Serdica

Bulgariacae mathematicae publicationes

Сердика

Българско математическо списание

О НЕКОТОРЫХ МОДИФИКАЦИЯХ МЕТОДА НЬЮТОНА ДЛЯ ПРИБЛИЖЕННОГО РЕШЕНИЯ АЛГЕБРАИЧЕСКИХ УРАВНЕНИЙ

СПАС П. ТАШЕВ, НИКОЛАЙ КЮРКЧИЕВ

В статье выводится итерационная формула, являющаяся модификацией метода Ньютона и производится оценка скорости сходимости итерационного процесса. Приведены некоторые результаты экспериментальных расчетов.

Известно, что метод Ньютона является одним из старейших вычислительных методов уточнения корней нелинейного уравнения $f(x)=0$.

Скорость сходимости процесса:

$$x_{k+1} = -x_k f(x_k)/f'(x_k); \quad k=0,\ 1,\ 2, \dots \tag{1}$$

квадратичная, когда величина $M/|f'(x_k)|$ ограничена, где M — наибольшее значение $|f''(x)|$ в рассматриваемой окрестности корня. Метод очень часто применяется, и поэтому естественно, что было сделано ряд попыток, преследующих цель изменить его либо в сторону увеличения скорости сходимости, либо — упрощения вычислений.

Пусть дано алгебраическое уравнение $f(x)=0$ степени n, не имеющее кратных корней:

$$f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0 = 0. \tag{2}$$

Оказывается, что итерационный процесс (1) можно модифицировать для одновременного приближенного вычисления всех корней $x_1, \dots, x_n$ уравнения (2), чтобы сохранить его сходимость.

К. Дочев [1; 2] предложил следующий метод:
Пусть для итерационного решения уравнения (2) задан процесс:

$$x_i^{(k+1)} = x_i^{(k)} - f(x_i^{(k)}) / \prod_{\substack{j=1\\ j\neq i}}^{n} (x_i^{(k)} - x_j^{(k)}; \ i=1,\ 2,\ \dots,\ n;\ k=0,1,\dots \tag{3}$$

Если начальные приближения $\{x_i^{(0)}\}_1^n$ удовлетворяют неравенствам $|x_i^{(0)} - x_i| \le \rho$, $i=1, \dots, n$ и $d=\min_{i\neq j}|x_i - x_j|$, где $\rho = d((1+q)^{1/(n-1)}-1)/(2(1+q)^{1/(n-1)}-1)$ и $0<q<1$, то процесс (3) обладает квадратической скоростью сходимости, то есть $|x_i^{(k)} - x_i| \le \rho q^{2^k-1}$.

Существуют другие модификации метода Ньютона. Как отметили К. Дочев и П. Бырнев [3], можно получить итерационные методы с кубической и с большей скоростью сходимости (напр., модифицированные схемы

СЕРДИКА Българско математическо списание. Том 9, 1983, с. 67—72.

Чебышева), использующие на каждом шаге информацию $f(x_i^{(k)})$, $f'(x_i^{(k)})$ или производные высших порядков. Так, например, процесс:

$$x_i^{(k+1)}=x_i^{(k)}-f(x_i^{(k)})\,(2\prod_{\substack{j=1\\ j\neq i}}^{n}(x_i^{(k)}-x_j^{(k)})-f'(x_i^{(k)})+f(x_i^{(k)})\sum_{\substack{j=1\\ j\neq i}}^{n}(x_i^{(k)}-x_j^{(k)})^{-1})/\prod_{\substack{j=1\\ j\neq i}}^{n}(x_i^{(k)}-x_j^{(k)})^2; \tag{4}$$

$i=1,\ldots,n$; $k=0,1,2,\ldots$, сходится с кубической скоростью [3].

