Efficiency of a modified globally convergent algorithm for solving systems of nonlinear equations

1987 J. Phys. A: Math. Gen. 20 3219

(http://iopscience.iop.org/0305-4470/20/11/025)

J. Phys. A: Math. Gen. **20** (1987) 3219-3224. Printed in the UK

Efficiency of a modified globally convergent algorithm for solving systems of non-linear equations

W C Kok and S M Tang

Department of Physics, National University of Singapore, Kent Ridge, Singapore 0511

Received 9 October 1986, in final form 16 January 1987

Abstract. A study is made of the efficiency of an iterative procedure devised for solving a system of non-linear equations $\boldsymbol{F}(\boldsymbol{x}) = 0$, in which the choice of an iterative scheme randomly selected from a set of schemes is generally based on satisfying the criterion of a smaller norm $\|\boldsymbol{F}\|$ in the subsequent iteration. The practicality of using such a procedure is illustrated by solving a set of $n = 5$ equations.

1. Introduction

Many different methods (Ostrowski 1960, Ortega and Rheinboldt 1970, Blue 1980, Allgower and Georg 1980, Zirilli 1982, Dennis and Schnabel 1983) have been proposed for solving algebraic systems of equations. While those for solving linear systems are well established, the problem of solving non-linear systems is far less tractable. Many iterative methods are convergent only when the starting vector is close to a zero of the equations to be solved. For problems where the solutions are not known with any degree of accuracy, methods which extend the domain of convergence would be desirable and much effort has been expended in the past in this direction. In our recent paper (Tang and Kok 1985, hereafter referred to as TK), we presented a computational procedure which has the advantage of global convergence for solving non-linear algebraic systems of the general form

$$F_j(x_1, x_2, \ldots, x_n) = 0 \qquad j = 1, 2, \ldots, n. \tag{1}$$

The iteration equations employed in this procedure are

$$x_1^{i+1} = x_{1'}^{i+1} \tag{2a}$$

$$x_j^{i+1} = \tfrac{1}{2}(x_j^i + x_{j'}^{i+1}) \qquad j = 2, 3, \ldots, n \tag{2b}$$

where

$$x_{j'}^{i+1} = x_j^i - F_{j'}^i / F_{j',j}^i \tag{3}$$

with

$$F_{j'}^i = F_j(x_{1'}^{i+1}, x_{2'}^{i+1}, \ldots, x_{(j-1)'}^{i+1}, x_j^i, \ldots, x_n^i). \tag{4}$$

It can be seen from the structure of these formulae that there are $(n!)^2$ possible iteration schemes, each corresponding to a different iteration sequence of x_j associated with a different ordering of F_j. Each of these schemes will lead to a distinct iteration path.

0305-4470/87/113219+06$02.50

With an arbitrarily chosen initial vector x^0, the procedure starts with the search of the scheme that yields the smallest norm $\|F\|$ at x^1. Once found, the scheme is adopted for subsequent iterations as long as it produces a norm which is smaller than the one at the previous iteration point. At the occurrence of increasing $\|F\|$, the search for the scheme that gives the smallest norm of F is repeated using the iterates of the previous stage as the starting vector.

This algorithm has been extensively tested for two- and three-variable systems and found to be reliable and reasonably efficient. For systems with $n > 3$, however, its efficiency is poor since each time a search for the scheme that yields the smallest norm of F is required, the computation of $\|F\|$ has to be repeated $(n!)^2$ times. It can be proved that the $(n!)^2$ iteration schemes of a n-variable system have only $n!(n-1)!$ distinct convergence factors. In other words, each convergence factor is associated with n schemes. Therefore the search for the scheme that gives the smallest $\|F\|$ at the next point may be restricted to those $n!(n-1)!$ having distinct convergence factors. But $n!(n-1)!$ is still a large number for $n > 3$. In this paper, we discuss a modification of the method to improve the efficiency for large n and this is illustrated by solving a non-linear system of $n = 5$ with several different starting points.

