EESTI NSV TEADUSTE AKADEEMIA TOIMETISED. 25. KÖIDE
FÜÜSIKA * MATEMAATIKA. 1976, NR. 2

ИЗВЕСТИЯ АКАДЕМИИ НАУК ЭСТОНСКОЙ ССР. ТОМ 25
ФИЗИКА * МАТЕМАТИКА. 1976, № 2

https://doi.org/10.3176/phys.math.1976.2.03

УДК 518 : 512.39

В. ПОЛЛЬ

МОДИФИКАЦИИ ИНТЕРПОЛЯЦИОННОГО МЕТОДА ДЛЯ НАХОЖДЕНИЯ СТАЦИОНАРНЫХ ТОЧЕК ФУНКЦИИ МНОГИХ ПЕРЕМЕННЫХ

В последнее время было предложено несколько различных беспроизводных аналогов метода Ньютона [1–7] со сверхлинейной или квадратичной скоростью сходимости. В настоящей работе для нахождения стационарных точек функции многих переменных исследуются некоторые методы такого же типа, представляющие собой новые модификации интерполяционного аналога метода Ньютона, но дающие более высокую скорость сходимости, чем ранее предложенные α- и ε-методы [8–10]. При этом количество значений функции, подлежащих вычислению на каждом итерационном шаге, не увеличивается по сравнению с последними.

Итак, для нахождения стационарных точек функции многих переменных $f(x)=f(x_1, \ldots, x_d)$, где y^n, z^n являются величинами, зависящими от x^n и x^{n-1}, рассмотрим метод

$$x^{n+1}=x^n-[f_{11}(x^n;y^n;z^n)+f_{12}(x^n;y^n;z^n)]^{-1}\times$$
$$\times[f_1(x^n;y^n)+f_{11}(x^n;y^n;z^n)(x^n-y^n)] \tag{1}$$

($n=1,2,\ldots$; x^0, x^1 — начальные приближения). В работах [8, 9] был исследован α-метод, где $y^n=\alpha x^n+(1-\alpha)x^{n-1}$, $\alpha\in(0,1)$, $z^n=x^{n-1}$, и в работе [10] — ε-метод, где $y^n=x^{n-1}$, $z^n=x^n-\varepsilon f_1(x^n;\ x^{n-1})$, ε — достаточно малое положительное число. Оба эти метода дают порядок сходимости $k=\sqrt{2}$.

Предложим три модификации метода (1) с порядками сходимости больше, чем $\sqrt{2}$:

а) $$y^n=\alpha_n x^n+(1-\alpha_n)x^{n-1},\ \alpha_n\in(0,1);\quad \alpha_n\to 1\quad (n\to\infty);$$
$$z^n=x^{n-1}; \tag{2}$$

б) $$y^n=\alpha_n x^n+(1-\alpha_n)x^{n-1};\quad z^n=\beta_n x^n+(1-\beta_n)x^{n-1};$$
$$\alpha_n,\beta_n\in(0,1),\quad \alpha_n\neq\beta_n;\quad \alpha_n\to 1,\quad \beta_n\to 1\quad (n\to\infty); \tag{3}$$

в) $$y^n=\alpha x^n+(1-\alpha)\bar{x}^n,\quad \alpha\in(0,1);\quad z^n=\bar{x}^n,$$
$$\|x^n-\bar{x}^n\|\leqslant c_n\|x^n-x^{n-1}\|^2,\quad \inf_n c_n>0,\quad \sup_n c_n<\infty. \tag{4}$$

Для дальнейшего изложения предположим, что для любых x', x'', x''', $x^{\mathrm{IV}}\in S\subset m_d$ (S — некоторая замкнутая сфера) определены непрерывные обобщенные разделенные разности $f_1(x'; x'')$; $f_{11}(x'; x''; x''')$; $f_{12}(x'; x''; x''')$ и выполнены следующие неравенства:

$$\|f_{11}(x'; x''; x''') - f_{11}(x''; x'''; x^{IV})\| \leqslant$$
$$\leqslant a\|x' - x''\| + b\|x'' - x'''\| + c\|x''' - x^{IV}\|;$$
$$\|f_{12}(x'; x''; x''') - f_{12}(x''; x'''; x^{IV})\| \leqslant$$
$$\leqslant d\|x' - x''\| + e\|x'' - x'''\| + f\|x''' - x^{IV}\|;$$
$$\|[f_{11}(x'; x''; x''') + f_{12}(x'; x''; x''')]^{-1}\| \leqslant p, \quad \text{(см. [9])}. \tag{5}$$