Пусть для итерационного решения уравнения (2) задан процесс:

(5)

$$x_i^{(k+1)}=x_i^{(k)}-f(x_i^{(k)})/(f'(x_i^{(k)})-f(x_i^{(k)})\sum_{\substack{j=1\\ j\neq i}}^{n}(x_i^{(k)}-x_j^{(k)})^{-1});\quad i=1,\ldots,n;\quad k=1,2,\cdots$$

Итерацию (5) исследовал Л. Ерлих [4]. Он показал, что она обладает свойством кубической сходимости при некоторых ограничительных условиях. Работа [4] является дальнейшим развитием работ [1; 2; 3]. В методе Л. Ерлиха главная доля труда затрачивается на нахождение значений $f'(x_i^{(k)})$ на каждом шаге, и было бы желательно избежать вычисление f' вполне или отчасти за счет некоторой потери в скорости сходимости последовательности $\{x_i^{(k)}\}_1^n$. В связи с этим иногда может оказаться практически более целесообразным рассмотреть следующую процедуру:

$$x_i^{(k+1)}=x_i^{(k)}-f(x_i^{(k)})/(\prod_{\substack{j=1\\ j\neq i}}^{n}(x_i^{(k)}-x_j^{(k)})+\delta_i f(x_i^{(k)})\sum_{\substack{j=1\\ i\neq i}}^{n}(x_i^{(k)}-x_j^{(k)})^{-1}); \tag{6}$$

$i=1,\ldots,n$; $k=0,1,\ldots$, где

Таблица 1

k	i	$x_i^{(k)}$ по схеме (3)	$x_i^{(k)}$ по схеме (5)	$x_i^{(k)}$ по схеме (6)
0	1	1.01	1.01	1.01
1		1.000206879	0.999994111	1.000002915
0	2	2.01	2.01	2.01
1		2.000084158	1.999997636	2.000001539
0	3	3.01	3.01	3.01
1		3.00000125	2.999997499	3.00000125
0	4	4.01	4.01	4.01
1		3.999917508	3.999997674	4.000001516
0	5	5.01	5.01	5.01
1		4.999790204	4.999998586	5.000002848

$$\delta_i = \begin{cases} 1, & \text{если} \quad \operatorname{sgn} f(x_i^{(k)}) \sum\limits_{\substack{j=1\\ j\neq i}}^{n} (x_i^{(k)} - x_j^{(k)})^{-1} = (-1)^i, \\ -1, & \text{если} \quad \operatorname{sgn} f(x_i^{(k)}) \sum\limits_{\substack{j=1\\ j\neq i}}^{n} (x_i^{(k)} - x_j^{(k)})^{-1} \neq (-1)^i. \end{cases}$$

Приведем пример на применение этого метода. Пусть дано уравнение $f(x)=(x-1)(x-2)(x-3)(x-4)(x-5)$. В таблице 1 приведены вычислительные эксперименты, проведенные по схемам (3), (5) и (6).

Из таблицы 1 видно, что приближения, полученные по схемам (5) и (6), сравнимы.

Выведем итерационную формулу, являющуюся модификацией метода Ньютона. Возможно получить скорость сходимости выше квадратической, используя только информацию $f(x_i^{(k)})$. Общие итерационные схемы можно получить из следующих формул:

(7)
$$x_i^{(k+1)} = x_i^{(k)} - f(x_i^{(k)}) / (\alpha \prod_{\substack{j=1\\ j\neq i}}^{n} (x_i^{(k)} - x_j^{(k)}) + \beta \sum_{\substack{j=1\\ j\neq i}}^{n} f(x_j^{(k)}) (x_j^{(k)} - x_i^{(k)})^{-1} \prod_{\substack{s=1\\ s\neq j\\ s\neq i}}^{n} (x_i^{(k)} - x_s^{(k)}) (x_j^{(k)} - x_s^{(k)})^{-1});$$

$$i=1, \ldots, n; \quad k=0, 1, 2, \ldots$$

Эти вычислительные схемы зависят от некоторых параметров, которые могут выбираться в зависимости от конкретных свойств многочлена $f(x)$. Большее влияние на скорость сходимости оказывает изменения параметров. Естествено, что уточнение этих параметров целесообразно, когда необходимо многократно решать однотипные уравнения. Отметим, что при $\alpha=1$, $\beta=0$ из (7) получаем формулу Дочева. Пусть для итерационного решения уравнения (2) задан процесс:

(8)
$$x_i^{(k+1)} = x_i^{(k)} - f(x_i^{(k)}) / (\prod_{\substack{j=1\\ j\neq i}}^{k} (x_i^{(k)} - x_j^{(k)}) + \sum_{\substack{j=1\\ j\neq i}}^{n} f(x_j^{(k)}) (x_j^{(k)} - x_i^{(k)})^{-1} \prod_{\substack{s=1\\ s\neq j\\ s\neq i}}^{k} (x_i^{(k)} - x_s^{(k)}) (x_j^{(k)} - x_s^{(k)})^{-1})$$

$$i=1, \ldots, n; \; k=0, 1, 2, \ldots,$$

которые можно получить в предельном случае из (7) при $\alpha=\beta=1$. Сейчас выясним вопрос о скорости сходимости последовательности $\{x_i^{(k)}\}_1^n$ из (8).

Для этого нам потребуется следующая теорема [6].

Теорема. *Пусть* $0<q<1$, $d=\min_{i\neq j}|x_i-x_j|>0$, *где* $\{x_i\}_1^n$ — *нули многочлена* $f(x)$ *и* $c>0$ *такое, что выполнено*

(9)
$$0<cen/(d-2c)<1.$$

Если начальные приближения $\{x_i^{(0)}\}_1^n$ *к корням* (2) *удовлетворяют неравенствам*

(10) $$|x_i^{(0)}-x_i|\leqq cq,\ i=1,\dots,n,$$

то для всех k будем иметь

(11) $$|x_i^{(k)}-x_i|\leqq cq^{3^k},\quad i=1,\dots,n,$$

то есть процесс (8) *сходится с кубической скоростью.*

Доказательство. Займемся оценкой разности $x_i^{(k+1)}-x_i$. При помощи простых преобразований для нее получаем приводимое ниже представление:

$$x_i^{(k+1)}-x_i=x_i^{(k)}-x_i-\prod_{j=1}^{n}(x_i^{(k)}-x_j)/\Big(\prod_{\substack{j=1\\ j\neq i}}^{n}(x_i^{(k)}-x_j^{(k)})$$

$$+\sum_{\substack{j=1\\ j\neq i}}^{n}f(x_j^{(k)})(x_j^{(k)}-x_i^{(k)})^{-1}\prod_{\substack{s=1\\ s\neq j\\ s\neq i}}^{n}(x_i^{(k)}-x_s^{(k)})(x_j^{(k)}-x_s^{(k)})^{-1}\Big)$$

$$=(x_i^{(k)}-x_i)\Big(1-\prod_{j=1}^{n}{}^{i}(x_i^{(k)}-x_j)\ (x_i^{(k)}-x_j^{(k)})^{-1}/\Big(1$$

$$+\ \sum_{j=1}^{n}{}^{i}(x_j^{(k)}-x_j)(x_i^{(k)}-x_j^{(k)})^{-1}\prod_{s=1}^{n}{}^{j}(x_j^{(k)}-x_s)(x_j^{(k)}-x_s^{(k)})^{-1}\Big)\Big)$$

$$=(x_i^{(k)}-x_i)\Big(1-\prod_{j=1}^{n}{}^{i}(x_i^{(k)}-x_j)(x_i^{(k)}-x_j^{(k)})^{-1}$$

$$+\ \sum_{j=1}^{n}{}^{i}(x_j^{(k)}-x_j)(x_i^{(k)}-x_j^{(k)})^{-1}\prod_{s=1}^{n}{}^{j}(x_j^{(k)}-x_s)(x_j^{(k)}-x_s^{(k)})^{-1}\Big).$$

$$\times\Big(1+\sum_{j=1}^{n}{}^{i}(x_j^{(k)}-x_j)(x_i^{(k)}-x_j^{(k)})^{-1}\prod_{s=1}^{n}{}^{j}(x_j^{(k)}-x_s)(x_j^{(k)}-x_s^{(k)})^{-1}.$$