2. The criterion for scheme selection

Basically, the algorithm presented in TK consists of two stages to finding a solution given an arbitrary starting vector. The first stage involves the application of an iterative process until the iterates approach the neighbourhood of the solution where a convergence factor may be defined. The convergence to the solution in this stage is governed by the criterion of decreasing norm. Changes of iterative scheme are often required in order to satisfy this criterion. In the second stage, the same iterative process is employed. But once a locally convergent scheme has been adopted, the criterion of decreasing norm is automatically satisfied in all subsequent iterations and there will not be any further change of iteration scheme. In TK we proposed that the iteration scheme giving the smallest $\|F\|$ at the next iteration point should be identified and adopted for subsequent iterations at the occurrence of increasing norm. The reason is that this scheme will give the fastest convergence within the linear region of the solution. As pointed out earlier, this scheme selection algorithm requires a lot of computing time when n is large. To reduce the time needed for scheme selection, we propose to relax the requirement of the choice of scheme based on the minimum $\|F\|$ and substitute it with the less stringent requirement that, whenever a choice of schemes is to be made, the one chosen needs only to yield a smaller $\|F\|$ at the next iteration point. The search for such a scheme may be done randomly. Although this might increase the number of iterations for reaching the solution, the computing time saved in searching for a new scheme is expected to be substantial.

While it is almost an impossible task to show that such a choice of scheme would render the method more efficient generally, a consideration of the probability for finding a scheme that gives a smaller $\|F\|$ at the next iteration point would throw some light on this. Within the linear region near a root, an iteration scheme having a convergence factor less than one will always yield a smaller $\|F\|$ at the next iteration point. In TK, it was shown rigorously that at least two out of the four iteration schemes for a two-variable system had convergence factors less than one in the linear region near a root. The probability for finding an acceptable scheme in this region is calculated

to be 0.65. For three-variable cases, the number of convergent schemes is system dependent. We showed in TK that out of $36(=3!\times 3!)$ possible iterative schemes, the number of convergent schemes might vary from 6 to 27. The result was obtained from computing the convergence factors from 10 000 sets of nine randomly generated numbers representing the nine partial derivatives at a root of a three-variable system. The corresponding probability for finding a convergent scheme is 0.33. Similar calculations for $n>3$ can be carried out but require much computer time. It involves either solving the secular equation shown in the appendix for the largest root or determining whether all its roots lie within the unit hyper-circle with centre at the origin. Several approaches (Hammarling 1970) can be used but we employed Schur-Cohn's criterion (Marden 1949) in our computation to determine whether a scheme is convergent. The probabilities obtained are 0.14 and 0.05 for $n=4$ and $n=5$, respectively. The results of our computations are detailed in table 1 for easy reference. For $n=5$ the chance of obtaining a convergent scheme is only 1 out of 20 attempts on the average. Thus this criterion of scheme selection gives a mean time-reduction factor of 144 (i.e. 2880/20).

Table 1.

n	2	3	4	5
Number of possible iteration schemes	4	36	576	14 400
Number of distinct convergence factors	2	12	144	2 880
Number of convergent schemes	2-4	6-27	24-180	290-1375
Average number of convergent schemes	2.6	12	81	706
Probability for finding a convergent scheme	0.65	0.33	0.14	0.05
Number of systems considered	—	10 000	5000	200

In the course of the above study, we also examined the probability of not finding any convergent scheme in a subset of $n!$ schemes obtained from changing the order of F_j but keeping the iteration sequence of the variables x_j fixed. It was found to be zero for $n=3$, 0.000 03 for $n=4$ and 0.001 for $n=5$. These results imply that it is most likely that the iterative process would lead to finding a solution even if the search of schemes is restricted to one of these subsets. Adopting this restriction in the algorithm mentioned in TK will reduce the scheme searching time by a factor of $n!$. However, confining the selection of the iteration schemes within this subset will not improve the efficiency when the above proposed criterion for choosing the iteration scheme is employed.