Из (1) и (5) получаем, что

$$\|x^* - x^{n+1}\| \leqslant A\|x^* - x^n\|\,\|x^* - y^n\| + B\|x^* - x^n\|\,\|x^* - z^n\| + C\|x^* - y^n\|\,\|x^* - z^n\|, \tag{6}$$

где $\qquad A = pe; \quad B = pf; \quad C = pc.$

Теорема 1. *Пусть*

1° *функция* $f(x)$ *имеет стационарную точку* $x^* \in S(x^0; \varrho) \subset m_d$;

2° $\|x^* - x^1\| \leqslant \varrho$;

3° *в сфере* $S(x^*; \varrho) \subset m_d$ *имеют место оценки* (5);

4° $(A + B + 2C)\varrho = F\varrho < 1$;

5° $\alpha_n = 1 - h^{\mu_{n-2}}, \quad \mu_n = \mu_{n-1} + \mu_{n-2} \quad (n \geqslant 2), \quad \mu_0 = \mu_1 = 1, \quad 0 < h < 1$;

α_1 — *произвольное число, меньшее единицы.*

Тогда последовательность (1)—(2) *сходится к* x^*, *причем в зависимости от выполнения* $F\varrho < h < 1$ *или* $h \leqslant F\varrho < 1$ *имеет место*

$$\|x^* - x^{n+1}\| \leqslant \varrho h^{\mu_{n+1} - 1}$$

или

$$\|x^* - x^{n+1}\| \leqslant F^{-1}(F\varrho)^{\mu_{n+1}} \quad (n = -1, 0, 1, \ldots).$$

Доказательство нетрудно провести по индукции (см. [8]), учитывая выводимое из (6) и (2) неравенство

$$d_{n+1} = \|x^* - x^{n+1}\| \leqslant A\alpha_n d_n^2 + [A(1 - \alpha_n) + B + C\alpha_n] d_n d_{n-1} + C(1 - \alpha_n) d_{n-1}^2.$$

Очевидно, что последовательность чисел μ_n ($n = 0, 1, \ldots$) обеспечивает методу (1)—(2) порядок сходимости $k = \dfrac{1 + \sqrt{5}}{2} \approx 1{,}618$.

Вариант (1)—(3) метода (1) будет сходиться к x^* со скоростью $k = 2$ (при тех же предположениях, что и в теореме 1, причем $F = 2(A + B + 2C)$), если принять

$$\alpha_n = 1 - h^{\mu_{n-1}}; \quad \mu_n = 2\mu_{n-1} \quad (n \geqslant 1), \quad \mu_0 = 1;$$
$$\beta_n = 1 - h^{\mu_{n-1} + \delta}; \quad \delta > 0, \quad 0 < h < 1.$$

Замечание 1. Если в методе (1) взять

$$y_n = x^{n-1}, \quad z^n = x^n - \varepsilon_n f_1(x^n; x^{n-1}), \quad \varepsilon_n = h^{\mu_{n-2}},$$
$$\mu_n = \mu_{n-1} + \mu_{n-2} \quad (n \geqslant 2), \quad \mu_0 = \mu_1 = 1, \tag{7}$$

ε_1 — произвольное число, меньшее единицы, $0 < h < 1$, то при определенных условиях, аналогичных условиям теоремы 1, метод (1)—(7) обеспечит сходимость к x^* с той же скоростью, что и метод (1)—(2). Если же задать

$$y^n = x^n - \varepsilon^1_n f_1(x^n; x^{n-1}) \quad \text{и} \quad z^n = x^n - \varepsilon^2_n f_1(x^n; x^{n-1}), \quad \varepsilon^1_n = h^{\mu_{n-1}},$$

$$\mu_n = 2\mu_{n-1} \quad (n \geqslant 1), \quad \mu_0 = 1, \quad \varepsilon^2_n = h^{\mu_{n-1}+\delta}, \quad \delta > 0, \quad 0 < h < 1,$$

то его скорость сходимости будет квадратичной.

