Доклады Академии наук СССР
1976. Том 229, № 2

УДК 512.25/26

МАТЕМАТИКА

А. А. КАПЛАН

О СКОРОСТИ СХОДИМОСТИ МЕТОДА ШТРАФОВ

(Представлено академиком Л. В. Канторовичем 5 III 1976)

Рассматривается задача минимизации выпуклой функции f на компакте $\Omega \subset R^n$, задаваемом системой неравенств

$$g^j(x) \leqslant 0, \quad j \in J = \{1, 2, \ldots, m\},$$

с выпуклыми функциями g^j. Предполагается существование такой точки $\bar{x}$, что $g^j(\bar{x}) < 0$ для всех $j \in J$.

Используя для решения приведенной задачи метод штрафов, будем строить последовательность $\{\Phi_k\}$ функций штрафа следующим образом. Пусть $\{A_k\}$ — числовая последовательность, $A_k > 0$ при всех k, $A_k \to \infty$. Обозначим $\psi_k^+(t) = A_k(t + |t|)$ и рассмотрим систему функций $\{\psi_k\}$, ψ_k: $R^1 \to R^1$, обладающих при всех k следующими свойствами:

а) ψ_k — выпуклая функция, $\psi_k \in C^1(R^1)$;

б) $\psi_k(t) \geqslant \psi_k^+(t)$ для всех $t \in R^1$;

в) $\psi_k'(t) \geqslant 0$ для всех $t \in R^1$;

г) существует функция λ_1: $R^1 \to R^1$ такая, что $\psi_k(t) - \psi_k^+(t) \leqslant \lambda_1(A_k)$ для всех $t \in R^1$, причем $\lambda_1(A_k) \to 0$;

д) $\psi_k'(0) \geqslant c_1 A_k$, $c_1 > 0 - \text{const}$;

е) при фиксированном $\delta > 0$ для $t < -\delta$ имеем $\psi_k'(t) \leqslant \lambda_2(A_k, \delta)$, где λ_2: $R^2 \to R^1$ — заданная функция, $\lambda_2(A_k, \delta) \underset{k \to \infty}{\longrightarrow} 0$, если $\delta > 0$.

Заметим, что, как следует из условий а), в) и г), при $t < -\delta$ справедливо неравенство

$$\psi_k'(t) \leqslant \frac{1}{\delta} \lambda_1(A_k)$$

и в качестве λ_2, в частности, годится функция $\frac{1}{\delta}\lambda_1(A_k)$. Однако условие е) имеет смысл сохранить как независимое, ибо от скорости стремления к 0 функции λ_2 существенно зависит характер полученных ниже оценок быстроты сходимости метода штрафов.

Рассмотрим теперь систему функций $\{\Phi_k\}$, Φ_k: $R^n \to R^1$, задаваемых соотношением

$$\Phi_k(x) = \sum_{j \in J} \psi_k(g^j(x)). \tag{1}$$

Относительно сконструированной таким образом системы будем говорить, что она обладает ψ-свойством.

Исследование свойств функций (1) с точки зрения использования их в методе штрафов существенно опирается на полученные в ([1], [2]) общие результаты о сходимости названного метода, на которые мы будем неоднократно ссылаться, имея в виду текст статьи ([1]).

Обозначим x^* решение исходной задачи, x^k — точку минимума функции $F_k = f + \Phi_k$; $J^1(x) = \{j \in J: g^j(x) = 0\}$, $J^2(x) = J \setminus J^1(x)$; $|J^1(x)|$, $|J^2(x)|$ — число элементов в множествах $J^1(x)$ и $J^2(x)$ соответственно.

Теорема 1. *Система функций* (1) *удовлетворяет всем условиям теоремы* 1 *работы* ([1]). *Если, кроме того,* $f\in C^1(R^n)$, $g^j\in C^1(R^n)$, $j\in J$, *то выполнены также условия теоремы* 4 *из* ([1]) *и существует такое* K, *что* $x^k\in\Omega$ *при* $k\geqslant K$.

Замечание 1. На основании теоремы 4 из ([1]) теперь можно заключить, что при $x^k\in\Omega$ для $f(x^*)$ имеются двухсторонние оценки

$$f(x^k)\geqslant f(x^*)\geqslant\sum_{j\in J}\psi_k'(g^j(x^k))\,g^j(x^k)+f(x^k),$$

причем

$$\sum_{j\in J}\psi_k'(g^j(x^k))\,g^j(x^k)\underset{k\to\infty}{\longrightarrow}0.$$

Точки (x^k, u^k), где $u^k=(u_1{}^k,\ldots, u_m{}^k)$, $u_j{}^k=\psi_k'(g^j(x^k))$, являются допустимыми в двойственной по Вулфу ([3]) задаче.