3. An example

As an illustration of the practicality of applying the algorithm to solving a system of equations with $n>3$, we consider the example of an isothermal irreversible second-order constant volume reaction $A+B\rightarrow C+D$. A solution containing A and B is fed into n continuous stirred tank reactors in series with a volumetric flow rate v l min^{-1}. The volume of each tank reactor is V l. If the reaction has a velocity constant k l g^{-1} mol^{-1} min^{-1} and the inlet concentrations of A and B are both equal to a_0 g mol l^{-1},

then the exit concentration, a_j, of A from the jth reactor satisfies the equation (Carnahan and Wilkes 1973):

$$f_j = (Vk/v)a_j^2 + a_j - a_{j-1} = 0. \tag{5}$$

For given values of a_0, a_n, k and v, the intermediate concentrations $a_1, a_2, \ldots, a_{n-1}$ and the tank volume V can be obtained from solving the n simultaneous non-linear equations defined by (5). For our illustration, we took $n = 5$. Since the algorithm requires all the first partial derivatives of f_j to be non-zero, we solved, instead, the following equations to obtain the intermediate concentrations and Vk/v:

$$\begin{aligned} F_1 &= f_1 + f_2 + f_3 + f_4 + f_5 = 0 \\ F_2 &= f_1 - f_2 + f_3 + f_4 + f_5 = 0 \\ F_3 &= f_1 + f_2 - f_3 + f_4 + f_5 = 0 \\ F_4 &= f_1 + f_2 + f_3 - f_4 + f_5 = 0 \\ F_5 &= f_1 + f_2 + f_3 + f_4 - f_5 = 0. \end{aligned} \tag{6}$$

With four different initial vectors $(a_1, a_2, a_3, a_4, Vk/v)$ and taking $a_0 = 2.0$ and $a_5 = 1.0$, we obtained the solution (1.677, 1.439, 1.258, 1.115, 0.1148) using a Sinclair microcomputer and interpretive BASIC. With each initial vector, the computation was repeated 40 times and each time a different sequence of random numbers was used for the scheme selection process. The average as well as the range of computer time and the average number of iterations required to obtain the solution accurate to four significant figures for each case are shown in table 2. In the same table, the amount of computer time and number of iterations for obtaining the solution with the same accuracy using the algorithm given in TK are also included for comparison. In all these computations, the selection of iterative schemes was restricted to one of the 5! subsets for the reason stated in the previous section. The frequency distribution of the time taken for the runs using the same initial vector is shown in figure 1. These results show that relaxing the scheme selection requirement does make the procedure more efficient. On the average, the computer time required to obtain the solution is reduced by a factor of 3. It is interesting to note that the number of iterations taken to find the solution is almost the same for both algorithms. This seems to indicate that outside the linear region the selection of the scheme that gives the smallest $\|F\|$ does not necessarily

Table 2.

	Modified algorithm		Algorithm of TK	
Initial vector	Average time (range) (s)	Average number of iterations	Time (s)	Number of iterations
(500, 100, 300, −400, −200)	1006 (396-2490)	60	2899	54
(−300, 100, 500, −200, 400)	1154 (523-2591)	61	6870	62
(100, 200, 300, 400, 500)	1042 (417-4717)	61	1388	49
(100, −400, 200, −500, −300)	1068 (400-2434)	63	1420	63

