'ENGLISH ELECTRIC'

NELSON RESEARCH LABORATORIES

BEACONHILL, STAFFORD

Report NS u

©

Date

Order N.5197/11

Reference

Title A Method for the Solution of Non-Linear Simultaneous Equations.

Report by D. Bolton, C.
Mathematics and Theoretical Studies Group.

1. SUMMARY.

The report describes a method for the solution of non-linear simultaneous equations, applicable particularly when the number of equations is small, and the initial estimates are poor.

An example is included in which the method was successfully applied.

With one equation, this is equivalent to "Aitken"

Contents Sheets 1-8.

2. OBJECT OF THE INVESTIGATION.

The investigation was carried out in order to find a reliable method of solving non-linear simultaneous equations which would, in particular, for a small number of equations give convergence even with poor initial estimates.

3. INTRODUCTION.

Let there be n simultaneous equations in n variables $x_1, x_2, \dots x_n$, the equations being

$$\left.\begin{aligned} z_1 &= f_1(x_1, x_2, \dots x_n) = 0 \\ z_2 &= f_2(x_1, x_2, \dots x_n) = 0 \\ &\dots\dots\dots\dots\dots\dots \\ z_n &= f_n(x_1, x_2, \dots x_n) = 0 \end{aligned}\right\} \qquad (1)$$

where $f_1, f_2, \dots f_n$ are any n functions,

$z_1, z_2, \dots z_n$ are values of these functions for specified values of $x_1, x_2, \dots x_n$.

The basis of the method is multi-dimensional inverse interpolation, i.e. finding functions to represent approximately the 'inverse' functions g_j in the region of a root of equations (1), where

$$x_j = g_j(z_1, z_2, \dots z_n) \quad (j = 1 \text{ to } n) \qquad (2)$$

Substituting $z_1 = z_2 = \dots = z_n = 0$ in these functions will then give an approximation to a root $(X_1, X_2, \dots X_n)$ of equations (1). This result is then used to improve the approximations to the 'inverse' functions, and so on, until a root is found with sufficient accuracy.

A simple form for the approximate 'inverse' functions is linear in all the z_j; this has proved unsatisfactory in certain cases. Instead therefore the method described uses polynomial functions:

$$\left.\begin{aligned} &x_j = \text{constant} + \text{linear terms} + \text{quadratic terms} + \text{cubic terms} + \dots \\ &\text{where} \\ &\quad \text{constant} = a_{j0} \\ &\quad \text{linear terms} = a_{j1} z_1 + a_{j2} z_2 + \dots + a_{jn} z_n \\ &\quad \text{quadratic terms} = a_{j,n+1} z_1^2 + a_{j,n+2} z_1 z_2 + \dots + a_{j,2n} z_1 z_n \\ &\qquad\qquad + a_{j,2n+1} z_2^2 + \dots + a_{j,3n-1} z_2 z_n \\ &\qquad\qquad + \dots \qquad\qquad + a_{jp} z_n^2 \\ &\quad \text{cubic terms} = a_{j,p+1} z_1^3 + a_{j,p+2} z_1^2 z_2 + \dots \quad \text{etc.} \end{aligned}\right\} \qquad (3)$$

$$\left[\, p = \tfrac{1}{2}(n^2 + 3n) \,\right]$$

401/2

As many terms, i, of the series (3) are taken as there are estimates of the root available.

4. AN ITERATION - EXAMPLE.

In the description which follows, it will be assumed for simplicity that there are two equations (n = 2) and 4 estimates available (i = 4), although the method is perfectly general. Equations (3) then become

$$\left.\begin{aligned} x_1 &= a_{10} + a_{11} z_1 + a_{12} z_2 + a_{13} z_1^2 \\ x_2 &= a_{20} + a_{21} z_1 + a_{22} z_2 + a_{23} z_1^2 \end{aligned}\right\} \qquad (4)$$

Suppose the estimates are (x_{l1}, x_{l2}) : l = 1 to 4, and corresponding to these, the function values (as calculated by (1)) are (z_{l1}, z_{l2}) : l = 1 to 4.