Теорема 2. *Пусть*

1° *функция* $f(x)$ *имеет стационарную точку* $x^* \in S(x^0; \varrho) \subset m_d$;

2° $\|x^* - x^1\| \leqslant \varrho$;

3° *в сфере* $S(x^*; \varrho(1 + 4\bar{c}\varrho))$ $(\bar{c} = \sup\limits_n c_n)$ *имеют место оценки* (5);

4° $F\varrho < 1$; F *является положительным корнем уравнения*

$$F^3 - (a_0 + 3b_0\varrho + 5c_0\varrho^2) F^2 - (b_0 + 10c_0\varrho) F - c_0 = 0,$$

где

$$a_0 = A + B + C;$$
$$b_0 = \bar{c}[(1-\alpha)A + B + (2-\alpha)C];$$
$$c_0 = \bar{c}^2(1-\alpha)C.$$

Тогда последовательность (1)—(4) *сходится к* x^*, *причем*

$$\|x^* - x^{n+1}\| \leqslant F^{-1}(F\varrho)^{\mu_{n+1}}, \quad n = -1, 0, 1, \ldots,$$

где $\mu_{n+1} = 2\mu_n$ $(n \geqslant 1)$, $\mu_0 = \mu_1 = 1$.

Доказательство. Учитывая (4) и (6), получим

$$d_{n+1} \leqslant \alpha A d_n^2 + [(1-\alpha)A + B + \alpha C] d_n \|x^* - \bar{x}^n\| + (1-\alpha) C \|x^* - \bar{x}^n\|^2. \tag{8}$$

Поскольку

$$\|x^* - \bar{x}^n\| \leqslant d_n + \bar{c}(d_n^2 + 2d_n d_{n-1} + d_{n-1}^2),$$

то (8) преобразуется к виду

$$d_{n+1} \leqslant a_0 d_n^2 + b_0 d_n^3 + 2b_0 d_n^2 d_{n-1} + b_0 d_n d_{n-1}^2 + c_0 d_n^4 + 4c_0 d_n d_{n-1}^3 +$$
$$+ 6c_0 d_n^2 d_{n-1}^2 + 4c_0 d_n^3 d_{n-1} + c_0 d_{n-1}^4. \tag{9}$$

Исходя из (9), легко проверить по индукции справедливость утверждения теоремы, а вместе с тем и принадлежность используемых точек к указанной в условиях теоремы сфере. Теорема доказана.

Замечание 2. Если в методе (1) принять

$$y^n = \bar{x}^n, \quad z^n = x^n - \varepsilon f_1(x^n; \bar{x}^n), \tag{10}$$

ε — достаточно малое положительное число, то при определенных условиях метод (1)—(10) также даст квадратичную скорость сходимости к x^*. При выборе $y^n = \bar{x}^n$ и $z^n = x^{n-1}$ получим метод с $k \cong 1{,}618$. Точку $\bar{x}^n$ можно определить также из следующих неравенств (при этом сохраняется квадратичная сходимость методов):

$$\|x^n - \bar{x}^n\| \leqslant c_n \|f_1(x^n; x^{n-1})\|^2 \tag{11}$$

или

$$\|x^n - \bar{x}^n\| \leqslant c_n |f(x^n) - f(x^{n-1})|. \tag{12}$$

Замечание 3. Из неравенства

$$\|x^* - x^{n+1}\| \leqslant p\{\|[f_{11}(x^n; y^n; z^n) + f_{12}(x^n; y^n; z^n) - f''(x^n)](x^* - x^n) +$$
$$+ [f_1(x^n; y^n) + f_{11}(x^n; y^n; z^n)(x^n - y^n) - f'(x^n)] +$$
$$+ f'(x^n) + f''(x^n)(x^* - x^n)\|\} \leqslant$$
$$\leqslant p\{\|x^* - x^n\| [(b+e)\|x^n - y^n\| + (c+f)\|x^n - z^n\|] +$$
$$+ c\|x^n - y^n\| \|x^n - z^n\| + \frac{1}{2} N \|x^* - x^n\|^2\}.$$

где $\|f'''(x)\| \leqslant N$, $x \in [x^*, x^n]$, следует, что для получения из (1) метода второго порядка сходимости y^n и z^n следует выбирать из условия

$$\max\{\|x^n - y^n\|, \|x^n - z^n\|\} \leqslant k\|x^n - x^{n-1}\|^2, \tag{13}$$

где k — некоторая произвольная положительная константа. Иными словами, y^n и z^n следует выбирать так, чтобы сумма вторых обобщенных разделенных разностей $f_{11}(x^n; y^n; z^n) + f_{12}(x^n; y^n; z^n)$ аппроксимировала $f''(x^n)$ с точностью $O(\|x^n - x^{n-1}\|^2)$, а выражение $f_1(x^n; y^n) + f_{11}(x^n; y^n; z^n)(x^n - y^n)$ аппроксимировало $f'(x^n)$ с точностью $O(\|x^n - x^{n-1}\|^4)$ (ср. с [4]). Этим требованиям удовлетворяет метод (1)—(4) и его варианты (1)—(11) и (1)—(12).

