ON THE SEPARATION AND COMPUTATION OF THE ROOTS OF AN ALGEBRAIC EQUATION

(OB OTDELENII I VYCHISLENII KORNEI ALGEBRAICHESKOGO URAVNENIIA)

PMM Vol.26, No.4, 1962, pp. 772-774

L.M. MARKHASHOV
(Moscow)

(Received April 2, 1962)

In applied problems there arises frequently the necessity for the evaluation of roots of algebraic equations. In the present note there are considered certain methods for the separation and evaluation of roots of such equations.

1. The well-known methods for locating the roots of an algebraic equation [1,2], such as the classical methods of Sturm and Budan-Fourier, yield solutions to the problem of the determination of the number of roots lying within an arbitrarily given region (or within an interval of the real axis).

A problem that is in a certain sense the inverse of the above problem, is that of finding regions (or intervals) each of which will contain exactly one root, independently of the specialization of the coefficients of the equation.

Let us consider the following equation with real coefficients:

$$z^n + a_1 z^{n-1} + \ldots + a_n = 0$$

all of whose roots are real, $\lambda_1 < \lambda_2 < \ldots < \lambda_n$.

The problem consists of finding $n - 1$ rational functions $\mu_\nu(a_1, \ldots, a_n)$ such that any definite relation $\mu_1 < \mu_2 < \ldots < \mu_{n-1}$, that may arise for any particular choice of the a's, will imply the following inequalities

$$-\infty < \lambda_1 < \mu_1 < \lambda_2 < \mu_2 < \ldots < \lambda_{n-1} < \mu_{n-1} < \lambda_n < \infty \qquad (1.1)$$

That is, if there exist such functions μ then they represent "movable" boundaries for the roots; in this case the structure of the functions μ

may change only with the degree of the equations. We shall show that such functions μ exist at least for third degree equations

$$f(z) \equiv z^3 + a_1 z^2 + a_2 z + a_3 = 0 \tag{1.2}$$

Indeed, let us choose for the functions μ the following functions of the coefficients of Equation (1.2):

$$\mu' = -\frac{a_1}{3}, \qquad \mu'' = \frac{a_1 a_2 - 9a_3}{6a_2 - 2a_1^2} \tag{1.3}$$

If we denote by Δ_1, Δ_2, Δ_3 the sequence of the principal minors in ascending order of the Hankel matrix of Equation (1.2)

$$\begin{Vmatrix} 3 & s_1 & s_2 \\ s_1 & s_2 & s_3 \\ s_2 & s_3 & s_4 \end{Vmatrix}$$

then one can show the validity of the following identities

$$f(\mu') = (\mu'' - \mu')\frac{\Delta_2}{\Delta_1^2}, \qquad f(\mu'') = (\mu' - \mu'')\frac{\Delta_1 \Delta_3}{\Delta_2^2} \tag{1.4}$$

If all the roots of the Equation (1.2) are real, then

$$\Delta_i > 0 \qquad (i = 1, 2, 3) \tag{1.5}$$

and from (1.4) it follows that

$$\begin{aligned} -\infty < \lambda_1 < \mu' < \lambda_2 < \mu'' < \lambda_3 < \infty, \quad & \text{if} \quad \mu' < \mu'' \\ -\infty < \lambda_1 < \mu'' < \lambda_2 < \mu' < \lambda_3 < \infty, \quad & \text{if} \quad \mu'' < \mu' \\ \mu' = \lambda_2 = \mu'', \quad & \text{if} \quad \mu' = \mu''. \end{aligned}$$

We note that the denominators in (1.3) differ only in sign from Δ_1 and Δ_2, and that they do not vanish, in view of (1.5).

It should be mentioned that the expressions chosen for μ'' and μ' coincide with the values of double or triple roots of Equation (1.2) whenever such roots exist. The latter are found by means of simple rational operations; μ'' is found by solving for the system of equations

$$f(\mu) = 0, \qquad f'(\mu) = 0, \qquad \mu f'(\mu) = 0$$

in which the different powers of μ are treated as independent variables; μ' is found from the equation $f''(\mu) = 0$. The expression for a root of multiplicity k of an nth degree equation can be found in an analogous way. The corresponding process is, however, not single-valued, the value of the multiple root can be determined to within an additive term that will go to zero together with the resultants $R(f, f')$, $R(f', f'')$, ..., which correspond to the multiplicity of the root.

This circumstance complicates the application of the described procedure in case $n \geqslant 4$.

2. In automatic control theory it is frequently required to find the trajectories of the roots of the characteristic equation constructed for a linear system with a varying parameter. Sometimes it is necessary to find the roots of an algebraic equation for the application of a certain analytic algorithm.

In such cases, it is convenient to have a sufficiently simple and precise analytic relation between the roots and coefficients of the given equation.

For the derivation of such a relation, one can use a procedure similar to the method of a small parameter of Poincaré. The only difference between its use for differential and algebraic equations is that the expansion of the solution of an algebraic equation into a power series of the parameter can be accomplished directly by evaluating the derivatives whose existence is guaranteed. One can determine the regions of analyticity of the solution, and hence the regions of convergence of the series representing the solution.