Substituting these, we obtain 4 linear equations each for a_{1k}, a_{2k} (k = 0 to 3):

$$\left.\begin{aligned} x_{l1} &= a_{10} + a_{11} z_{l1} + a_{12} z_{l2} + a_{13} z_{l1}^2 \quad (l = 1 \text{ to } 4) \\ \text{and} \quad x_{l2} &= a_{20} + a_{21} z_{l1} + a_{22} z_{l2} + a_{23} z_{l1}^2 \quad (l = 1 \text{ to } 4) \end{aligned}\right\} \qquad (5)$$

These equations may be written together in matrix form:

$$\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \\ x_{41} & x_{42} \end{bmatrix} = \begin{bmatrix} 1 & z_{11} & z_{12} & z_{11}^2 \\ 1 & z_{21} & z_{22} & z_{21}^2 \\ 1 & z_{31} & z_{32} & z_{31}^2 \\ 1 & z_{41} & z_{42} & z_{41}^2 \end{bmatrix} \begin{bmatrix} a_{10} & a_{20} \\ a_{11} & a_{21} \\ a_{12} & a_{22} \\ a_{13} & a_{23} \end{bmatrix} \qquad (6)$$

Hence

$$\begin{bmatrix} a_{10} & a_{20} \\ a_{11} & a_{21} \\ a_{12} & a_{22} \\ a_{13} & a_{23} \end{bmatrix} = \begin{bmatrix} 1 & z_{11} & z_{12} & z_{11}^2 \\ 1 & z_{21} & z_{22} & z_{21}^2 \\ 1 & z_{31} & z_{32} & z_{31}^2 \\ 1 & z_{41} & z_{42} & z_{41}^2 \end{bmatrix}^{-1} \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \\ x_{41} & x_{42} \end{bmatrix} \qquad (7)$$

Let the inverse matrix in (7) be denoted by $[b_{\alpha\beta}]$, so that

$$\begin{bmatrix} a_{10} & a_{20} \\ a_{11} & a_{21} \\ a_{12} & a_{22} \\ a_{13} & a_{23} \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ b_{31} & b_{32} & b_{33} & b_{34} \\ b_{41} & b_{42} & b_{43} & b_{44} \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \\ x_{41} & x_{42} \end{bmatrix} \qquad (8)$$

Then:

$$\begin{bmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ b_{31} & b_{32} & b_{33} & b_{34} \\ b_{41} & b_{42} & b_{43} & b_{44} \end{bmatrix} \begin{bmatrix} 1 & z_{11} & z_{12} & z_{11}^2 \\ 1 & z_{21} & z_{22} & z_{21}^2 \\ 1 & z_{31} & z_{32} & z_{31}^2 \\ 1 & z_{41} & z_{42} & z_{41}^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{9}$$

or, transposing the matrices

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ z_{11} & z_{21} & z_{31} & z_{41} \\ z_{12} & z_{22} & z_{32} & z_{42} \\ z_{11}^2 & z_{21}^2 & z_{31}^2 & z_{41}^2 \end{bmatrix} \begin{bmatrix} b_{11} & b_{21} & b_{31} & b_{41} \\ b_{12} & b_{22} & b_{32} & b_{42} \\ b_{13} & b_{23} & b_{33} & b_{43} \\ b_{14} & b_{24} & b_{34} & b_{44} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{10}$$

Now, putting $z_1 = z_2 = 0$ in (4), the values of x_1 and x_2 provide the new estimate (X_1, X_2) for the root:

$$X_1 = a_{10} \qquad X_2 = a_{20} \tag{11}$$

where, by (8)

$$\begin{bmatrix} a_{10} & a_{20} \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & b_{13} & b_{14} \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \\ x_{41} & x_{42} \end{bmatrix} \tag{12}$$

and by (10)

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ z_{11} & z_{21} & z_{31} & z_{41} \\ z_{12} & z_{22} & z_{32} & z_{42} \\ z_{11}^2 & z_{21}^2 & z_{31}^2 & z_{41}^2 \end{bmatrix} \begin{bmatrix} b_{11} \\ b_{12} \\ b_{13} \\ b_{14} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{13}$$

which is a set of 4 linear simultaneous equations for $b_{11}, b_{12}, b_{13}, b_{14}$.