Теорема 2. *Пусть* $f\in C^2(R^n)$, $g^j\in C^1(R^n)$, $j\in J$, $(f''(x)\xi,\xi)\geqslant\gamma\|\xi\|^2$, $\gamma>0$, *при любых* $x\in R^n$, $\xi\in R^n$.

Тогда, если $x^k\in\Omega$, *то*

$$\|x^k-x^*\|^2\leqslant\frac{2m}{\gamma}\lambda_1(A_k),\tag{2}$$

а если, кроме того *, $g^j(x^*)<-\delta$, $g^j(x^k)<-\delta$ *при* $j\in J^2(x^*)$ *и некотором* $\delta>0$, *то*

$$\|x^k-x^*\|^2\leqslant\frac{2|J^1(x^*)|}{\gamma}\lambda_1(A_k)+\frac{2|J^2(x^*)|}{\gamma}\max_{j\in J^2(x^*)}\max_{x\in\Omega}|g^j(x)|\lambda_2(A_k,\delta).\tag{3}$$

Отметим, что пользоваться второй из приведенных оценок имеет смысл лишь в том случае, когда $\lambda_2(A_k,\ \delta)\ll\lambda_1(A_k)$ и можно достаточно точно оценить сверху величину $|J^1(x^*)|$.

Примером системы, обладающей ψ-свойством, является рассматриваемый ниже класс функций $\{\Phi_k^{(\theta)}\}$

$$\Phi_k^{(\theta)}(x)=A_k\sum_{j\in J}(g^j(x)+\sqrt{g^{j2}(x)+A_k^{-2-\theta}})\tag{4}$$

в случае, когда $\theta>0-\text{const}$.

Функции (4), очевидно, представимы в виде

$$\Phi_k^{(\theta)}(x)=\sum_{j\in J}\psi_k^{(\theta)}(g^j(x)),$$

где

$$\psi_k^{(\theta)}(t)=A_k(t+\sqrt{t^2+A_k^{-2-\theta}}),\tag{5}$$

причем имеет место

Теорема 3. *Функции* (5) *удовлетворяют условиям* а)—е). *При этом в качестве* λ_1 *и* λ_2 *в условиях* г) *и* е) *можно принять*

$$\lambda_1(A_k)=A_k^{-\theta/2},\quad \lambda_2(A_k,\delta)={}^1\!/_2\delta^{-2}A_k^{-1-\theta}.$$

Отметим, что $\psi_k^{(\theta)}\in C^\infty(R^1)$, следовательно, гладкость порожденных ими функций $\Phi_k^{(\theta)}$ определяется гладкостью функций g^j.

* На основании определения $J^2(x^*)$ существует такое $\delta>0$, что $g^j(x^*)<-\delta$, $g^j(x^k)<-\delta$ при всех k, начиная с некоторого номера.

Насколько известно автору, рассматриваемые в литературе системы функций штрафа не обладают ψ-свойством *. Интересно отметить, что функции (4) при $\theta=0$ также не обладают указанным свойством, в частности, нарушено условие г), так как $\psi_k^{(0)}(0)-\psi_k^{+}(0)=1$. Система $\{\Phi_k^{(0)}\}$, однако, удовлетворяет условиям теорем 1 и 4 из [1] (последней — в предположении о дифференцируемости функций f и g^j), а также (в отличие от системы $\{\Phi_k^{(\theta)}\}$ при $\theta>0$) условиям теоремы 3 из [1] и представляет класс функций штрафа с весьма интересными свойствами. В частности,

$$\Phi_k(x) \underset{k\to\infty}{\longrightarrow} A_k \sum_{j\in J}(g^j(x)+|g^j(x)|)$$

равномерно относительно x в области $Q_\varepsilon=\{x\in R^n: |g^j(x)|\geqslant\varepsilon, j\in J\}$ при любом фиксированном $\varepsilon>0$.

Обозначим $\sigma=\min\limits_{j\in J}|g^j(\bar{x})|$ (по определению $\bar{x}$ имеем $\sigma>0$), $\Omega_\varepsilon=\{x: g^j(x)\leqslant-\varepsilon, j\in J\}$.

Теорема 4. *Пусть* $f(\bar{x})-f(x)<(\sqrt{5}-1)/2$. *Тогда при использовании функций штрафа* $\Phi_k^{(0)}$ *для*

$$A_k>\frac{m}{2\sigma((\sqrt{5}-1)/2-f(\bar{x})+f(x^*))}$$

имеем $x^k\in\Omega_{1/2A_k}$.

Отметим, что условие $f(\bar{x})-f(x^*)<(\sqrt{5}-1)/2$ всегда может быть достигнуто за счет предварительного умножения целевой функции исходной задачи на подходящее положительное число.

Приводимая ниже теорема 5 дает оценку скорости сходимости метода штрафов с использованием системы функций $\{\Phi_k^{(\theta)}\}$ при любом фиксированном $\theta\geqslant0$. Следует отметить, что полученная здесь оценка и для случая $\theta>0$ не является следствием теорем 2, 3.

