Доклады Академии наук СССР
1971. Том 201, № 3

УДК 518.5 *МАТЕМАТИКА*

В. В. ИВАНОВ

ОБ ОПТИМАЛЬНЫХ АЛГОРИТМАХ МИНИМИЗАЦИИ В КЛАССАХ ДИФФЕРЕНЦИРУЕМЫХ ФУНКЦИЙ

(Представлено академиком А. Н. Тихоновым 5 V 1971)

Пусть F — класс функций, определенных в кубе π_n: $0 \leqslant x_i \leqslant 1$, $i = 1, 2, \ldots, n$. Легко видеть, что любой алгоритм минимизации $f(X) \in F$, $X \in \pi_n$, использующий информацию о значениях f не более чем в N точках $X_\nu \in \pi_n$, эквивалентен (в смысле результата) некоторому алгоритму A следующего вида:

$$X_0;\ f_0 = f(X_0),\ X_1 = \varphi_1(X_0, f_0);\ f_1 = f(X_1),\ X_2 = \varphi_2(X_0, f_0; X_1, f_1); \ldots$$
$$\ldots; f_{N-2} = f(X_{N-2}),\ X_{N-1} = \varphi_{N-1}(X_0, f_0; \ldots; X_{N-2}, f_{N-2});\ f_{N-1} = f(X_{N-1}),$$
$$r_N(A, f) = s_N(X_0, f_0; \ldots; X_{N-1}, f_{N-1}),$$

где φ_ν и s_N — вычислимые функции своих аргументов.

Обозначим по аналогии с ([1])

$$v(f, N, A) = |r_N(A, f) - \inf_{X \in \pi_n} f(X)|, \quad v(F, N, A) = \sup_{f \in F} v(f, N, A),$$
$$v(F, N) = \inf_A v(F, N, A) = \inf_{X_0, \varphi_1, \ldots, \varphi_{N-1}, s_N} v(F, N, A). \tag{1}$$

Алгоритм, для которого достигается $v(F, N)$, назовем о п т и м а л ь н ы м. Условия $v(F, N, A) / v(F, N) \to 1$, $N \to \infty$, и $v(F, N, A) / v(F, N) \leqslant \mathrm{const}$, $N \to \infty$, будут означать соответственно асимптотическую оптимальность и оптимальность по порядку алгоритма A.

Обозначим еще

$$\delta(f, N, A') = \min_{0 \leqslant \nu \leqslant N-1} f(X_\nu) - \inf_{X \in \pi_n} f(X), \quad \delta(F, N, A') = \sup_{f \in F} \delta(f, N, A'),$$
$$\delta(F, N) = \inf_{A'} \delta(F, N, A') = \inf_{X_0, \varphi_1, \ldots, \varphi_{N-1}} \delta(F, N, A'). \tag{2}$$

Заметим, что в множество алгоритмов A' без итоговой операции S_N входят, в частности, многочисленные одношаговые и многошаговые конечные (с заданным числом итераций) итеративные процессы минимизации функций.

Т е о р е м а 1. *Пусть $F \equiv C^n_{s+1, L}$ — множество функций, имеющих s производных в π_n, причем s-е производные удовлетворяют в π_n условию Липшица с константой L:*

$$\left| \frac{\partial^{k_1 + \ldots + k_n} f}{\partial x_1^{k_1} \ldots \partial x_n^{k_n}}(X_1) - \frac{\partial^{k_1 + \ldots + k_n} f}{\partial x_1^{k_1} \ldots \partial x_n^{k_n}}(X_2) \right| \leqslant L \| X_1 - X_2 \|_1 = L \cdot \max_{1 \leqslant i \leqslant n} |x_i^{(1)} - x_i^{(2)}|,$$
$$\sum_{i=1}^n k_i = s.$$

Тогда справедливы соотношения

$$v(C^n_{s+1, L}, N) = O(L / N^{(s+1)/n}), \quad \delta(C^n_{s+1, L}, N) = O(L / N^{(s+1)/n}). \tag{3}$$

