Internat. J. Math. & Math. Sci.
Vol. 7 No. 4 (1984) 793-802



# MULTIPARAMETER EXTRAPOLATION AND DEFLATION METHODS FOR SOLVING EQUATION SYSTEMS

**A.J. HUGHES HALLETT**

Department of Economics
Erasmus University
PO Box 1738, 3000 DR Rotterdam
The Netherlands

(Received May 21, 1984 and in revised form August 9, 1984)

ABSTRACT. Most models in economics and the applied sciences are solved by first order iterative techniques, usually those based on the Gauss-Seidel algorithm. This paper examines the convergence of multiparameter extrapolations (accelerations) of first order iterations, as an improved approximation to the Newton method for solving arbitrary nonlinear equation systems. It generalises my earlier results on single parameter extrapolations. Richardson's generalised method and the deflation method for detecting successive solutions in nonlinear equation systems are also presented as multiparameter extrapolations of first order iterations. New convergence results are obtained for those methods.

KEY WORDS AND PHRASES. *First order iterative techniques, extrapolation methods, approximation theory, multiple solutions.*
*AMS MATHEMATICS SUBJECT CLASSIFICATION CODES.* 65F10

## 1. INTRODUCTION.

Solving systems of equations is one of the most commonly performed operations of applied science. Every forecast and every simulation made with the aid of a numerical model will be based on a sequence of solutions to the corresponding set of equations. Moreover many optimisation techniques, such as maximum likelihood estimation or stochastic nonlinear optimal control, involve repeated equation solutions. It is therefore necessary to be able to solve an arbitrary equation system efficiently. Indeed the British Treasury, for example, recently complained that inefficient solution methods had seriously inhibited them in their search for better policies for the British Economy (HMSO [5]).

Iterative techniques, and the Gauss-Seidel method in particular, have been used most frequently by applied scientists and economists for solving both nonlinear and large linear systems. It is obviously important to know which methods can guarantee convergence in an arbitrary equation system, and how accelerations can be introduced so as to minimise computational costs. General convergence results and optimal

extrapolations first appeared in the econometrics literature (Hughes Hallett [6]). Subsequent results have stressed that the usual first order iterations with one parameter extrapolations necessarily diverge in a large class of equation systems (see Hughes Hallett [7], Hadjidimos [3], Hughes Hallett [9], [10]). In common with many other areas of the applied science, econometric models can contain no restrictions which would normally exclude them from that class.

The aim of this paper is to design first order iterative techniques which are guaranteed to converge in an arbitrary equation system. This is done by assigning separate extrapolation parameters to each equation. The resulting techniques generalise on Richardson's method, and on the deflation method for detecting successive solutions to nonlinear equation systems. Thus convergence results for the latter two methods are derived together with those for multiparameter extrapolation techniques.

## 2. STATIONARY INTERACTIONS AND ONE PARAMETER EXTRAPOLATIONS.

Consider the linear equation system

$$Ay = b \tag{2.1}$$

where $A \in \mathbf{R}^{n,n}$ is a known real matrix of order n with nonvanishing diagonal elements, and where y and b respectively are real vectors of unknown and predetermined variables. Stationary first order iterative techniques,

$$y^{(s+1)} = G\, y^{(s)} + c \qquad s = 0,1,2,\ldots \tag{2.2}$$

with an arbitrary start $y^{(o)}$, are widely used to construct the numerical solution to (2.1); see for example Hageman and Young [4] or Varga [14]. They are computationally efficient if A is a large, sparse, or ill-conditioned matrix. They are also widely used for solving nonlinear equation systems, in which case A represents the system's Jacobian matrix.

These first order methods are based on the splitting $A = P - Q$ where P is nonsingular; and they are completely consistent with (2.1) when $G=P^{-1}Q$ and $c=P^{-1}b$ define the iteration matrix and forcing function (Young [16]). The best known examples of (2.2) are the Jacobi, Gauss-Seidel, and Successive Overrelaxation (SOR) methods defined by

$$P = D, \text{ and } P = (D-L), \text{ and } P = \frac{1}{\alpha} D\,(I-\alpha D^{-1}L) \tag{2.3}$$

respectively, where D and L are matrices of order n such that

$$D_{ij} = \begin{cases} A_{ij} & \text{if } i=j \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad L_{ij} = \begin{cases} -A_{ij} & i<j, \\ 0 & \text{otherwise} \end{cases} \quad \text{and } \alpha \text{ is a}$$

scalar (Varga [14]). Without loss of generality, let (2.1) contain the normalisation $D = I$ and let $B = I - A$.