Теорема 5. *Пусть* $f\in C^2(R^n)$, $g^j\in C^1(R^n)$, $j\in J$, $(f''(x)\xi,\xi)\geqslant\gamma\|\xi\|^2$, $\gamma>0$, *при любых* $x\in R^n$, $\xi\in R^n$.

Тогда при реализации метода штрафа с использованием системы функций $\{\Phi_k^{(\theta)}\}$ ($\theta\geqslant0$ — const), *начиная с некоторого* K, *имеем*

$$\|x^k-x^*\|^2<\frac{2}{\gamma}\left[\frac{A_k^{\tau-1}}{\sigma}(f(\bar{x})-f(x^*))+\frac{m}{2}A_k^{-\tau-\theta}\right], \qquad (6)$$

где τ — *фиксированное число*, $0<\tau<1$.

Замечание 2. Область применения оценки (6) на самом деле шире, чем указано в теореме 5. Именно, оценка (6) справедлива при всех таких k, что $x^k\in\Omega$ и $A_k^{\tau-1}/\sigma<1$.

Замечание 3. Метод доказательства теоремы 5 позволяет получить оценки быстроты сходимости при использовании некоторых известных функций штрафа. В частности, если $\Phi_k(x)=\sum\limits_{j\in J}e^{A_k g^j(x)}$, в тех же условиях справедлива оценка

$$\|x^k-x^*\|^2<\frac{2}{\gamma}\left[\frac{A_k^{\tau-1}}{\sigma}(f(\bar{x})-f(x^*))+me^{-A_k^{\tau}}\right],$$

* В принципе не может существовать функций, обладающих ψ-свойством, тождественно равных 0 в допустимой области, так как в этом случае условие непрерывной дифференцируемости ψ_k и условие д) оказываются противоречивыми.

а при $\Phi_k(x)=-A_k^{-1}\sum_{j\in J}(1/g^j(x))$ — оценка

$$\|x^k-x^*\|^2<\frac{2}{\gamma}\left[\frac{A_k^{\tau-1}}{\sigma}(f(\bar{x})-f(x^*))+mA_k^{-\tau}\right].$$

При конечномерной аппроксимации некоторых вариационных задач с помощью прямых методов приходится иметь дело с ограничениями вида

$$g^j(x, y)\leqslant 0 \text{ для всех } y\in Q\subset R^l, \quad j\in J.$$

(Q — область интегрирования в исходной задаче), где g^j — непрерывные функции, выпуклые по x, причем

$$\Omega=\{x\in R^n: g^j(x, y)\leqslant 0,\ y\in Q,\ j\in J\},$$

Ω — компакт и существует точка $\bar{x}$, в которой

$$\sup_{y\in Q}\max_{j\in J} g^j(\bar{x}, y)<0.$$

Для решения полученной конечномерной задачи методом штрафов оказывается весьма удобным применить функции штрафа интегрального типа, порожденные функциями (5).

Именно, при весьма естественных для встречающихся на практике задач предположениях относительно области Q (например, если Q — компакт, удовлетворяющий условию конуса ([4])) справедлива

Теорема 6. *Функции*

$$\tilde{\Phi}_k^{(\theta)}(x)=A_k\sum_{j\in J}\int_Q (g^j(x, y)+\sqrt{g^{j2}(x, y)+A_k^{-2-\theta}})\,dy$$

удовлетворяют условиям теоремы 1 *из* ([1]). *При* $\theta>0$ *разность*

$$\tilde{\Phi}_k^{(\theta)}(x)-A_k\sum_{j\in J}\int_Q (g^j(x, y)+|g^j(x, y)|)\,dy$$

равномерно стремится к 0 *на всем* R^n, *а при* $\theta=0$ — *на любом таком компакте* S, *что* mes $\{y\in Q: g^j(x, y)=0\}=0$ *для всех* $x\in S$, $j\in J$.

Отметим, что высокая степень гладкости функций (5) позволяет (если, конечно, достаточно гладкими являются сами g^j) применить при вычислении интегралов кубатурные формулы высокой точности и, следовательно, обойтись значительно меньшим числом узлов, чем если бы мы пользовались сеткой по y для перехода к конечному числу ограничений.

Институт математики
Сибирского отделения Академии наук СССР
Новосибирск

Поступило
4 III 1976

ЛИТЕРАТУРА

[1] *А. А. Каплан*, ДАН, т. 210, № 5, 1018 (1973). [2] *А. А. Каплан*, Оптимизация, в. 8 (25), 13 (1972). [3] *Ph. Wolfe*, Quart. Appl. Math., v. 19, № 3, 239 (1961). [4] *С. Г. Михлин*, Вариационные методы в математической физике, М., 1970, стр. 284.