Доказательство. Пусть X_ν, $\nu = 0, 1, \ldots, N-1$,— произвольная фиксированная последовательность точек из π_n. Всегда найдется такая точка $X^* \in \pi_n$, для которой $\min\limits_{0 \leqslant \nu \leqslant N-1} \|X^* - X_\nu\|_1 \geqslant {}^1/_2 N^{1/n}$. Действительно, если допустить противное, то X_ν будет ε-сетью в π_n с $\varepsilon < 1/(2N^{1/n})$. Для кубической нормы это означает, что $N(2\varepsilon)^N \geqslant 1$, чего не может быть. Теперь допустим, что указанная выше последовательность точек X_ν получена алгоритмом A для функции $f_1(X) \equiv 0$, и введем функцию

$$f_2(X) = \begin{cases} 0, \ \|X - X^*\|_3 = \left(\sum\limits_{i=1}^{n} (x_i - x_i^*)^2\right)^{1/2} \geqslant 1/(2N^{1/n}), \\ -\dfrac{LN^{(s+1)/n}}{(s+1)!n} [(1/4)N^{-2/N} - \|X - X^*\|_3^2]^{s+1}, \|X - X^*\|_3 \leqslant 1/(2N^{1/n}). \end{cases} \tag{4}$$

Можно проверить, что $f_1, f_2 \in C^n_{s+1, L}$. Далее, имеем $v(F, N, A) \geqslant$ $\geqslant \max\{v(f_1, N, A), \ v(f_2, N, A)\} \geqslant L/((s+1)!2^{2s+3}N^{(s+1/n}n)$. Поэтому и $v(F, N) \geqslant L/((s+1)!2^{2s+3}N^{(s+1)/n}n)$. Для доказательства неравенства вида $v(F, N) \leqslant C_1L/N^{(s+1)/n}$ (здесь и дальше C_i будут означать константы, не зависящие от L и N), построим такую кусочно-многочленную аппроксимацию $T(f, N, X)$ любой функции $f \in C^n_{s+1, n}$, для которой

$$|f(X) - T(f, N, X)| \leqslant C_2L/N^{(s+1)/n}, \quad X \in \pi_n.$$

Следствием этого неравенства является

$$\left|\inf_{X \in \pi_n} T(f, N, X) - \inf_{X \in \pi_n} f(X)\right| \leqslant C_2L/N^{(s+1)/n}$$

и требуемое неравенство для $v(C^n_{s+1, L}, N)$ можно считать доказанным. Один из способов построения $T(f, N, X)$ состоит в следующем. Примем $N > (s+1)^n$ узлов с равномерным распределением в π_n и в окрестности каждой точки X_ν выберем $(s+1)^n$ ближайших к X_ν узлов, включая X_ν. Обозначим их через $(x_1^{(\nu, \nu_1)}, x_2^{(\nu, \nu_2)}, \ldots, x_n^{(\nu, \nu_n)})$ и положим

$$f(X) = \sum_{\nu_1=0}^{s} \ldots \sum_{\nu_n=0}^{s} f(x_1^{(\nu, \nu_1)}, \ldots, x_n^{(\nu, \nu_n)}) P_{s,1}^{(\nu, \nu_1)}(x_1) \ldots P_{s,n}^{(\nu, \nu_n)}(x_n) + R_{s, \nu}(X),$$

$$X \in \Gamma_s(X_\nu): \ \min_{\nu_r}(x_r^{(\nu, \nu_r)} - x_r^{(\nu)}) \leqslant x_r - x_r^{(\nu)} \leqslant \max_{\nu_r}(x_r^{(\nu, \nu_r)} - x_r^{(\nu)}), \tag{5}$$

где $P_{s,r}^{(\nu,\nu_r)}(x_r)$ — алгебраические многочлены степени s, $f(X) - R_{s,\nu}(X) =$ $= T_{s,\nu}(f, X)$ — интерполяционный многочлен по указанной системе узлов. Нетрудно убедиться, что для $X \in \Gamma_s(X_\nu)$ $|R_{s,\nu}(X)| \leqslant (L/N)^{(s+1)/n}) \times$ $\times (1 + L_s + \ldots + L_s^n)$, где L_s — постоянная Лебега, которая для равномерно распределенных узлов легко оценивается константой, не зависящей от N. Выбрав наименьшее число таких окрестностей $\Gamma_s(X_\nu)$, чтобы они в совокупности покрывали π_n, получим искомую более экономную (в смысле числа многочленов) кусочно-многочленную аппроксимацию $f(X)$ в π_n. Таким образом, первое из соотношений (3) доказано. Для доказательства второго соотношения следует повторить сказанное с заменой N на $N-1$, после чего положить

$$X_{N-1} = X^*_{N-1}, \ T(f, N-1, X^*_{N-1}) \leqslant \inf_{X \in \pi_n} T(f, N-1, X) + C_3L/(N-1)^{(s+1)/n}.$$

Тогда

$$0 \leqslant f(X^*_{N-1}) - \inf_{X \in \pi_n} f(X) \leqslant f(X^*_{N-1}) - T(f, N-1, X^*_{N-1}) + \inf_{X \in \pi_n} T(f, N-1, X) +$$
$$+ C_3L/(N-1)^{(s+1)/n} - \inf_{X \in \pi_n} f(X) \leqslant C_4L/(N-1)^{(s+1)/n} \leqslant C_5L/N^{(s+1)/n}.$$

Отсюда, учитывая, что в силу определения $v(F, N) \leqslant \delta(F, N)$, получаем доказательство и второго из соотношений (3).