The rate of convergence in (2.2) is often significantly increased by the one parameter extrapolation

$$y^{(s+1)} = \gamma(G\, y^{(s)} + c) + (1-\gamma)y^{(s)} = H\, y^{(s)} + \gamma c \ . \qquad (2.4)$$

Two particular versions of (2.4) are routinely used; the Jacobi overrelaxation (JOR) method with G=B, and the Fast Gauss-Seidel (FGS) method based on $G = (I-\alpha L)^{-1}(\alpha U+(1-\alpha)I)$. Both techniques are specialisations of the standard Newton method for solving systems of equations (Hughes Hallett [7]).

It is well known that (2.2) converges if and only if $\rho(G)<1$, where $\rho(.)$ denotes spectral radius (Young [16], theorem 5.1). Similarly

LEMMA 1: (2.4)converges from some $\gamma>0$ if and only if $a_j<1$, j=1...n, where G has eigenvalues $\lambda_j=a_j+ib_j$ and $i = \sqrt{-1}$.

PROOF: Hughes Hallett [6].

The latter result implies(2.4)also converges for some $\gamma<0$ if $a_j>1$ all j; but that it is necessarily divergent if $a_j<1<a_k$ for some $j\neq k$. Hence it is not possible to achieve convergence using (2.4) in a wide class of equation systems. A simple iteration guaranteed to converge for arbitrary values of G is therefore required.

Finally, if $\rho(H)<1$, the number of steps for convergence such that

$$\max_i \left|(y_i^{(s)} - y_i^{(s-1)})/y_i^{(s-1)}\right| < \tau \qquad (2.5)$$

is approximately $\log \tau/\log \rho(H)$. That criterion is equivalent to defining convergence in $\|y^{(s)} - y^{(s-1)}\|_\infty$. The corresponding speed to convergence can then be measured as an asymptotic rate ($-\log \rho(H)$) or as an average rate ($\frac{1}{s}\log \|H^s\|$, for some norm). We can therefore aim to minimise $\rho(H)$ in order to ensure convergence.

## 3. MULTIPARAMETER FIRST ORDER EXTRAPOLATIONS.

There are two ways forward when faced with the inevitable divergence of one parameter first order extrapolations in an equation system where G has roots spanning unity in real parts. One possibility is to try multiple extrapolation parameters. The other possibility is to use second order iterations, with extrapolations as required. The question then arises, do both approaches guarantee convergence for arbitrary values of A, G and $y^{(o)}$? If so, which method converges fastest and at the lower computational cost?

Although these two approaches may be combined, they are presented here as dichotomous options for three reasons:

(i) Second order iterations are now known to offer guaranteed convergence for any value of G, and also to involve relatively economical calculations (Hughes Hallett [8]). The question is whether multiparameter first order extrapolations have the same properties.

(ii) The multiparameter first order extrapolation method is easily shown to include Richardson's generalised method for solving equation systems (the GRF method in Young [16]) as a special case. No convergence results have been presented for Richardson's method, except for the particular case where A is either positive definite or symmetric (Young [16]). Indeed the method appears to have been neglected in applications, probably for lack of a general convergence theory since it is no

more difficult to use than (2.4). A general convergence theory for this class of iterations is therefore required.

(iii) In an attempt to design an algorithm to locate all the real solutions of a nonlinear equation system, Salmon [13] has proposed a multiparameter first order nonlinear iteration due to Brown and Gearhart [1]; the deflation method. However that algorithm can only be certain to locate those solutions if the iterations are convergent for arbitrary values of G, since the underlying Jacobian matrix is then strictly - to the extent of sign changes - path dependent. Once again a general convergence theory is needed.

3.1 STATIONARY MULTIPARAMETER EXTRAPOLATIONS.

The multiparameter extrapolation of the general first order stationary iteration, (2.2), is

$$y^{(s+1)} = R\,G\,y^{(s)} + (I-R)y^{(s)} + R\,c \qquad (3.1)$$

where R is a real nonsingular diagonal matrix. This iteration is a generalisation of (2.4) in which each equation can be assigned a separate extrapolation parameter. But if $R = \gamma I$, then (3.1) specialises to (2.4). Of course any G which is completely consistent with (2.1) may be used in (3.1); and the Jacobi, Gauss-Seidel and SOR iteration matrices defined by the values of P in (2.3) will be typical examples. The iteration matrix of (3.1) is I+R(G-I), so the multiparameter extrapolation is convergent if and only if $\rho(I+R(G-I))<1$.