Ввиду представления

$$\prod_{j=1}^{n}{}^{i}(x_i^{(k)}-x_j)(x_i^{(k)}-x_j^{(k)})^{-1}-1=\prod_{j=1}^{n}{}^{i}(x_i^{(k)}-x_j)(x_i^{(k)}-x_i^{(k)})^{-1}$$

$$-\prod_{j=1}^{n-1}{}^{i}(x_i^{(k)}-x_j)(x_i^{(k)}-x_j^{(k)})^{-1}+\prod_{j=1}^{n-1}{}^{i}(x_i^{(k)}-x_j)(x_i^{k}-x_j^{(k)})^{-1}-1$$

$$=(x_n^{(k)}-x_n)(x_i^{(k)}-(x_n^{(k)})^{-1}\prod_{j=1}^{n-1}{}^{i}(x_i^{(k)}-x_j)(x_i^{(k)}-x_j^{(k)})^{-1}$$

$$+\prod_{j=1}^{n-1}{}^{i}(x_i^{(k)}-x_j)(x_i^{(k)}-x_j^{(k)})^{-1}-1=\cdots$$

$$=\sum_{s=1}^{n}{}^{i}(x_s^{(k)}-x_s)(x_i^{(k)}-x_s^{(k)})^{-1}\prod_{j=1}^{s-1}{}^{i}(x_i^{(k)}-x_j)(x_i^{(k)}-x_j^{(k)})^{-1}$$

$$\Big(\prod_{j=1}^{0}(x_i^{(k)}-x_j)(x_i^{(k)}-x_j^{(k)})^{-1}=1\Big),$$

получаем

$$S = 1 - \prod_{j=1}^{n}{}^{i}(x_i^{(k)} - x_j)(x_i^{(k)} - x_j^{(k)})^{-1}$$

$$+ \sum_{j=1}^{n}{}^{i}(x_j^{(k)} - x_j)(x_i^{(k)} - x_j^{(k)})^{-1} \prod_{s=1}^{n}{}^{j}(x_j^{(k)} - x_s)(x_j^{(k)} - x_s^{(k)})^{-1}$$

$$= \sum_{j=1}^{n}{}^{i}(x_j^{(k)} - x_j)(x_i^{(k)} - x_j^{(k)})^{-1}\Big(\prod_{s=1}^{n}{}^{j}(x_j^{(k)} - x_s)(x_j^{(k)} - x_s^{(k)})^{-1}$$

$$- \prod_{s=1}^{j-1}{}^{i}(x_i^{(k)} - x_s)(x_i^{(k)} - x_s^{(k)})^{-1} - 1 + 1\Big)$$

$$= \sum_{j=1}^{n}{}^{i}(x_j^{(k)} - x_j)(x_i^{(k)} - x_j^{(k)})^{-1}\Big(\prod_{s=1}^{n}{}^{j}(x_j^{(k)} - x_s)(x_j^{(k)} - x_s^{(k)})^{-1} - 1\Big)$$

$$+ \sum_{j=1}^{n}{}^{i}(x_j^{(k)} - x_j)(x_i^{(k)} - x_j^{(k)})^{-1}\Big(1 - \prod_{s=1}^{j-1}{}^{i}(x_i^{(k)} - x_s)(x_i^{(k)} - x_t^{(k)})^{-1}\Big)$$

