Доклады Академии Наук СССР
1951. Том LXXIX, № 4

МАТЕМАТИКА

С. И. ЗУХОВИЦКИЙ

АЛГОРИФМ ДЛЯ РЕШЕНИЯ ЧЕБЫШЕВСКОЙ ЗАДАЧИ ПРИБЛИЖЕНИЯ В СЛУЧАЕ КОНЕЧНОЙ СИСТЕМЫ НЕСОВМЕСТНЫХ ЛИНЕЙНЫХ УРАВНЕНИЙ

(Представлено академиком С. Н. Бернштейном 7 VI 1951)

Пусть дана система несовместных линейных уравнений

$$a_{i1}\xi_1 + a_{i2}\xi_2 + \ldots + a_{in}\xi_n = b_i \quad (i = 1, 2, \ldots, m > n) \tag{1}$$

и пусть все определители n-го порядка матрицы коэффициентов $\| a_{ik} \|$ отличны от нуля (условие Хаара).

Обозначим

$$x \equiv (\xi_1, \xi_2, \ldots, \xi_n), \quad \Delta_i(x) = \sum_{k=1}^{n} a_{ik}\xi_k - b_i \quad (i = 1, 2, \ldots, m). \tag{2}$$

Величину $|\Delta_i(x)|$ назовем уклонением i-го уравнения в точке x, а величину $\max\limits_i |\Delta_i(x)|$ — уклонением системы (1) в точке x.

Задача приближенного решения системы (1) в смысле П. Л. Чебышева („наилучшего приближения", „минимального приближения" ([1])) заключается в отыскании точки $x \equiv (\xi_1, \xi_2, \ldots, \xi_n)$, для которой достигается величина

$$\min_{\xi} \max_{i} \left| \sum_{k=1}^{n} a_{ik}\xi_k - b_i \right| = \min_{x} \max_{i} |\Delta_i(x)|, \tag{3}$$

называемая наименьшим уклонением системы (1).

Как известно ([2]), выполнение условия Хаара гарантирует для любой системы правых частей b_i существование и единственность точки x, для которой достигается (3).

Первые шаги предлагаемого алгорифма представляют собой некоторое приспособление к нашей задаче метода наискорейшего спуска ([3]).

Начинаем с произвольной точки $x_0 \equiv (\xi_1^{(0)}, \xi_2^{(0)}, \ldots, \xi_n^{(0)})$ и отыскиваем то из уравнений, пусть i_0-е, для которого уклонение в точке x_0 наибольшее:

$$|\Delta_{i_0}(x_0)| > |\Delta_i(x_0)| \quad (i \neq i_0).$$

Затем строим для функционала $|\Delta_{i_0}(x_0 + \varepsilon z)|$ направление градиента $z_1 \equiv (\zeta_1^{(1)}, \zeta_2^{(1)}, \ldots, \zeta_n^{(1)})$:

$$\zeta_k^{(1)} = \operatorname{sign} \Delta_{i_0}(x_0) \frac{a_{i_0, k}}{\lambda} \quad (k = 1, 2, \ldots, n); \quad \lambda = \pm \sqrt{\sum_{k=1}^{n} a_{i_0, k}}, \tag{4}$$

причем знак у корня берем такой, чтобы

$$\frac{d}{d\varepsilon} | \Delta_{i_0}(x_0 + \varepsilon z_1) |_{\varepsilon=0} = \operatorname{sign} \Delta_{i_0}(x_0) \sum_{k=1}^{n} a_{i_0 k} \zeta_k^{(1)} < 0.$$

Рассматриваем уравнения $\Delta_{i_0}(x_0 + \varepsilon z_1) = \delta_i \Delta_i (x_0 + \varepsilon z_1)$, где $i \neq i_0$ и $\delta_i = \pm 1$, из которых получаем

$$\varepsilon = \frac{\delta_i \Delta_i (x_0) - \Delta_{i_0}(x_0)}{\sum\limits_{k=1}^{n} (a_{i_0 k} - \delta_i a_{ik}) \zeta_k^{(1)}} \tag{5}$$

и пусть ε_1 — наименьшее положительное значение ε, достигаемое, например, при $i = i_1$.