Из способа доказательства теоремы 1 вытекает

Теорема 2. *Любой алгоритм отыскания с абсолютной погрешностью $O(L/N^{(s+1)/n})$ наименьшего в π_n значения кусочно-многочисленной функции, которая аппроксимирует $f(X)$ с абсолютной погрешностью того же порядка путем использования информации о значениях $f(X)$ в N равномерно распределенных точках π_n, является оптимальным по порядку в классе $C^n_{s+1, L}$ среди всех алгоритмов.*

Утверждения, аналогичные теоремам 1 и 2, будут, очевидно, справедливы и в более широком классе кусочно-дифференцируемых функций в π_n, s-е производные которых удовлетворяют условию Липшица в своих областях дифференцируемости. В важных подклассах дифференцируемых функций E_s^n и H_s^n ([2], [3]) могут быть получены значительно лучшие оценки для $v(F, N)$.

В частном случае $s = 0$ нетрудно построить оптимальные и асимптотически оптимальные алгоритмы. Возьмем целое положительное m и построим кубическую сетку с m^n узлами так, что ближайшие к ребрам π_n ряды сетки будут отстоять от них на расстоянии (в кубической метрике) $1/(2m)$. Такую сеть естественно назвать $1/(2m)$-минимальной кубической сетью в π_n.

Теорема 3. *Пусть $N = m^n$, где m — целое число, и X_ν, $\nu = 0, 1, \ldots, N-1$,— минимальная кубическая сетка в π_n.*

Тогда

$$\delta(C^n_{1,L}, N) = \sup_{f \in C^n_{1,L}} [\min_{0 \leqslant \nu \leqslant N-1} f(X_\nu) - \inf_{X \in \pi_n} f(X)] = L/(2N^{1/n}). \qquad (6)$$

Доказательство этой теоремы аналогично доказательству теоремы 1. Надо только вместо функции (4) ввести функцию $f^*(X) = 0$ для $\|X - X^*\|_1 \geqslant 1/(2m)$ и $= L\|X - X^*\|_1 - L/(2m)$ для $\|X - X^*\|_1 \leqslant 1/(2m)$. При помощи $f^*(X)$ легко показываем, что $\delta(C^n_{1,L}, N) \geqslant L/(2m)$. С другой стороны, очевидно, что $\min\limits_{0 \leqslant \nu \leqslant N-1} f(X_\nu) - \inf\limits_{X \in \pi_n} f(X) \leqslant L/(2m)$ для любой функции $f \in C^n_{1,L}$. Отсюда непосредственно выводим требуемое.

Теорема 4. *В случае $N = m^n$ с целым m алгоритм простого перебора значений $f(X) \in C^n_{1,L}$ в N узлах X_ν минимальной кубической сетки в π_n и значение $r_N = \min\limits_{0 \leqslant \nu \leqslant N-1} f(X_\nu) - L/(4m)$ дают оптимальный алгоритм A^* в классе $C^n_{1,L}$ среди всех алгоритмов. При этом*

$$v(C^n_{1,L}, N) = v(C^n_{1,L}, N, A^*) = \sup_{f \in C^n_{1,L}} \Big| \min_{0 \leqslant \nu \leqslant N-1} f(X_\nu) - L/(4m) - \inf_{X \in \pi_n} f(X) \Big| = L/(4m). \qquad (7)$$

Доказательство. При помощи двух функций $f_1(X) \equiv 0$ и $f_2(X) = f^*(X)$ по аналогии с тем, как это делается при доказательстве теоремы 1, получаем $v(C^n_{1,L}, N) \geqslant L/(4m)$. С другой стороны, нетрудно убедиться, что $|\min\limits_{0 \leqslant \nu \leqslant N-1} f(X_\nu) - L/(4m) + \inf\limits_{X \in \pi_n} f(X)| \leqslant L/(4m)$. Из указанных двух неравенств легко вытекает требуемое.