5. AN ITERATION-GENERAL.

Re-writing the example in general terms, the stages required are

(a) Having i estimates, set up the $i \times i$ matrix of polynomial terms in the function values. Let this matrix be denoted Z.

(b) Find $b_{11}, b_{12}, \ldots b_{1i}$ by solving the i linear equations

$$\begin{bmatrix} \\ Z \\ \\ \end{bmatrix} \begin{bmatrix} b_{11} \\ b_{12} \\ \vdots \\ b_{1i} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{14}$$

(from (13))

(c) Find the new estimate $(X_1, X_2, \ldots, X_n)$ from

$$\begin{bmatrix} X_1 & X_2 & \cdots & X_n \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1i} \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & & \vdots \\ x_{i1} & x_{i2} & \cdots & x_{in} \end{bmatrix} \qquad (15)$$

(from (12) and (11)).

6. SOLUTION OF THE EQUATIONS.

Suppose now that there are n equations to be solved. $(n + 1)$ initial estimates will be required, so that i is set $= n + 1$.

Applying the general method, array Z will initially consist of only constant and linear terms (corresponding to the constant and linear terms of equation (3)), and will be of size $i \times i = (n + 1) \times (n + 1)$. Using this, the new estimate $(X_1, X_2, \ldots, X_n)$ can be found.

By setting $x_{n+1,j} \equiv X_j$ ($j = 1$ to n), there are now $n + 2$ estimates available, so that i is increased by 1 to $n + 2$. Array Z may be set up, of size $i \times i = (n + 2) \times (n + 2)$ and containing quadratic terms, and a new estimate may again be obtained.

This process may be repeated until a tolerance test is met, i increasing by 1 on each occasion. The method has been used in the System 4 ALGOL program HNSA0018Z, and an example of the results obtained is given in appendix 1. This method has the disadvantage that each iteration takes longer than the preceding one.

However other alternatives are available such as increasing Z as above until it consists of all possible terms of degree not greater than 2 (corresponding to all the constant, linear and quadratic terms in (3)). Succeeding iterations would then each be followed by eliminating the 'worst' of the old estimates, which is replaced by the new one. This method would have the advantage of limiting the time taken by an iteration, but it is not known what would be the effect on the total time taken to obtain convergence.

7. RESULTS.

The method which has been adopted in the System 4 ALGOL program has given satisfactory results in several cases. However, ill-conditionality has arisen in the following respects:

(a) In later iterations, when terms of cubic or even higher degree occur in (3), the matrix Z will have elements ranging widely in numerical value. The effect may be minimised by 'scaling' the function values (multiplying by a constant), so that the initial functions have an R.M.S. value of 10^4. Iteration should reduce the 'scaled' function values to about 10^{-4}. In this way the range of the elements of Z has been restricted in most cases to $10^{\pm 20}$, which does not appear to cause any ill-effect.

(b) When initial guesses were bad, the process converged slowly, or not at all.

The method has been applied to finding more than one root to a set of equations. Once a root, say $(R_{h1}, R_{h2}, \ldots, R_{hn})$, to the equations has been found, succeeding function values z_j are 'corrected' to eliminate this old root. The only satisfactory method which has been found for 'correcting' is to divide each of the function values z_j by

$$c_h = (|X_1 - R_{h1}| + |X_2 - R_{h2}| + \dots + |X_n - R_{hn}|)$$

so that

$$z_j := z_j/c_h \qquad (j = 1 \text{ to } n)$$

The z_j are hence divided by the appropriate c_h for each old root, R_{hj}, which has been found.