3.2 RICHARDSON'S ITERATIONS.

Richardson's generalised method, applied to (2.1), is

$$y^{(s+1)} = y^{(s)} + R^*\,(Ay^{(s)}-b) \qquad (3.2)$$

where R* is a real diagonal matrix. Each iterate is therefore constructed from the previous step plus a proportion of the previous step's error $Ay^{(s)}-b$. Evidently this method is identical to a multiparameter extrapolation of the Jacobi method when $R=-R^*$, since (3.2) has iteration matrix $I+R^*A = -R^*B + (I+R^*)$ and forcing function $-R^*b$. Similarly the multiparameter Jacobi extrapolation, (3.1), has iteration matrix RB+(I-R) and forcing function Rb. Moreover if $R^*=rI$, yielding Richardson's ordinary method, then (3.2) specialises to the JOR method given by (2.4) with $r=-\gamma$. Convergence conditions and optimal values of r, for this method and for (2.4), have been given by Hughes Hallett [9].

Likewise Gauss-Seidel and SOR based versions of Richardson's generalised method are also possible:

$$y^{(s+1)} = y^{(s)} + \{R^*\,(1-G)\,y^{(s)} - c\} \qquad (3.3)$$

where $G = (I-\alpha L)^{-1}\,(\alpha U + (1-\alpha)I)$ and $c = \alpha^{-1}(I-\alpha L)$ for suitable $\alpha$ values, and R* is a real nonsingular diagonal matrix. This obviously coincides with the multiparameter SOR extrapolation defined by (3.1). Indeed, a Richardson method can be constructed

from any first order iteration described by (2.2) which is completely consistent with (2.1) since the same splittings are used. For that reason, the Richardson method is best interpreted as a multiparameter extrapolation of any underlying first order iteration represented by (2.2).

3.3 THE DEFLATION METHOD'S ITERATIONS.

Salmon [13] examined the deflation method for detecting the real solutions of nonlinear equation systems of the form

$$y = f(y,b) \ . \tag{3.4}$$

A convergent first order iterative technique

$$y^{(s+1)} = g(y^{(s)}, b), \tag{3.5}$$

which can be the nonlinear counterpart to any of the iterations described by (2.2), is adopted to locate any solution to (3.4). Let that first solution be $y_{(1)}$. Then (3.5) is restarted in the form

$$y^{(s+1)} = R\, g(y^{(s)},b) + (I-R)\, y^{(s)} \tag{3.6}$$

where $R = rI$, with $r^{-1} = \|y^{(s)} - y_{(1)}\|_\rho$ for a suitable choice of norm. That extrapolation ensures that another real solution $y_{(2)} \neq y_{(1)}$, if one exists, will be found whenever the method converges. Then (3.6) is repeated, where R is now a diagonal matrix with elements $R_{ii} = (\|y^{(s)} - y_{(1)}\|_\rho)^{-1}$ but $R_{jj} = (\|y^{(s)} - y_{(2)}\|_\rho)^{-1}$ for at least one $j \neq i$. Ultimately, (3.6) will be run n-1 times with R specified to be a diagonal matrix having at least one element equal to the inverse of the chosen norm applied to each solution vector so far detected. The alternative version with $R_{ii} = |y_i^{(s)} - y_{(k)i}|^{-1}$, where $y_i^{(s)}$ and $y_{(k)i}$ are respectively the $i^{th}$ elements of $y^{(s)}$ and the $k^{th}$ solution $y_{(k)}$, for i=1...n, must be ruled out because it cannot prevent a previous solution being "rediscovered" when k>2.

There are two points to note about Salmon's iterations:

(a) They are just multiparameter extrapolations (or Richardson's generalised method) applied to the underlying nonlinear iteration (3.5). They therefore form the nonlinear counterpart to (3.1), and may be analysed for convergence accordingly.