Из изложенного нетрудно заключить, что простой перебор на равномерных сетках в π_n и уменьшение наименьших на сетках значений f на $L/(4N^{1/n})$ будут давать асимптотически оптимальные алгоритмы в классе $C^n_{1.L}$.

В случае произвольного s важное практическое значение имеет построение наиболее точной кусочно-многочленной аппроксимации $f(X) \in C^n_{s+1,L}$ использующей информацию о значениях f не более чем в N точках π_n. Значительно более точную аппроксимацию по сравнению с той, которая предложена при доказательстве теоремы 1, можно получить за счет кусочно-многочленной аппроксимации с применением узлов Чебышева. Еще более точные способы аппроксимации могут быть основаны на фейеровских или валле-пуссеновских процессах интерполяции, а также на известных методах суммирования интерполяционных многочленов ([4]).

С другой стороны, можно существенно повысить оценку снизу для $v(F, N)$. В частности, в классе кусочно-дифференцируемых функций $\widetilde{C}^n_{s+1,L}$, вводя вместо (4) функцию

$$\widetilde{f}_2(X) = \begin{cases} 0, & \|X - X^*\|_1 \geqslant 1/(2N^{1/n}), \\ -\dfrac{L}{(s+1)!}(1/(2N^{1/n}) - \|X - X^*\|_1)^{s+1}, & \|X - X^*\| \leqslant 1/(2N^{1/n}), \end{cases} \quad (8)$$

получаем $v(\widetilde{C}^n_{s+1,L}, N) \geqslant L/((s+1)!\cdot 2^{s+2}\cdot N^{(s+1)/n})$. Тем не менее последняя оценка вряд ли может быть достигнута сверху при использовании лишь значений функций из $\widetilde{C}^n_{s+1,L}$ не более чем в N точках π_n.

Допустим в отличие от первоначальной постановки, что дан класс $W^n_{s+1,L}$ дифференцируемых в π_n функций $f(X)$, для которых $\varphi_{X_0}(t) = f(X_0 + t\nabla) = f(X)$, $\|\nabla\|_3 = 1$, обладает свойством $|\varphi^{(s)}_{X_0}(t_1) - \varphi^{(s)}_{X_0}(t_2)| \leqslant L|t_1 - t_2|$, где L не зависит от X_0, ∇ и f, и что алгоритмы (обозначаемые ниже через $\widetilde{A}$) могут использовать значения функций из $W^n_{s+1,L}$ и ее производных до s-го порядка включительно не более чем в N точках π_n. Обозначим результат применения алгоритма $\widetilde{A}$ к функции $f \in W^n_{s+1,L}$ через $r_N(\widetilde{A}, f)$.

Теорема 5. *При указанных предположениях и обозначениях и при $N = m^n$ с целым m*

$$L/((s+1)!\,2^{2s+3}N^{(s+1)/n}) \leqslant \inf_{\widetilde{A}} \sup_{f \in W^n_{s+1,L}} \left| r_N(\widetilde{A}, f) - \inf_{X \in \pi_n} f(X) \right| \leqslant$$

$$\leqslant \sup_{f \in W^n_{s+1,L}} \left| \inf_{X \in \pi_n} \left[\varphi_{X_{\nu(X)}}(0) + \varphi'_{X_{\nu(X)}}(0)\cdot t + \ldots + \frac{\varphi^{(s)}_{X_{\nu(X)}}(0)}{s!} t^s \right] - \right.$$

$$\left. - \inf_{X \in \pi_n} f(X) \right| \leqslant L/((s+1)!\,2^{s+1}\cdot N^{(s+1)/n}), \quad (9)$$

где $X_{\nu(X)}$ — ближайший к X узел минимальной кубической сетки в π_n, $X = X_{\nu(X)} + t\nabla$.

Доказательство аналогично вышеприведенным, если воспользоваться двумя функциями $f_1(X) \equiv 0$ и $f_2(X)$, где $f_2(X)$ дано формулой (4), и применить известную оценку для остаточного члена формулы Тейлора.

Институт кибернетики
Академии наук УССР
Киев

Поступило
29 IV 1971

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

[1] Н. С. Бахвалов, Журн. вычислит. матем. и матем. физ., 10, № 3, 555 (1970). [2] Н. М. Коробов, Теоретикочисловые методы в приближенном анализе, 1963. [3] С. А. Смоляк, ДАН, 148, № 5, 1042 (1963). [4] Н. И. Ахиезер, Лекции по теории аппроксимации, «Наука», 1965.