За первое приближение принимаем точку $x_1 = x_0 + \varepsilon_1 z_1$. Ясно, что

$$| \Delta_{i_0}(x_1) | = | \Delta_{i_1}(x_1) | > | \Delta_i (x_1) | \quad (i \neq i_0, i_1), \quad \Delta_{i_0}(x_1) = \delta_{i_1} \Delta_{i_1}(x_1).$$

Если уже построено x_p $(p < n)$ такое, что

$$| \Delta_{i_0}(x_p) | = | \Delta_{i_1}(x_p) | = \cdots = | \Delta_{i_p}(x_p) | > | \Delta_i (x_p) | \quad (i \neq i_0, i_1, \ldots, i_p)$$

и

$$\Delta_{i_0}(x_p) = \delta_{i_1} \Delta_{i_1}(x_p) = \cdots = \delta_{i_p} \Delta_{i_p}(x_p),$$

то относительный градиент $z_{p+1} = (\zeta_1^{(p+1)}, \zeta_2^{(p+1)}, \ldots, \zeta_n^{(p+1)})$ для функционала $| \Delta_{i_0}(x_p + \varepsilon z) |$, двигаясь по которому уменьшается $| \Delta_{i_0}(x_p + \varepsilon z_{p+1}) |$ при сохранении равенств

$$\Delta_{i_0}(x_p + \varepsilon z_{p+1}) = \delta_{i_1} \Delta_{i_1}(x_p + \varepsilon z_{p+1}) = \cdots = \delta_{i_p} \Delta_{i_p}(x_p + \varepsilon z_{p+1}),$$

определяется формулами:

$$\zeta_k^{(p+1)} = \frac{\operatorname{sign} \Delta_{i_0}(x_p)\, a_{i_0 k} + \sum\limits_{\nu=1}^{p} \mu_\nu (a_{i_0 k} - \delta_{i_\nu} a_{i_\nu k})}{\lambda} \quad (k = 1, 2, \ldots, n), \tag{6}$$

где μ_ν определяются из уравнений

$$\sum_{k=1}^{n} \left\{ (a_{i_0 k} - \delta_{i_\tau} a_{i_\tau k}) \left[\operatorname{sign} \Delta_{i_0}(x_p)\, a_{i_0 k} + \sum_{\nu=1}^{p} \mu_\nu (a_{i_0 k} - \delta_{i_\nu} a_{i_\nu k}) \right] \right\} = 0$$

$$(\tau = 1, 2, \ldots, p) \tag{7}$$

и

$$\lambda = \pm \sqrt{\sum_{k=1}^{n} \left[\operatorname{sign} \Delta_{i_0}(x_p)\, a_{i_0 k} + \sum_{\nu=1}^{p} \mu_\nu (a_{i_0 k} - \delta_{i_\nu} a_{i_\nu k}) \right]^2}.$$

Знак у корня берем такой, чтобы $\frac{d}{d\varepsilon} | \Delta_{i_0}(x_p + \varepsilon z_{p+1}) |_{\varepsilon=0} < 0$.

Рассмотрим уравнения

$$\Delta_{i_0}(x_p + \varepsilon z_{p+1}) = \delta_i \Delta_i (x_p + \varepsilon z_{p+1}) \quad (i \neq i_0, i_1, \ldots, i_p;\ \delta_i = \pm 1)$$

и пусть ε_{p+1} — наименьшее положительное значение ε, достигаемое, например, при $i = i_{p+1}$.

За $p+1$-е приближение принимаем точку $x_{p+1} = x_p + \varepsilon_{p+1} z_{p+1}$. Ясно, что

$$| \Delta_{i_0}(x_{p+1}) | = | \Delta_{i_1}(x_{p+1}) | = \cdots = | \Delta_{i_{p+1}}(x_{p+1}) | > | \Delta_i (x_{p+1}) | \quad (i \neq i_0, i_1, \ldots, i_{p+1})$$

и

$$\Delta_{i_0}(x_{p+1}) = \delta_{i_1} \Delta_{i_1}(x_{p+1}) = \cdots = \delta_{i_{p+1}} \Delta_{i_{p+1}}(x_{p+1}).$$

Продолжаем процесс до построения так называемой стационарной точки x_n, для которой

$$|\Delta_{i_0}(x_n)| = |\Delta_{i_1}(x_n)| = \cdots = |\Delta_{i_n}(x_n)| > |\Delta_i(x_n)| \quad (i \neq i_0, i_1, \ldots, i_n). \qquad (8)$$

Если стационарная точка x_n лежит внутри n-мерного симплекса, ограниченного плоскостями $i_0, i_1, \ldots, i_n$ системы (1), то она искомая, и процесс закончен.

Если же в процессе применения алгорифма мы натолкнемся на стационарную точку x', для которой

$$\Delta_{i_0}(x')| = |\Delta_{i_1}(x')| = \cdots = |\Delta_{i_p}(x')| > |\Delta_i(x')| \quad (p \geqslant n;\ i \neq i_0, i_1, \ldots, i_p),$$

и x' лежит вне всех n-мерных симплексов, образованных любыми $n+1$ плоскостями из плоскостей $i_0, i_1, \ldots, i_p$, то всегда существует, по крайней мере, одна вершина x'' у этих симплексов такая, что при движении к ней по отрезку, соединяющему x' с x'', будут уменьшаться все $p+1$ наибольших уклонений, причем не меньше чем n из них уменьшаясь остаются равными между собой, а остальные уменьшаются быстрее.