(b) The success of the method depends entirely on first picking an underlying iterative method (3.5) which converges; and then picking values for R which allow (3.6) to converge at each of the n-1 repetitions. However it is not proven that any values of R do in fact exist to guarantee convergence for an arbitrary iteration matrix, $G = [\partial g/\partial y]$, associated with each step $y^{(s)}$ in (3.6). Nor is it proven that the particular values proposed for R would necessarily imply convergence. In fact, divergence appeared in numerical tests.

4. THE CONVERGENCE RESULTS.

4.1 DIAGONAL EXTRAPOLATION MATRICES.

THEOREM 1: A real diagonal matrix R does not always exist such that the multiparameter extrapolations (3.1), and Richardson's generalised method (3.3), are guaranteed to converge for an arbitrary real matrix G.

PROOF: Iterations (3.1) and (3.3) are identical when $R=-R^*$. Thus they are convergent iterations, given arbitrary values of G and $y^{(o)}$, if and only if a real diagonal R matrix exists such that R(G-I) has all its eigenvalues in the left half-plane. That follows because the iteration matrix is I + R(G-I). Hence, by lemma 1, $\gamma>0$ exists such that the additional extrapolation

$$y^{(s+1)} = \{\gamma[R(G-I) + I] + (1-\gamma)I\}\, y^{(s)} + \gamma\, Rc \qquad (4.1)$$

is convergent if and only if the roots of R(G-I) lie in the left half-plane. But (4.1) just reproduces (3.1) with $\gamma R$ for R. (Similarly if R exists such that the roots of R(G-I) all lie in the right half-plane, then substituting -R for R provides roots with negative real parts). The Lienard-Chipart stability criterion (Gantmacher [2]) states that the polynomial

$$f(\lambda) = a_0\lambda^n + a_1\lambda^{n-1} + \ldots + a_n = 0 \qquad \text{with} \quad a_0 = 1 \qquad (4.2)$$

has all roots in the left half-plane if and only if

$$a_n > 0,\ a_{n-1} > 0,\ a_{n-3} > 0 \ \ldots \ \underline{\text{and}}\ \Delta_1 > 0,\ \Delta_3 > 0,\ldots \qquad (4.3)$$

or

$$a_n > 0,\ a_{n-2} > 0, \qquad \ldots \ \underline{\text{and}}\ \Delta_2 > 0,\ \Delta_4 > 0,\ldots \qquad (4.4)$$

where $\Delta_1 = a_1$, $\Delta_i = \begin{vmatrix} a_1 & a_3 & a_5 & \cdots \\ a_0 & \ddots & \ddots & \\ & \ddots & \ddots & a_{i-1} \\ 0 & & a_{i-1} & a_i \end{vmatrix}$ and $\Delta_n = \begin{vmatrix} a_1 & a_3 & \cdots & 0 \\ a_0 & \ddots & \ddots & \\ & \ddots & \ddots & 0 \\ 0 & & a_{n-1} & a_n \end{vmatrix}$

and $|\cdot|$ denotes a determinant. (4.3) and (4.4) are implied by the Routh-Hurwicz stability criterion, $\Delta_i > 0$ for i=1...n. But $a_n = (-1)^n|R||G-I|$ and

$$a_k = \sum_{j=k}^{n} (r_{j-k+1} \cdots r_j) d_j^{(k)} \qquad k=1\ldots n-1 \qquad (4.5)$$

where $d_j^{(k)}$ is the $j^{th}$ principle minor of G-I, with leading element $a_{j-k+1,j-k+1}$, of order k. The theorem now follows from two counterexamples where (4.3) and (4.4) are violated for <u>every</u> value of R; namely (i) where G-I is singular, and/or (ii) where $d_j^{(k)} = 0$ for all j=k...n and some $2 \leq k \leq n$. Notice that case (ii) is entirely consistent with nonsingular G-I matrices.

COMMENTS: (a) Theorem 1 demonstrates that the convergence of a first order iteration cannot be guaranteed by choice of a real diagonal matrix R, given an arbitrary matrix A. An alternative interpretation of theorem 1 is that, given an arbitrary matrix G, it is not possible to assign the eigenvalues of I + R(G-I) to an

arbitrary subspace of the complex field (to the left half-plane in this instance) by choice of a diagonal R.

(b) Theorem 1 is consistent with the standard result that convergent single parameter extrapolations of the Jacobi and Gauss-Seidel methods exist when either A or I-G are positive definite with unit diagonal elements (Young [16]), because $d_1^{(k)} > 0$ for all k and G-I = -A when G=B.

