J(\bar{x})\neq\emptyset$).

Действительно, непрерывность в 0 траектории $(x^1(t), y^1(t))$, где $y^1(t)=(y_1^1(t),\ldots,y_m^1(t))$, $y_j^1(t)=t^{-1}e^{t^{-1}f^j(x(t))}$, легко следует из теоремы 4 в [4]. Для $j\in J(\bar{x})$ имеем $\frac{g^j(x^1(t))-g^j(x^1(0))}{t}=\ln(ty_j^1(t))$. Если $x^1(t)$ дифференцируема, то существует конечный предел $\frac{g^j(x^1(t))-g^j(x^1(0))}{t}$ при $t\to+0$, но это невозможно, ибо $\lim_{t\to+0}\ln(ty_j^1(t))=-\infty$ в силу условия 1.

Аналогичные рассуждения проходят и для $x^3(t)$.

ЛИТЕРАТУРА

1. ФИАККО А., МАК-КОРМИК Г. Нелинейное программирование. Пер. с англ. М., "Мир", 1972.

2. LOOTSMA F.A. Boundary properties of penalty functions for constrained minimization. - "Philips Res. Rept.", 1970, suppl. N 3.

3. КАПЛАН А.А. К вопросу о реализации метода штрафов. - В кн.: Оптимизация. Вып. I9 (36). Новосибирск, I977, с. 58-73.
4. КАПЛАН А.А. Характеристические свойства штрафных функций. - "ДАН СССР", I973, т.2I0, № 5, с.I0I8-I02I.

Поступила в ред.-изд. отдел
I7.04.I978 г.