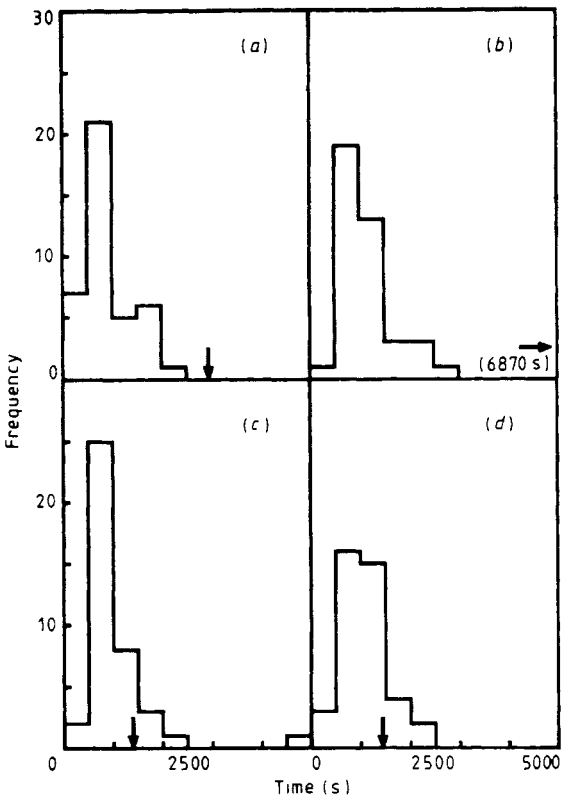

Figure 1. Frequency distribution of the time taken for convergence to the solution for different starting vectors. ((*a*) (500, 100, 300, −400, −200); (*b*) (−300, 100, 500, −200, 400); (*c*) (100, 200, 300, 400, 500); (*d*) (100, −400, 200, −500, −300)). The arrow indicates the time required using the algorithm of TK.

lead to a shorter path towards a solution. In the computations using the modified algorithm, we also monitored the average numbers of schemes tested before an acceptable one was encountered. The result found was 40. This is twice the expected average value obtained in § 2.

4. Remarks

The approach presented here and in TK differs entirely from that of modified Newton methods (Dennis and Schnabel 1983). In the latter techniques, a subsequent iterate is obtained with an adjustable step size along a descent direction of a model function. On the other hand, our algorithm involves mainly a search for a descent path from a finite set of iterative schemes.

The existence of a local minimum near an iterate poses a problem for all descent methods as it tends to restrict the domain of global convergence. This problem can be overcome in our method as described in TK.

The modified Newton methods have the advantage that they are quadratically convergent whereas our algorithms exhibit only linear convergence. However, each

basic iterative step of our algorithm is simple; it only requires evaluating n times a variation of the one-dimensional Newton-Raphson formula.

Acknowledgment

This work is supported by a grant from the National University of Singapore.

Appendix

The elements of the $(n-1)\times(n-1)$ matrix S in the secular equation $\det(S-\alpha I)=0$ (TK, equation (4.5)) can be calculated from the formula

$$M_{j,k}=\tfrac{1}{2}(M'_{j,k}+\delta_{j,k}) \qquad j,k=2,3,\ldots,n$$

where

$$M'_{j,k}=-\frac{1}{F_{j,j}}\left(\sum_{l=1}^{j-1}F_{j,l}M'_{l,k}+F_{j,k}s_{j<k}\right)$$

with

$$s_{j<k}=\begin{cases}1 & j<k\\ 0 & j\geqslant k\end{cases}$$

$$M'_{1,k}=-F_{1,k}/F_{1,1}.$$

The secular equation has $(n-1)$ roots and the largest one is the convergence factor. Schur-Cohn's criterion allows the determination of whether the absolute values of all the roots are less than one without solving the secular equation.

References

Allgower E and Georg K 1980 *SIAM Rev.* **22** 28

Blue J L 1980 *SIAM J. Sci. Stat. Comput.* **1** 22

Carnahan B and Wilkes J O 1973 *Digital Computing and Numerical Methods* (New York: Wiley)

Dennis J E Jr and Schnabel R B 1983 *Numerical Methods for Unconstrained Optimization and Nonlinear Equations* (Englewood Cliffs, NJ: Prentice-Hall)

Hammarling S J 1970 *Latent Roots and Latent Vectors* (Bristol: Adam Hilger)

Marden M 1949 *The Geometry of the Zeros of a Polynomial in a Complex Variable* (New York: Am. Math. Soc.) p 152

Ortega J M and Rheinboldt W C 1970 *Iterative Solution of Nonlinear Equations in Several Variables* (New York: Academic)

Ostrowski A 1960 *Solution of Equations and Systems of Equations* (New York: Academic)

Tang S M and Kok W C 1985 *J. Phys. A: Math. Gen.* **18** 2691

Zirilli F 1982 *SIAM J. Numer. Anal.* **19** 800