MATHEMATICS

1960

SovietRxiv

Abstract

Full Text

MATHEMATICS

A. N. KOSTOVSKII

FORMULAS FOR TRANSFORMING COEFFICIENTS IN LEHMER' S METHOD FOR THE NUMERICAL SOLUTION OF ALGEBRAIC EQUATIONS

(Presented by Academician S. L. Sobolev, 23 XI 1959)

Let an equation with real coefficients be given,

$$f(x) = a_0 + a_1 x + \cdots + a_{nx}^{n} = 0, \qquad a_0 \neq 0, \qquad a_n \neq 0, \tag{1}$$

whose roots are arranged in increasing order of their moduli,

$$0 < |x_1| \leq |x_2| \leq \cdots \leq |x_n|. \tag{2}$$

D. Lehmer (1) gave a new modification of the well-known Lobachevsky–Graeffe method for the numerical solution of equations. He proposed computing two sequences of symmetric functions

$$S_k^{(\nu)}\left(x_1^{-2^\nu} x_2^{-2^\nu} \cdots x_k^{-2^\nu}\right), \qquad \Sigma_k^{(\nu)}\left(x_1^{-2^\nu} \cdots x_{k-1}^{-2^\nu} x_k^{-2^\nu} + 1\right), \qquad k = 1, 2, \ldots, n;\ \nu = 1, 2, \ldots$$

The formulas for computing $\Sigma_k^{(\nu)}$ are different for $\nu = 1$ and $\nu \geq 2$ and differ in their external form from the formulas for computing $S_k^{(\nu)}$.

In the present note it is proposed, instead of computing $\Sigma_k^{(\nu)}$, to compute the symmetric functions

$$\sigma_k^{(\nu)}\left(x_1^{-2^\nu} \cdots x_k^{-2^\nu} x_{k+1}^{-1}\right),$$

which makes it possible, on electronic computers, to compute $S_k^{(\nu)}$ and $\sigma_k^{(\nu)}$ by one and the same program. In addition, an algorithm will be indicated that makes it possible to judge how long it is necessary to carry out the transformation of equations ($\nu = 1, 2, \ldots$) in order to obtain the roots of the given equation with a prescribed accuracy. Transform the given equation (1) by the formulas

$$f_p(x) = f_{p-1}(\varepsilon_0\sqrt{x})f_{p-1}(\varepsilon_1\sqrt{x}), \qquad p = 1, 2, \dots, \nu,$$

where ε_0 and ε_1 are the roots of the equation $x^2+1=0$. We obtain the equation

$$f_\nu(x) = a_0^{(\nu)} + a_1^{(\nu)}x + \dots + a_n^{(\nu)}x^n = a_0^{(\nu)}\left(1+\frac{x}{x_1^m}\right)\left(1+\frac{x}{x_2^m}\right)\cdots\left(1+\frac{x}{x_n^m}\right) = 0, \tag{3}$$

where $m = 2^\nu$, $a_0^{(\nu)} = a_0^m$,

$$a_k^{(p)} = \left(a_k^{(p-1)}\right)^2 + 2\sum_{i=1}^{r}(-1)^i a_{k-i}^{(p-1)}a_{k+i}^{(p-1)}, \qquad r = \min(k, n-k), \tag{4}$$

$$a_k^{(0)} = a_k, \qquad k = 0, 1, 2, \dots, n;$$

$$S_k^{(\nu)}\left(x_1^{-m}x_2^{-m}\cdots x_k^{-m}\right) = a_k^{(\nu)} : a_0^{(\nu)}. \tag{5}$$

Let h be an infinitesimal quantity. Neglecting terms containing h^2 and higher powers of h, we obtain

$$f(x-h) = A_0 + A_1x + \dots + A_{nx}^n = 0, \tag{6}$$

$$b_k = -(k+1)a_{k+1}, \quad A_k = a_k + hb_k, \quad k = 0, 1, 2, \dots, n \quad (b_n = 0). \tag{7}$$

Transforming equation (6) by formulas (4), we obtain

$$f_1(x-h) = A_0^{(1)} + A_1^{(1)}x + \dots + A_n^{(1)}x^n = 0,$$

$$A_k^{(1)} = A_k^2 + 2\sum_{i=1}^{r}(-1)^i A_{k-i}A_{k+i} = \left[a_k^2 + 2\sum_{i=1}^{r}(-1)^i a_{k-i}a_{k+i}\right] +$$

$$+2h\left[a_{kb}k + \sum_{i=1}^{r}(-1)^i\left(a_{k-i}b_{k+i} + a_{k+i}b_{k-i}\right)\right] = a_k^{(1)} + 2hb_k^{(1)}.$$