8. RECOMMENDATIONS FOR FURTHER WORK.

8.1 Form of Inverse Functions.

The form of the inverse functions given by (3) has been arbitrarily taken to consist of: a constant, all linear terms, then all quadratic terms, then all cubic terms etc., in the order given. This is not the only possible form, and it may be that better convergence will be obtained by taking:

either (a) other forms of polynomials. The polynomial for two equations with four estimates has been taken as

$$x_i = a_{i0} + a_{i1} z_1 + a_{i2} z_2 + a_{i3} z_1^2$$

Another possible form is

$$x_i = a_{i0} + a_{i1} z_1^2 + a_{i2} z_1 z_2 + a_{i2} z_2^2$$

Note, however, that to take terms of a higher degree than the highest necessary appears to hinder convergence (Hence do not take

$$x_i = a_{i1} + a_{i1} z_1^3 + a_{i2} z_2^3 + a_{i3} z_1^2 z_2^2 \qquad \text{etc.)}$$

or (b) Non-linear terms, e.g.

$$x_i = a_{i0} + a_{i1} \log z_i + a_{i2} \log z_2 + a_{i3} \sin (z_1 + z_2)$$

8.2 Number of Estimates Used.

As mentioned in Section 6, the program may either retain all estimates or eliminate old estimates. Some research is needed to find which is the best alternative.

9. CONCLUSION.

The method presented provides a useful means of solving non-linear simultaneous equations, and improvements may be possible.

D. Bolton.

MATHEMATICS AND THEORETICAL STUDIES GROUP
S2 - THEORETICAL DESIGN

Approved by:

HEP

APPENDIX 1.

An example of the application of the first method described to the solution of a small number of equations with poor initial guesses.

The equations were :

$$\left.\begin{aligned} x_1^3 - 3\,x_1\,x_2^2 - 1 &= 0 \\ 3x_1^2\,x_2 - x_2^3 &= 0 \end{aligned}\right\} \qquad (1)$$

(obtained by separating real and imaginary parts of $z^3 - 1 = 0$). All three roots were found to a tolerance of 10^{-6}; the largest number of iterations taken to find a root was 8.

First Root.

		Estimates		Functions	
Initial Guesses		1.000000;	0.250000;	-0.187500;	0.734375;
		0.750000;	0.500000;	-1.140625;	0.718750;
		1.250000;	0.250000;	0.718750;	1.156250;
Successive estimates	1	1.026904;	-0.229861;	-0.079871;	-0.715043;
	2	1.049588;	-0.018608;	0.155174;	-0.061491;
	3	1.035265;	-0.002973;	-0.109544;	-0.009560;
	4	1.003679;	0.003116;	0.011048;	0.009418;
	5	0.999966;	0.000189;	-0.000101;	0.000568;
	6	0.999989;	0.000068;	-0.000033;	0.000205;
	7	1.000000;	10^{-8};	-2×10^{-9};	4×10^{-8};
Root	8	1.000000;	10^{-13};	-3×10^{-14};	3×10^{-13};

Second Root.

		Estimates		Functions	
Initial Guesses		-0.500000;	1.000000;	0.375000;	-0.250000;
		-1.000000;	0.500000;	-1.250000;	1.375000;
		-0.700000;	0.700000;	-0.314000;	0.686000;
Successive iterations	1	-0.623016;	0.924603;	0.356009;	0.286216;
	2	-0.528941;	0.863947;	0.036427;	0.080288;
	3	-0.497870;	0.865537;	-0.004464;	-0.004787;
	4	-0.499978;	0.865974;	-0.000165;	0.000020;
	5	-0.500001;	0.866027;	7×10^{-6};	7×10^{-7};
	6	-0.500000;	0.866025;	5×10^{-10};	-3×10^{-12};
Root	7	-0.500000;	0.866025;	-4×10^{-16};	1×10^{-15};

Third Root.