Вершина x'' характеризуется тем, что для нее имеет максимальное значение сумма числа плоскостей, проходящих через нее, и числа плоскостей, отделяющих ее от стационарной точки x'.

Двигаемся теперь из x' к этой вершине x'' по отрезку, соединяющему эти две точки, пока, подобно предыдущим шагам, не натолкнемся на новую стационарную точку с уменьшенными наибольшими уклонениями.

Так, монотонно уменьшая наибольшие уклонения, мы после конечного числа шагов придем к искомой точке (являющейся, очевидно, также стационарной).

Замечание 1. Нетрудно убедиться, что нашим алгорифмом можно пользоваться и в том случае, когда система (1) не удовлетворяет условию Хаара, но тогда решение может оказаться не единственным.

Замечание 2. Внутренняя точка n-мерного симплекса, ограниченного плоскостями

$$a_{i_\nu 1}\xi_1 + a_{i_\nu 2}\xi_2 + \ldots + a_{i_\nu n}\xi_n = b_{i_\nu} \quad (\nu = 0, 1, 2, \ldots, n) \qquad (9)$$

характеризуется тем, что при подстановке ее координат в левую часть ν-го уравнения получим значение отрицательное, если C и C_ν имеют разные знаки, и положительное, если C и C_ν имеют одинаковые знаки, где

$$C = \begin{vmatrix} a_{i_0 1} & a_{i_0 2}, & \ldots, & a_{i_0 n} & -b_{i_0} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{i_n 1} & a_{i_n 2}, & \ldots, & a_{i_n n} & -b_{i_n} \end{vmatrix},$$

а C_ν — алгебраическое дополнение элемента $-b_{i_\nu}$.

Замечание 3. Предыдущий алгорифм может быть использован ([4]) для приближенного построения многочлена $P(t) = \xi_0 t^n + \xi_1 t^{n-1} + \ldots + \xi_n$, наименее уклоняющегося от данной непрерывной на сегменте $[a, b]$ функции $f(t)$.

Действительно, пусть $\max\limits_{a \leqslant t \leqslant b} |f(t)| = M$ и ρ_0 — нижняя грань для наименьшего уклонения $\rho = \min\limits_{x} \max\limits_{t} |f(t) - P(t)|$. Пусть $Q(t)$ — многочлен

n-й степени и $|Q(t)| \leqslant 2M$. Тогда, по известной теореме А. А. Маркова ([5]), в интервале (a, b) будет $|Q'(t)| \leqslant \frac{4n^2 M}{b-a}$. Каждому $\varepsilon > 0$ соответствует такое $\delta > 0$, что из $|t' - t''| < \delta$ следует $|f(t') - f(t'')| < \frac{\varepsilon \rho_0}{2}$. Потребуем, кроме того, чтобы $\delta < \frac{\varepsilon \rho_0 (b-a)}{8n^2 M}$. Тогда

$$|f(t') - Q(t') - f(t'') + Q(t'')| < \varepsilon \rho_0, \text{ если } |t' - t''| < \delta.$$

Определив таким образом $\delta > 0$ по данному $\varepsilon > 0$, разобьем $[a, b]$ на равные сегменты $[t_i, t_{i+1}]$ системой точек $F = \{a = t_1 < t_2 < \cdots < t_m = b\}$ так, чтобы $t_{i+1} - t_i < \delta$.

Разыщем по нашему алгорифму многочлен $Q_1(t)$ степени $\leqslant n$ ($|Q_1(t)| \leqslant 2M$), дающий наименьшее уклонение от $f(t)$ на F. Тогда

$$\max_{t \in F} |f(t) - Q_1(t)| = \rho' \leqslant \rho.$$

Далее, в некоторой точке t' ($t_i \leqslant t' \leqslant t_{i+1}$) достигается $\max\limits_{a \leqslant t \leqslant b} |f(t) - Q_1(t)|$ и мы получим

$$\max_{a \leqslant t \leqslant b} |f(t) - Q_1(t)| = |f(t') - Q_1(t')| \leqslant$$

$$\leqslant |f(t') - Q_1(t') - f(t_i) + Q_1(t_i)| + |f(t_i) + Q_1(t_i)| < \varepsilon \rho_0 + \rho' < \rho(1 + \varepsilon).$$

Киевский государственный педагогический институт им. А. М. Горького

Поступило 10 V 1951

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

[1] Е. Я. Ремез, Про методи найкращого, в розумінні Чебишова, наближенного представлення функцій, Київ, 1935. [2] Н. И. Ахиезер, Лекция по теории апроксимации, М.— Л., 1947. [3] Л. В. Канторович, Усп. матем. наук, 3, в. 6 (28) (1948). [4] C. J. de la Vallée-Poussin, Leçons sur l'approximation des fonctions d'une variable réelle, Paris, 1919. [5] А. Марков, Зап. Российск. Акад. наук, 62 (1889).