Having performed ν such transformations of equation (6), we find

$$f_\nu(x-h) = A_0^{(\nu)} + A_1^{(\nu)}x + \dots + A_n^{(\nu)}x^n = 0, \tag{8}$$

$$b_k^{(p)} = a_k^{(p-1)} b_k^{(p-1)} + \sum_{i=1}^{r} (-1)^i \left(a_{k-i}^{(p-1)} b_{k+i}^{(p-1)} + a_{k+i}^{(p-1)} b_{k-i}^{(p-1)} \right), \tag{9}$$

$$r = \min(k, n-k), \quad k = 0, 1, \dots, n; \quad p = 1, 2, \dots, \nu, \quad A_k^{(\nu)} = a_k^{(\nu)} + mhb_k^{(\nu)}. \tag{10}$$

The roots of equation (8) are

$$-\left(x_k^m + mhx_k^{m-1}\right), \quad k = 1, 2, \dots, n, \quad m = 2^\nu. \tag{11}$$

Taking into account that $A_n^{(\nu)} = a_n^{(\nu)}$, we obtain

$$\frac{A_k^{(\nu)}}{A_n^{(\nu)}} = \frac{a_k^{(\nu)} + mhb_k^{(\nu)}}{a_n^{(\nu)}} = S_{n-k}^{(\nu)}(x_1^m \dots x_{n-k}^m) + mhS_{n-k}^{(\nu)}(x_1^m \dots x_{n-k-1}^m x_{n-k}^{m-1}).$$

Dividing the last equality by $a_0^{(\nu)}/a_n^{(\nu)} = x_1^m \dots x_n^m$, we find

$$\frac{b_k^{(\nu)}}{a_0^{(\nu)}} = S_k^{(\nu)}(x_1^{-m} \dots x_k^{-m} x_{k+1}^{-1}). \tag{12}$$

Let $f(x)$ be an integral transcendental function of genus zero (^{2})

$$f(x) = a_0 \prod_{k=1}^{\infty} \left(1 + \frac{x}{x_k}\right) = a_0 + a_1 x + \dots + a_{nx}^n + \dots = 0,$$

$$0 < |x_1| \le |x_2| \le \cdots.$$

The above arguments remain valid in this case. We shall obtain formula (12), taking into account that $b_0^{(\nu)} = -a_0^{m-1} a_1$, and assuming that the infinitely small quantity $h < |x_1|$, i.e. $A_0^{(\nu)} \ne 0$.

$$\frac{a_0^{(\nu)}}{A_0^{(\nu)}} = \frac{a_0^m}{a_0^m - mha_0^{m-1} a_1} = \frac{1}{1 + mh\left(\frac{1}{x_1} + \frac{1}{x_2} + \cdots\right)};$$

$$\frac{A_k^{(\nu)}}{A_0^{(\nu)}} = S_k^{(\nu)} \left(\frac{1}{(x_1^m + mhx_1^{m-1}) \dots (x_k^m + mhx_k^{m-1})} \right) =$$

$$= S_k^{(\nu)} \left(\frac{1}{x_1^m x_2^m \dots x_k^m + mh\left(x_1^m \dots x_{k-1}^m x_k^{m-1} + x_1^{m-1} x_2^m \dots x_k^m\right)} \right),$$

whence

$$\frac{A_k^{(\nu)}}{a_0^{(\nu)}} = \frac{a_k^{(\nu)}}{a_0^{(\nu)}} + mh\frac{b_k^{(\nu)}}{a_0^{(\nu)}} =$$

$$= S_k^{(\nu)} \left(\frac{\left[1 + mh\left(\frac{1}{x_1} + \frac{1}{x_2} + \cdots\right)\right]\left[x_1^m \ldots x_k^m - mh\left(x_1^m \ldots x_{k-1}^m x_k^{m-1} + \cdots\right)\right]}{x_1^{2m} x_2^{2m} \ldots x_k^{2m}} \right) =$$

$$= S_k^{(\nu)}\left[x_1^{-m} \ldots x_k^{-m} + mhx_1^{-m} \ldots x_k^{-m}(x_{k+1}^{-1} + x_{k+2}^{-1} + \cdots)\right] =$$

$$= S_k^{(\nu)}(x_1^m \ldots x_k^m) + mhS_k^{(\nu)}(x_1^{-m} \ldots x_k^{-m} x_{k+1}^{-1}),$$

i.e. formula (12) remains valid.

