The Effect of Parameters on the Roots of an Equation System

By C. E. Maley

One of the consequences of automatic computation is that results are given to the full digital capacity of the engine, whether or not such precision is warranted by the quality of the input parameters or by the cancellation and roundoff errors encountered. It is not difficult, however, in the case of equation systems, to program an evaluation of the results. The remarks of this paper also apply to manual computation.

Wilkinson (1960) has recently considered the effects of computational error, in the usual floating-point mode. Let one here consider the sensitivity of the roots to possible errors in the given parameters.

Because it presents a special case, the solution of a polynomial equation will be first examined:

$$f(x) \equiv a_0x^n + a_1x^{n-1} + \dots + a_jx^{n-j} + \dots + a_n = 0 \tag{1}$$

$$\equiv a_0(x - r_1)(x - r_2)\dots(x - r_i)\dots(x - r_n) \tag{2}$$

The possibility of factorization (2) in terms of the roots is peculiar to (1).

Referring to the function $f(x)$:

$$(\partial f/\partial a_j)_{x=r_i} = r_i^{n-j}; \tag{3}$$

$$\left.\begin{aligned}(\partial f/\partial x)_{x=r_i} &= a_0 \prod_{k \neq i}^{1,\,n} (r_i - r_k), \quad r_i \text{ a simple root,}\\ &= 0, \qquad\qquad\qquad\quad r_i \text{ a multiple root.}\end{aligned}\right\} \tag{4}$$

Referring to the equation $f(x) = 0$, x is thereby defined as an n-valued function of the $(n + 1)$ coefficients or, equivalently, as n single-valued functions,

$$x(a_0, a_1, \dots, a_j, \dots, a_n) \equiv r_i(a_0, a_1, \dots, a_j, \dots, a_n), \tag{5}$$

for any value of x for which $\partial f/\partial x$ does not vanish. As (4) indicates a multiple root at such exceptional points, these exceptions may be removed by adopting the usual convention of multiplicity.

A n-dimensional vector field of roots over the $(n + 1)$-dimensional region of coefficients, (5) is the polynomial root space that has claimed the attention of mathematicians for the past two thousand years.

The key to its topography is its vector gradient, a rectangular Jacobian matrix:

$$\partial r_i/\partial a_j \equiv (\partial x/\partial a_j)_{x=r_i} = \\ -[(\partial f/\partial a_j)/(\partial f/\partial x)]_{r=r_i} = -\left\{r_i^{n-j}\Big/\Big[a_0 \prod_{k \neq i}^{1,\,n} (r_i - r_k)\Big]\right\} \tag{6}$$

But this is a precise measure of root sensitivity.

Then the total differential of (5) is, for $x = r_i$,

$$dr_i = -\Big(\sum_j^{0,\,n} r_i^{n-j} da_j\Big)\Big/\Big[a_0 \prod_{k \neq i}^{1,\,n} (r_i - r_k)\Big], \tag{7}$$

the formula for absolute error dr_i due to small errors da_j in the parameters. The formula for relative error is got by dividing both sides of (7) by r_i.

An application of (7) to an example, as

$$f(x) \equiv x^3 - 11\cdot 1x^2 + 11\cdot 1x - 1 = 0, \\ r_1 = 0\cdot 1,\ r_2 = 1,\ r_3 = 10, \tag{8}$$

is likely to reveal unsuspected weaknesses in a calculating engine of fixed digital capacity. For,

$$\frac{dr_1}{r_1} = -\frac{0\cdot 1^3 da_0 + 0\cdot 1^2 da_1 + 0\cdot 1 da_2 + da_3}{0\cdot 1(0\cdot 1 - 1)(0\cdot 1 - 10)}, \text{ etc.,}$$

or, more concisely by (6),

$$\begin{bmatrix} ar_1/r_1 \\ dr_2/r_2 \\ dr_3/r_3 \end{bmatrix} = \begin{bmatrix} -0\cdot 0011223 & -0\cdot 0112233 & -0\cdot 1122335 & -1\cdot 1223345 \\ 0\cdot 12345679 & 0\cdot 12345679 & 0\cdot 12345679 & 0\cdot 12345679 \\ -1\cdot 1223345 & -0\cdot 1122335 & -0\cdot 0112233 & -0\cdot 0011223 \end{bmatrix} \begin{bmatrix} da_0 \\ da_1 \\ da_2 \\ da_3 \end{bmatrix}.$$

These are relative errors. In contrast, the example

$$f(x) \equiv x^3 - 3x^2 + 2\cdot 99x - 0\cdot 99 = 0, \\ r_1 = 0\cdot 9,\ r_2 = 1,\ r_3 = 1\cdot 1, \tag{9}$$

produces

$$\begin{bmatrix} dr_1/r_1 \\ dr_2/r_2 \\ dr_3/r_3 \end{bmatrix} = \begin{bmatrix} -40\cdot 5 & -45 & -50 & -55\cdot 555556 \\ 100 & 100 & 100 & 100 \\ -60\cdot 5 & -55 & -50 & -45\cdot 454546 \end{bmatrix} \begin{bmatrix} da_0 \\ da_1 \\ da_2 \\ da_3 \end{bmatrix}.$$

When r_i is a root of multiplicity s, formula (7) is replaced by

$$dr_i = -\lim_{x \to r_i} \left\{\Big(\sum_j^{0,\,n} x^{n-j} da_j\Big)\Big/\Big[s a_0 (x - r_i)^{s-1} \prod_{k \neq i}^{1,\,n} (x - r_k)\Big]\right\} \tag{10}$$

Thus the root sensitivity is infinite in the neighbourhood of a multiple root, which is otherwise evident by (6).