Замечание 4. Предложенные модификации метода (1) не увеличивают объема вычислительной работы, необходимой на каждом итерационном шаге, по сравнению с α- и ε-методами. Дополнительно они требуют лишь определения значений параметров α_n, β_n или нахождения $\bar{x}^n$. В случае же применения варианта (1)—(11) метода (1)—(4) нужно вычислить еще $d+1$ значений функции $f(x)$.

Пример. Найти

$$\min f(x) = \min \sum_{i=1}^{3} \{ \sum_{j=1}^{3} (A_{ij} \sin x_j + B_{ij} \cos x_j) - E_i \}^2, \tag{14}$$

где

$$A = \begin{pmatrix} -13 & -30 & 8 \\ 44 & -29 & -82 \\ -76 & 72 & -17 \end{pmatrix}; \quad B = \begin{pmatrix} -40 & -39 & 63 \\ 6 & 29 & 67 \\ 14 & -56 & -51 \end{pmatrix}; \quad E = \begin{pmatrix} -117{,}224 \\ -72{,}82396 \\ 81{,}71924 \end{pmatrix};$$

$x^* = (-1{,}014147;\ 0{,}1808786;\ -3{,}081409)$, $f(x^*) = 0$.

Начальные приближения:

$x^1 = (-1{,}131226;\ 0{,}0260196;\ -2{,}944214)$,

$x^0 = (-0{,}7465715;\ 0{,}4833373;\ -3{,}242432)$,

вычисления продолжались до $\|x^n - x^{n-1}\| \leqslant 10^{-5}$.

α-Методом с $\alpha = 0{,}5$ минимум функции (14) был найден с такой точностью через семь итераций, ε-методом с $\varepsilon = 0{,}0001$ — через восемь, методом Ньютона (с начальным приближением x^1) — через пять. Предложенные здесь методы дали нужную точность: метод (1)—(2) с $h = 0{,}5$ — через шесть итераций, метод (1)—(3) с $h = 0{,}5$ и $\delta = 1$ — через четыре и метод (1)—(4) — через пять.

Вычисления были произведены на ЭВМ «Минск-22».

ЛИТЕРАТУРА

1. Шаманский В. Е., Методы численного решения краевых задач на ЭЦВМ, ч. II, Киев, 1966, с. 61.
2. Огнева В. А., Чернышенко В. М., Матем. записки, **8**, 487 (1970).
3. Бартиш М. Я., № 5653-73 Деп., Новосибирск, 1973.
4. Полль В., Изв. АН ЭССР, Физ. Матем., **22**, 409 (1973).
5. Brown K. M., Dennis J. E. jr., Numer. Math., **18**, 289 (1972).
6. Schwetlick H., ЖВМ и МФ, **14**, 278 (1974).
7. Schmidt J. W., Math. Nachr., **B. 59**, H. 1—6, 95 (1974).
8. Полль В., Изв. АН ЭССР, Физ. Матем., **16**, 35 (1967).
9. Полль В., Изв. АН ЭССР, Физ. Матем., **16**, 157 (1967).
10. Полль В., Изв. АН ЭССР, Физ. Матем., **16**, 382 (1967).

Институт кибернетики
Академии наук Эстонской ССР

Поступила в редакцию
3/VII 1975

V. POLL

MITME MUUTUJA FUNKTSIOONI STATSIONAARSETE PUNKTIDE LEIDMISE INTERPOLATSIOONIMEETODI MODIFIKATSIOONE

Konstrueeritakse Newtoni meetodi interpolatsioonianaloogi kaht tüüpi modifikatsioonid, mis tagavad kiirema koondumise (koondumiskiirus ulatub ruutkiiruseni) kui teadaolevad analoogilised varasemad meetodid [8–10], kuid samal ajal ei suurene ühel iteratsioonil arvutatavate funktsiooni väärtuste arv. Tuuakse numbriline näide.

V. POLL

SOME MODIFICATIONS OF THE INTERPOLATION METHOD FOR FINDING STATIONARY POINTS OF A FUNCTION OF SEVERAL VARIABLES

The modifications of two types of interpolation analogue of Newton's method are proposed. These modifications converge faster than the modifications (α-method and ε-method [8–10]), known up to the present time, but at the same time they require as many function evaluation at each iteration as the latter modifications do. A numerical example is given.