Let us consider an algebraic equation whose real ooefficients depend on a parameter

$$f(\mu, z) = 0$$

Differentiating the left-hand side with respect to the parameter, we obtain

$$f_z' \frac{\partial z}{\partial \mu} + f'_\mu = 0, \quad f_z' \frac{\partial^2 z}{\partial \mu^2} + f_z'' \left(\frac{\partial z}{\partial \mu}\right)^2 + 2 f''_{\mu z} \frac{\partial z}{\partial \mu} + \frac{\partial^2 f}{\partial \mu^2} = 0, \ldots$$

It can be seen that the sequence of derivatives $\partial z/\partial \mu$, $\partial^2 z/\partial \mu^2$, ... can be evaluated for arbitrary values μ° for which z°, the root of the equation $f(\mu^\circ, z^\circ) = 0$, is finite*, and for which $f_z'(\mu^\circ, z^\circ) \neq 0$. Hence it follows that the series

$$z(\mu) = z^\circ + \left(\frac{\partial z}{\partial \mu}\right)^\circ (\mu - \mu^\circ) + \frac{1}{2!}\left(\frac{\partial^2 z}{\partial \mu^2}\right)^\circ (\mu - \mu^\circ)^2 + \ldots$$

---

* The critical algebraic points which are branch points of the solution are the only singular points that the solutions of the algebraic equation in this case can have; the existence of poles is excluded by the requirement that z° be finite.

converges and represents an analytic function for all μ for which the resultant $R(f,\ f')$ retains the sign which it has when $\mu = \mu^0$.

Other parameters on which the coefficients of the equation might depend must be considered as fixed quantities. Such parameters may be, in particular, the coefficients themselves. It is desirable to select for μ^0 a value that will simplify as much as possible the evaluation of z^0.

We shall now demonstrate the described procedure on an example involving the cubic equation

$$z^3 + pz + q = 0$$

Considering q as a parameter, and starting out from its zero value, we obtain

$$z_1 = -\frac{q}{p} - \frac{q^3}{p^4} - 3\frac{q^5}{p^7} + \ldots$$

$$z_2 = \sqrt{-p} + \frac{q}{2p} + \frac{3}{8}\frac{\sqrt{-p}}{p^3}q^2 - \frac{1}{2}\frac{q^3}{p^4} - \frac{105}{128}\frac{\sqrt{-p}}{p^6}q^4 + \ldots$$

$$z_3 = -\sqrt{-p} + \frac{q}{2p} - \frac{3}{8}\frac{\sqrt{-p}}{p^3}q^2 - \frac{1}{2}\frac{q^3}{p^4} + \frac{105}{128}\frac{\sqrt{-p}}{p^6}q^4 + \ldots$$

The series converge when $R(f,\ f') \equiv -4p^3 + 27q^2 < 0,\ p > 0$ (or else when $-4p^3 + 27q^2 > 0,\ p < 0$). The convergence takes place faster if the inequality $pR > 0$ is satisfied with a greater margin.

3. In one of the more effective procedures for evaluating the roots of an algebraic equation, the method of Lobachevskii, it is necessary to compute the symmetric functions of the powers of the roots in terms of the coefficients of the equation. For this purpose there is usually suggested an algorithmic process, which leads, however, to a rather complicated result.

Making use of the recurrence relations of Newton, one can prove the validity of the following simple matrix representations for the required symmetric functions.

For a second degree polynomial one has

$$\lambda_1^k + \lambda_2^k = \begin{vmatrix} a_1 & 1 & 0 & 0 & \ldots \\ 2a_2 & a_1 & 1 & 0 & \ldots \\ 0 & a_2 & a_1 & 1 & \ldots \\ 0 & 0 & a_2 & a_1 & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \end{vmatrix}_k, \qquad \lambda_1^k \lambda_2^k = \begin{vmatrix} a_2 & a_1 & 1 & \ldots \\ 0 & a_2 & a_1 & \ldots \\ 0 & 0 & a_2 & \ldots \\ \ldots & \ldots & \ldots & \ldots \end{vmatrix}_k$$

For a third degree polynomial one obtains

$$\sum_{j=1}^{3} \lambda_j^k = \begin{vmatrix} a_1 & 1 & 0 & 0 & \cdots \\ 2a_2 & a_1 & 1 & 0 & \cdots \\ 3a_3 & a_2 & a_1 & 1 & \cdots \\ 0 & a_3 & a_2 & a_1 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{vmatrix}_k \qquad \lambda_1^k \lambda_2^k \lambda_3^k = \begin{vmatrix} a_3 & a_2 & a_1 & \cdots \\ 0 & a_3 & a_2 & \cdots \\ 0 & 0 & a_3 & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{vmatrix}_k$$

$$\sum_{i=j}^{3} \lambda_i^k \lambda_j^k = \begin{vmatrix} a_2 & a_1 & 1 & 0 & \cdots \\ 2a_3 & a_2 & a_1 & 1 & \cdots \\ 0 & a_3 & a_2 & a_1 & \cdots \\ 0 & 0 & a_3 & a_2 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{vmatrix}_k + \begin{vmatrix} a_3 & a_2 & a_1 & \cdots \\ 0 & a_3 & a_2 & \cdots \\ 0 & 0 & a_3 & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{vmatrix}_{k-1}$$

For a polynomial of the nth degree this representation has the form

$$\sum_{j=1}^{n} \lambda_j^k = \begin{vmatrix} a_1 & 1 & 0 & 0 & \cdots \\ 2a_2 & a_1 & 1 & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ na_n & a_{n-1} & a_{n-2} & a_{n-3} & \cdots \\ 0 & a_n & a_{n-1} & a_{n-2} & \cdots \end{vmatrix}_k$$

(The subscript indicates the order of the determinants.)

The easily recognized forms of the above given determinants indicate how one can obtain more general results for nth degree polynomials.

BIBLIOGRAPHY

1. Parodi, M., *Lokalizatsiia kharacteristicheskikh chisel matrits i ee primenenie (Locating the Characteristic Numbers of a Matrix and its Application)*. Izd-vo inostr. lit-ry, 1960.

2. Lanczos, C., *Applied Analysis*. Prentice-Hall, 1956.

3. Krylov, A.N., *Lektsii o priblizhennykh vychisleniiakh (Lectures on Approximate Computations)*. Gostekhteorizdat, 1950.

Translated by H.P.T.