$$= \sum_{j=1}^{n}{}^{i}(x_j^{(k)} - x_j)(x_i^{(k)} - x_j^{(k)})^{-1} \sum_{l=1}^{n}{}^{j}(x_l^{(k)} - x_l)(x_j^{(k)} - x_l^{(k)})^{-1}$$

$$\times \prod_{s=1}^{l-1}{}^{j}(x_j^{(k)} - x_s)(x_j^{(k)} - x_s^{(k)})^{-1} + (-1)\sum_{j=2}^{n}{}^{i}(x_j^{(k)} -$$

$$- x_j)(x_i^{(k)} - x_j^{(k)})^{-1} \sum_{l=1}^{j-1}{}^{i}(x_l^{(k)} - x_l)(x_i^{(k)}$$

$$- x_l^{(k)})^{-1} \prod_{s=1}^{l-2}{}^{i}(x_i^{(k)} - x_s)(x_i^{(k)} - x_s^{(k)})^{-1}.$$

Докажем, что для последовательности $\{x_j^{(k)}\}_1^n$ верны неравенства (11). Для этого воспользуемся индукцией. При $k=0$, ввиду (10), неравенство (11) выполняется. Допустим теперь, что (11) выполняется для $k=m$.

Далее, учитывая ограничение (9) и неравенства

$$|x_i^{(n)} - x_i| \leq cq^{3^n}; \quad i = 1, \ldots, n,$$

$$|x_i^{(m)} - x_j^{(m)}| \geqq |x_i - x_j| - |x_i^{(m)} - x_i| - |x_j^{(m)} - x_j| > d - 2c \quad \text{при} \quad i \neq j,$$

$$\Big|\prod_{j=1}^{n}{}^{i}(x_i^{(k)} - x_j)(x_i^{(k)} - x_j^{(k)})^{-1}\Big| \leqq (1 + cq^{3^m}/(d-2c))^n \leqq (1 + c/(d-2c))^n$$

$$\leqq (1 + (en)^{-1})^n \leqq e^{1/e},$$

$$\Big|\sum_{j=1}^{n}{}^{i}(x_j^{(k)} - x_j)(x_i^{(k)} - x_j^{(k)})^{-1} \prod_{s=1}^{n}{}^{j}(x_j^{(k)} - x_s)(x_j^{(k)} - x_s^{(k)})^{-1}$$

$$\leqq ncq^{3^m} e^{1/e}(d-2c)^{-1} < e^{1/e} - 1,$$

получаем

$$| x_i^{(m+1)} - x_i | \leq cq^{3^m} 2\pi^2 c^2 q^{2.3^m} e^{1/e} (1 - e^{1/e-1} (d-2c)^{-2}$$

$$\leq cq^{3^{m+1}} 2n^2 c^2 e^2 e^{1/e-2} (1 - e^{1/e-1})(d-2c)^{-2} \leq cq^{3^{m+1}} 2e^{1/e-2} (1 - e^{1/e-1})^{-1} < cq^{3^{m+1}}.$$

Этим завершается индуктивное доказательство оценки скорости сходимости.

Приведем пример применения этого метода. Пусть дано уравнение $f(x) = x^3 - 6x^2 + 11x - 6 = 0$. Корни этого уравнения $x_1 = 1$, $x_2 = 2$, $x_3 = 3$.

В качестве начальных приближений выбрано $x_1^{(0)} = 0 \cdot 9$, $x_2^{(0)} = 2 \cdot 1$, $x^{(30)} = 2 \cdot 9$.