(c) Theorem 1 is not reworked to establish the conditions on A which ensure that a real diagonal R matrix can be found to make (3.1) convergent because: (i) such conditions would still not indicate what should be done in the unavoidably divergent cases; and (ii) convergence is already guaranteed by a computationally simple second order method, so that (3.1) is not generally competitive with second order methods.

(d) The case where G-I (or equivalently A, see theorem 2 below) is singular is not uninteresting. Equations where A is ill-conditioned, or where there are multiple (indeterminate) solutions, are exactly the situations where iterative techniques are essential even in linear systems. In nonlinear systems the iteration matrix is path dependent, and convergence should be preserved even if $G^{(s)}$-I passes through a singular value. However theorem 1 also applies to nonsingular G-I matrices, and an equality in one of the conditions in (4.3) or (4.4) generally implies others in the sequence are violated. For example, suppose $a_{n-1} = 0$ but $\{ G-I , a_n \} \neq 0$. The matrix in $\Delta_n$ then implies

$$\Delta_n = - a_{n-3} a_n^2 \Delta_{n-3} \quad . \tag{4.6}$$

The sign of $\Delta_n$ is therefore determined by $-a_{n-3} \Delta_{n-3}$. Using the fact that $x_{ii} x^{ii} \geq 1$ where $x_{ii}$ and $x^{ii}$ are corresponding diagonal elements from X and $X^{-1}$ respectively (Rao [12]), we have $a_1 a_{n-3} / \Delta_{n-3} \geq 1$. So if $a_1 > 0$ is chosen to satisfy the Routh-Hurwicz criterion, then $\Delta_n < 0$.

4.2 NONDIAGONAL EXTRAPOLATION MATRICES.

DEFINITION: Let K and L be n x n and n x m matrices respectively. They are said to be a controllable pair if and only if some $t \geq n \geq 2$ exists such that $r(L, KL, K^2L, \ldots, K^{t-1}L) = n$, where r(.) denotes rank (Wonham [15]).

LEMMA 2: (I,G-I) form a controllable pair if and only if G-I is nonsingular.

PROOF: $r(eR(G-I)) = n$ if $r(G-I) = n$ where $e=(1,\ldots,1)$ is of length t.

THEOREM 2: Let A be nonsingular. Then a real matrix R exists such that the eigenvalues of $I' = R(G-I)$ assume an arbitrary symmetric set of complex numbers, where G is the iteration matrix of any first order method, (2.2), which is completely consistent with (2.1).

PROOF: By definition A = P - Q = P(I-G), where P is nonsingular, when G is completely consistent with A. Hence the nonsingularity of A implies that of G-I, and vice versa. Consequently (I,G-I) form a controllable pair. Moreover the nonzero eigenvalues of I + R(G-I) are identical to those of I + (G-I)R (Young [16], theorem 2.1, p. 48). R always exists such that I + (G-I)R has eigenvalues which form any arbitrary symmetric set of complex numbers if (I,G-I) is a controllable pair. That completes the proof.

COMMENTS: Theorem 2 implies that, given an arbitrary nonsingular A, a real nondiagonal R may be chosen to make $\rho(I + R(G-I))$ arbitrarily small, and hence to guarantee the convergence of (3.1). But it does not specify how such an R should be determined. The obvious procedure of setting $R = (D-I)(G-I)^{-1}$, where D is a nilpotent matrix, suggests that finding such an R matrix will involve no less computation than solving (2.1) by matrix inversion. This method would not be used in practice - at least not for linear equation systems.

4.3 NONLINEAR ITERATIONS.

Now consider iterative techniques designed to solve the nonlinear equation system (3.4); e.g. (3.6) where R may be any real diagonal matrix. As usual, the convergence of nonlinear iterative processes must be based on a linearisation of the system in the neighbourhood of some solution y*.

THEOREM 3: There exists a continuous neighbourhood about y*, with an interior point for $y^{(o)}$, such that the nonlinear iteration (3.5) is convergent to y* only if $\rho(G^*) < 1$ where $G^* = [\partial g/\partial y]$ at $y = y^*$.

PROOF: Ostrowski [11], Appendix K.

In the nonlinear case, theorems 1 and 2 can be reformulated as:

THEOREM 4: (i) A real diagonal matrix R does not always exist to make (3.6) convergent for an arbitrary nonlinear function g(.,.) and solution y*; (ii) A real nondiagonal matrix R exists such that (3.6) is convergent in a neighbourhood of y*, for an arbitrary g(.,.) and y*. Moreover a sequence of matrices, $R^{(s)}$, exists such that (3.6) is convergent from any $y^{(o)}$ for an arbitrary function g(.,.).