Formula (4) can be written in the form

$$a_k^{(\nu)} = a_k^{(\nu-1)} a_k^{(\nu-1)} + \sum_{i=1}^{r} (-1)^i \left(a_{k-i}^{(\nu-1)} a_{k+i}^{(\nu-1)} + a_{k+i}^{(\nu-1)} a_{k-i}^{(\nu-1)}\right);$$

Table 1

	a_0	b_0	a_1	b_1	a_2	b_2	a_3	b_3
Initial values	$a_0 = 7$	$b_0 = -a_1 = -8$	$a_1 = 8$	$b_1 = -2a_2 = -16$	$a_2 = 8$	$b_2 = -3a_3 = -3$	$a_3 = 1$	$b_3 = 0$
$\nu = 1$	$a_0^{(1)} = 49$	$b_0^{(1)} = -56$	$a_1^{(1)} = -48$	$b_1^{(1)} = -43$	$a_2^{(1)} = 48$	$b_2^{(1)} = -8$	$a_3^{(1)} = 1$	$b_3^{(1)} = 0$
$\nu = 2$	$a_0^{(2)} = 2401$	$b_0^{(2)} = -2744$	$a_1^{(2)} = -2400$	$b_1^{(2)} = 5144$	$a_2^{(2)} = 2400$	$b_2^{(2)} = -341$	$a_3^{(2)} = 1$	$b_3^{(2)} = 0$
$\nu = 3$	$a_0^{(3)} = 5\,764\,801$	$b_0^{(3)} = -6\,588\,344$	$a_1^{(3)} = -5\,764\,800$	$b_1^{(3)} = -4\,941\,255$	$a_2^{(3)} = 5\,764\,800$	$b_2^{(3)} = -823\,544$	$a_3^{(3)} = 1$	$b_3^{(3)} = 0$
$\nu = 4$	$a_0^{(4)} =$ 33 232 930 [illegible]	$b_0^{(4)} =$ [illegible]	$a_1^{(4)} <$ [illegible] 1 492 079 [illegible] 544	$b_1^{(4)} > 0$	$a_2^{(4)} =$ 33 232 930 [illegible]	$b_2^{(4)} =$ [illegible] 61 509 [illegible]44	$a_3^{(4)} = 1$	$b_3^{(4)} = 0$

Therefore, a program compiled for computing on an electronic digital machine the quantities $b_k^{(\nu)}$ by formulas (9) can also be used for computing the coefficients $a_k^{(\nu)}$, after first sending into the cells $\langle b_k^{(\nu)} \rangle$ the quantities $a_k^{(\nu)}$.

From (5) and (12), for a sufficiently large number of transformations ν, it follows:

a) if $\ll |x_{k-1}| < |x_k| < |x_{k+1}| \ll$, then

$$x_k = \left(\frac{b_{k-1}^{(\nu)}}{a_{k-1}^{(\nu)}} - \frac{b_k^{(\nu)}}{a_k^{(\nu)}} \right)^{-1}; \tag{13}$$

b) if $x_k = \overline{x}_{k+1} = \rho e^{i\varphi}$, $\ll |x_{k-1}| < \rho < |x_{k+2}| \ll$, then

$$\rho = \left[a_{k-1}^{(\nu)} / a_{k+1}^{(\nu)} \right]^{1/2^{\nu+1}}; \tag{14}$$

$$\cos\varphi = \frac{\rho}{2} \left(\frac{b_{k-1}^{(\nu)}}{a_{k-1}^{(\nu)}} - \frac{b_{k+1}^{(\nu)}}{a_{k+1}^{(\nu)}} \right). \tag{15}$$

Let now, in the inequalities (2), the strict inequality $|x_k| < |x_{k+1}|$ hold. Then, for a polynomial of degree n, from (5), (9), and (12) it follows that

$$\lim_{\nu\to\infty} \frac{b_k^{(\nu+1)}}{a_k^{(\nu)} b_k^{(\nu)}} = \lim_{\nu\to\infty} x_1^{-2^{\nu+1}} \dots$$

$$\dots x_k^{-2^{\nu+1}} x_{k+1} b_k^{(\nu+1)} = 1. \tag{16}$$

Thus, coefficients of the transformed equation that change correctly and incorrectly can be judged not only from the table for computing the coefficients $a_k^{(\nu)}$, but also from the table in which the auxiliary quantities $b_k^{(\nu)}$ are computed.