Таблица 2

k	i	$x_i^{(k)}$ по ехеме (3)	$x_i^{(k)}$ — по (8)	$f\left(x_i^{(k)}\right)$
0	1	0.9	0.9	—0.2310100627040882
1	1	0.(2 * 9)6250063478995	0.(3 * 9)4832087907639	—0.(2 * 0)34383775954770
2	1	1.(4 * 0)48711347103	1.(8 * 0)2594623	0.(8 * 0)51892534
3	1	1.(9 * 0)920280	0.(14 * 9) 89	—0.(14 * 0)2220446
4	1	0.(14 * 9)87	1	0
5	1	1		
0	2	2.1	2.1	—0.09899944496144997
1	2	1.9878822238242753	1.(2 * 9)1901649738485	0.(2 * 0)80978191451634
2	2	2.(4 * 0)18859067776	2.(5 * 0)4650573106	—0.(5 * 0)4650573107050567
3	2	2.(11 * 0)1210	2.(13 * 0)11	—0.(13 * 0)1067814103640148
4	2	2.(14 * 0)3	2	0
5	2	2		
0	3	2.9	2.9	—0.1710005455013182
1	3	2.(2 * 9)8477103768902	2.(2 * 9)3780228388789	—0.01232372716108541
2	3	3.(7 * 0)65885793	3.(4 * 0)53096002086	0.(3 * 0)1062004618894716
3	3	3.(14 * 0)9	3.(9 * 0)246963	0.(9 * 0)493926677336
4	3	3	3	0

В табл. 2 приведены вычислительные эксперименты, проведенные по схемам (3) и (8). Если цифра l последовательно повторяется в записи данного числа m раз, то в табл. 2 это записывается как $| m * l |$.

Итерационный процесс заканчиваем тогда с точностью $\varepsilon=10^{-16}$ совпадут значения $x_i^{(\nu)}$ и $x_i^{(\nu+1)}$. Все вычисления выполнены на ЭВМ ЕВ-1040 с двойной точностью. Заметим, что Таблица 2 приведена в полном обьеме только из-за методических соображений, проанализировав рост числа верных знаков в нескольких шагах. Вычислительные эксперименты, проведенные по алгоритму (8), показали его надежную и быструю работу.

Если 1 — объем работы при вычислении значения $f(x)$ в точке, а θ_1 — дополнительный объем работы при вычислении производной $f'(x)$, то индексами эффективности в смысле А. Островского [5] будут: в случае метода (3) — порядка $2^{1/n}$, в случае метода (5) — порядка $3^{(n(1+\theta_1))^{-1}}$, а в случае процесса (8) — порядка $3^{1/n}$. Видно, что при $\theta_1>0{,}58496\ldots$ целесообразнее использовать метод Дочева.

В заключение отметим, что из анализа численных экспериментов можно сделать следующие замечания:

а. Процесс (8) оказывается сходящим и в случае, когда неравенство (9) не выполнено. Результаты слабо зависели от способа выбора $\{x_i^{(0)}\}_1^n$.

б. Процесс (8) можно использовать для нахождения как действителных, так и комплексных корней, поскольку формула (8) остается справедливой и в комплексной области.

в. В процессе вычисления приближений $x_\nu^{(k+1)}$ можно использовать те $x_\mu^{(k+1)}$, которые уже вычислены (аналог метода Зайделя).

ЛИТЕРАТУРА

1. Дочев. Видоизменен метод на Нютон за едновременно приблизително пресмятане на всички корени на дадено алгебрично уравнение. *Физич. мат. сп.*, **5**, 1962, 136—139.
2. K. Dočev. Über Newtonsche Iterationen. *C. R. Acad. Bulg. Sci.*, **15**, 1962, 695—701.
3. Дочев, П. Бырнев. О некоторых модификациях метода Ньютона для приближенного решения алгебраических уравнений. *Ж. вычисл. мат. и мат. физ.*, **4**, 1964, 915—920.
4. L. W. Ehrlich. A Modified Newton Method for Polynomials. *CACM*, **10**, 1967, 107—108.
5. А. М. Островский. Решение уравнений и систем уравнений. Москва, 1963.
6. Н. Кюркчиев, С. Ташев. Один метод одновременного приближенного вычисления всех корней алгебраического уравнения. *Доклады БАН*, 1981 (в печати).

Единый центр математики и механики
1090 П. Я. 373

Поступила 17. 3. 1981