		Estimates		Functions	
Initial Guesses		-0.500000;	-1.000000;	0.375000;	0.250000;
		-1.000000;	-0.500000;	-1.250000;	-1.375000;
		-0.700000;	-0.700000;	-0.314000;	-0.686000;
Successive iterations	1	-0.624931;	-0.919014;	0.339367;	-0.300546;
	2	-0.518453;	-0.869227;	0.035803;	-0.044178;
	3	-0.498946;	-0.864475;	-0.005598;	0.000411;
	4	-0.499893;	-0.865957;	-0.000338;	-0.000175;
	5	-0.499999;	-0.866025;	-2×10^{-7};	2×10^{-7};
Root	6	-0.500000;	-0.866025;	-10^{-10};	-10^{-11};

Bolton's Method, as applied to a single equation:

$$x_1 = a_{10} + a_{11}z_1 + a_{13}z_1^2 + \ldots$$

3 estimates x_{11}, x_{21}, x_{31}. For values z_{11}, z_{21}, z_{31}

Then

$$\begin{bmatrix} 1 & 1 & 1 \\ z_{11} & z_{21} & z_{31} \\ z_{11}^2 & z_{21}^2 & z_{31}^2 \end{bmatrix} \times \begin{bmatrix} b_{11} \\ b_{12} \\ b_{13} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Solution gives:- (b_{11}, b_{12}, b_{13}).

Then new estimate, $x_{41} = b_{11}x_{11} + b_{12}x_{21} + b_{13}x_{31}$

Let $z_1 = (x_1^2+1)(x_1-2)^2(x_1+3)$

For $x_{11} = 0$, $z_{11} = 12.0$, & for $x_{21} = 0.5$, $z_{21} = 9.84375$.

$$\therefore \begin{bmatrix} 1 & 1 \\ 12.0 & 9.84375 \end{bmatrix} \times \begin{bmatrix} b_{11} \\ b_{12} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\left.\begin{aligned} (12.0 - 9.84375)\, b_{12} &= +12.0 \\ \& \ (12.0 - 9.84375)\, b_{11} &= -9.84375 \end{aligned}\right\}$$

$$\therefore x_{31} = \frac{12.0 \times 0.5 - 9.84375 \times 0}{12.0 - 9.84375} \quad \text{(linear extrapolation)}$$

$$= 2.7826087 \qquad z_{31} = 30.9648663$$

$$\begin{bmatrix} 1 & 1 & 1 \\ 12.0 & 9.84375 & 30.9648663 \\ 144.0 & 96.8994141 & 958.822945 \end{bmatrix} \times \begin{bmatrix} b_{11} \\ b_{12} \\ b_{13} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \qquad 21.84375 ✓$$

$$\therefore \begin{bmatrix} -2.15625 & 18.9648663 \\ -47.1005859 & 814.822945 \end{bmatrix} \times \begin{bmatrix} b_{12} \\ b_{13} \end{bmatrix} = \begin{bmatrix} -12.0 \\ -144.0 \end{bmatrix} ✓$$

$$\therefore \overset{400.559147}{\cancel{400.631180}}\, b_{13} = \overset{118.125}{\cancel{118.079420}}$$

$$\left.\begin{aligned} \therefore b_{13} &= \overset{0.2949003}{\cancel{0.2947335}} \\ \therefore b_{12} &= 8.1589538 \\ \& \ b_{11} &= -7.4538541 \end{aligned}\right\} \begin{aligned} &\text{eqn. 1 checks ✓} \\ &\text{" 2 " ✓} \\ &\text{" 3 does not check} \end{aligned}$$

$\therefore x_{41} = 4.9000690$ (as with "Aitken")