From (4) and (9) the rule (3) immediately follows.

To compute the moduli of the roots of equation (1) to no fewer than μ correct digits, the transformation of the given equation should be continued until, in each of the regularly changing coefficients, the sums

$$2\sum_{i=1}^{r} (-1)^i a_{k-i}^{(\nu)} a_{k+i}^{(\nu)} \quad \text{and} \quad \sum_{i=1}^{r} (-1)^i \left(a_{k-i}^{(\nu)} b_{k+i}^{(\nu)} + a_{k+i}^{(\nu)} b_{k-i}^{(\nu)} \right)$$

cease to have an effect on the μ leading digits of the quantities $\left(a_k^{(\nu)}\right)^2$ and $a_k^{(\nu)} \cdot b_k^{(\nu)}$.

For the case considered by Lehmer[1], $b_k = (k-1)a_{k-1}$, $k = 0, 1, \dots, \dots, n+1$, one should, instead of equation (6), take the equation $f\left(\frac{1}{x-h}\right) = 0$, whose roots will be $-\left(x_i^{-m} + mhx_i^{-m+1}\right)$, $i = 1, 2, \dots, n$. In a manner analogous to that presented above, we obtain the same coefficient transformation formulas; however, here one must set

$$r = \min(k, n-k+1), \tag{17}$$

since $b_{n+1}^{(\nu)} = na_n \neq 0$, $b_0^{(\nu)} = b_{n+2}^{(\nu)} = \cdots = 0$, $\nu = 0, 1, 2, \ldots$.

The r determined by equality (17) is needed only for computing the auxiliary quantities $b_k^{(1)}$; in all other cases r may be determined in the same way as in (4) and (9).

Example. Compute the roots of the equation

$$f(x) = 7 + 8x + 8x^2 + x^3$$

to no fewer than 5 correct digits.

From Table 1 we see that the coefficients $a_0^{(\nu)}$, $a_2^{(\nu)}$, and $a_3^{(\nu)}$, as well as the auxiliary quantities $b_0^{(\nu)}$, $b_2^{(\nu)}$, and $b_3^{(\nu)}$, change regularly; the quantities $a_1^{(\nu)}$ and $b_1^{(\nu)}$ change sign in the process of transforming the equations; hence $x_1 = \bar{x}_2 = \rho e^{i\varphi}$. By formulas (13), (14), and (15) we find ($\nu = 4$)

$$\rho = \sqrt[2^5]{\frac{a_0^{(4)}}{a_2^{(4)}}} = \sqrt[32]{\frac{33\,232\,930\,569\,601}{33\,232\,930\,569\,600}} = 1.000\,000\,000\,000,$$

$$\cos\varphi = \frac{\rho}{2}\left[\frac{b_0^{(4)}}{a_0^{(4)}} - \frac{b_2^{(4)}}{a_2^{(4)}}\right] = \frac{1}{2}\left[-\frac{37\,980\,492\,079\,544}{33\,232\,930\,569\,601} + \frac{32\,232\,930\,569\,600}{4\,747\,561\,509\,941}\right] = -0.500\,000\,000\,000,$$

$$x_3 = \left(\frac{b_2^{(4)}}{a_2^{(4)}} - \frac{b_3^{(4)}}{a_3^{(4)}}\right)^{-1} = \left(-\frac{4\,747\,561\,509\,941}{33\,232\,930\,569\,600} - \frac{0}{1}\right)^{-1} = -7.000\,000\,000\,000,$$

$$x_1 = \bar{x}_2 = -\frac{1}{2} + \frac{1}{2}i\sqrt{3}, \qquad x_3 = -7.$$

Remark. The reasoning given above obviously remains valid also for equation (1) with complex coefficients. D. Lehmer' s observation that his method can be applied to computing zeros of integral functions that are small in modulus remains valid also for the results of the present paper.

Received
16 XI 1959

REFERENCES

[1] D. Lehmer, *Mathematical Tables and Other Aids to Computation*, **1**, 377 (1945).
[2] G. Pólya, *Zs. Math. Phys.*, **63**, 1-2, 275 (1914).

[3] A. N. Krylov, *Lectures on Approximate Computations*, Moscow–Leningrad, 1950.

Note: Figure translations are in progress. See original paper for figures.