Multiple roots, which can exist as such only at points of a certain precise locus in the coefficient region, can have little meaning in applied problems where the coefficients are not precisely determinable. Rather, one knows only that the ranges of possibility for s roots centre at the same value. The difficulty may be practically resolved by changing ("perturbing") the value a_j of one or more coefficients within its possibility range $[a_j - da_j, a_j + da_j]$. After the new roots are got, (7) is again employed, using the unsymmetric translated ranges.

In the general case one has a system of possibly non-linear equations involving parameters:

$$\left.\begin{aligned} f^{(1)}(x_1, \ldots, x_n, a_1, \ldots, a_p) &= 0 \\ f^{(2)}(x_1, \ldots, x_n, a_{p+1}, \ldots, a_q) &= 0 \\ \cdots\cdots\cdots\cdots \\ f^{(n)}(x_1, \ldots, x_n, a_{s+1}, \ldots, a_t) &= 0. \end{aligned}\right\} \tag{11}$$

Since the total vector differential is null, it follows that

$$\left[\frac{\partial(f^{(1)}, \ldots, f^{(n)})}{\partial(x_1, \ldots, x_n)}\right]\begin{bmatrix} dx_1 \\ dx_2 \\ \cdot \\ dx_n \end{bmatrix} + \left[\frac{\partial(f^{(1)}, \ldots, f^{(n)})}{\partial(a_1, \ldots, a_t)}\right]\begin{bmatrix} da_1 \\ da_2 \\ \cdot \\ da_t \end{bmatrix} = 0$$

or

$$\{dr_i\} = -\left[\frac{\partial(f^{(1)}, \ldots, f^{(n)})}{\partial(r_1, \ldots, r_n)}\right]^{-1}\left[\frac{\partial(f^{(1)}, \ldots, f^{(n)})}{\partial(a_1, \ldots, a_t)}\right]\{da_j\}. \tag{12}$$

This again is a vector root sensitivity gradient. The second Jacobian matrix in (12) will usually contain zero elements and be rectangular.

This formula is particularly convenient as, whether one has inverted a linear system or employed the Newton–Raphson tangent or secant method (see Maley, 1959), the inverse Jacobian is already at hand. The direct Jacobian matrix will consist of elements that are either simple functions of the known coefficients and roots, or else zeros.

If difficulties are experienced in inverting the matrix that is the heart of the method of solution one may then elect to perturb coefficients.

The thoughtful reader has probably always wondered why the roots of a polynomial equation are invariably obtained one at a time. The reason for this inefficiency appears to be that it is not generally realized that the well-known relations between roots and coefficients may, by the tangent or secant method, be exploited to obtain simultaneously all the roots, real or complex, of a polynomial equation. Thus, the roots r, s, t of (8) are iteratively obtained as

$$\begin{aligned} m_i &= r_i + s_i + t_i - 11\cdot 1 \\ n_i &= r_i s_i + r_i t_i + s_i t_i - 11\cdot 1 \\ p_i &= r_i s_i t_i - 1 \end{aligned} \tag{13}$$

$$[r\ s\ t]_{i+1} = [1\ 0\ 0\ 0]\begin{bmatrix} 1 & m_{i-3} & n_{i-3} & p_{i-3} \\ 1 & m_{i-2} & n_{i-2} & p_{i-2} \\ 1 & m_{i-1} & n_{i-1} & p_{i-1} \\ 1 & m_i & n_i & p_i \end{bmatrix}^{-1}\begin{bmatrix} r_{i-3} & s_{i-3} & t_{i-3} \\ r_{i-2} & s_{i-2} & t_{i-2} \\ r_{i-1} & s_{i-1} & t_{i-1} \\ r_i & s_i & t_i \end{bmatrix} \tag{14}$$

Acknowledgement

The experimental part of this work was carried out on a Univac I, The Carborundum Company, Niagara Falls, New York.

References

WILKINSON, J. H. (1960). "Error Analysis of Floating-Point Computation," *Numerische Mathematik*, Vol. 2, p. 319.
MALEY, C. E. (1959). "A Simplified Numerical Analysis," *Journal of the Royal Aeronautical Society*, Vol. 63, p. 59.

Electronic Computer Exhibition

The second Electronic Computer Exhibition will be held in the National Hall, Olympia, from 3–12 October 1961. *The Computer Bulletin* for September will contain a preview of the Exhibition.

The Symposium on Electronic Data Processing

This Symposium will be held on 4–6 October 1961, and will also be at Olympia. On the first day, a number of users who spoke in the 1958 Symposium will explain the advances that have taken place since that time. On the second day, speakers from various industries will describe the preparation for and installation of up-to-date electronic data-processing systems. The third morning will be taken up by a Brains Trust devoted to the small users and the benefits they get from service centres. On the last afternoon, two or three of Britain's leading authorities will present papers describing some techniques for the immediate future.

The draft programme for the Symposium, and summaries of most of the papers will be published in *The Computer Bulletin* for June.