PROOF: (i) By theorem 1, a real diagonal matrix R does not exist to ensure that $\rho(I + R(G^*-I))<1$. In that case, by theorem 3, no neighbourhood exists about y* such that $y^{(s)}$ in that neighbourhood implies (3.6) converges.

(ii) By theorem 2, a real nondiagonal matrix exists such that $\rho(I + R(G^*-I))<1$ for any g(.,.) and y*. Hence (3.6) can be made convergent within a neighbourhood of y* because an $R^{(s)}$ can be found such that

$$\|y^{(s+1)} - y^*\| < \|I + R^{(s)}(G^{(s)}-I)\| \, \|y^{(s)}-y^*\| < \|y^{(s)}-y^*\| \qquad (4.7)$$

for some norm and <u>every</u> s, where $G^{(s)} = [\partial g/\partial y]$ at $y = y^{(s)}$ for any g(.,.) and $y^{(o)}$ (Young [16], p. 81).

COMMENTS: (a) This section has assumed throughout that (3.4) has at least one real solution. The iteration (3.6) is of course convergent if (4.7) holds for every s, but that does not mean every (or any particular) solution of (3.4) can be found by these means.

(b) Salmon's iterations <u>cannot</u> necessarily detect the remaining real solutions of (3.4) even if some iteration such as (3.6) has located the first solution. That is a consequence of theorem 4(i) with G* replaced by $G^{(k)} = [\partial g/\partial y]$ at $y = y_{(k)}$ for $k>1$.

(c) However a sequence of nondiagonal $R^{(s)}$ matrices exists which guarantees that the real solutions of (3.4) are all detected. One possibility is to set

$$R_k^{(s)} = \begin{bmatrix} \left|y_1^{(s)} - y_{(1)1}\right|^{-1} & \cdots & \left|y_n^{(s)} - y_{(1)n}\right|^{-1} \\ \vdots & & \vdots \\ \left|y_1^{(s)} - y_{(k-1)1}\right|^{-1} & \cdots & \left|y_n^{(s)} - y_{(k-1)n}\right|^{-1} \\ \hline & \cdots & \end{bmatrix} \qquad (4.8)$$

at run k=2...n of (3.6), where elements below the partition line may be set at any values which preserve convergence. It is not certain that $R_k^{(s)}$ will always satisfy (4.7) at each s, and further modifications to that end may be necessary. The key question therefore remains; how may a nondiagonal R matrix be chosen to guarantee convergence to each real solution?

## 5. CONCLUDING REMARKS.

The main result of this paper is that first order iterative techniques, which are guaranteed to converge in solving an arbitrary equation system, do indeed exist but they may require a nondiagonal matrix of extrapolation parameters. Thus multiparameter extrapolations which assign one parameter to each equation cannot necessarily resolve the cases where standard one parameter extrapolations diverge. Consequently neither the deflation method for detecting multiple solutions, nor Richardson's generalised method, have guaranteed convergence.

Simple iterative techniques are effective for solving equation systems because they approximate the standard Newton method in a way which avoids <u>any</u> derivative evaluations and because they exploit the sparseness of the system by avoiding any matrix inversions. Precisely the same advantage has established Powell's method as the most economical optimisation technique for typical econometric problems. The Newton method applied to (3.4) - rewritten as f(y,b) = 0 - would be

$$y^{(s)} = y^{(s-1)} - [F^{(s-1)}]^{-1} f(y^{(s-1)}, b) \qquad (5.1)$$

where $F^{(s-1)} = [\partial f/\partial y]$ evaluated at $y^{(s-1)}$. Hence, in the absence of normalisation, JOR can be regarded as a <u>restricted</u> Newton method with $F^{(s-1)} = \gamma^{-1}I$ for all s. Similarly Gauss-Seidel is (5.1) with $f(y^{(s)}, y^{(s-1)}, b)$. In either case, a stationary approximation to the inverted matrix of partial derivatives, proportional to the identity matrix, is imposed throughout. Multiparameter extrapolations impose the more flexible approximation of $R^{-1}$, which is still stationary and usually diagonal, for $F^{(s-1)}$. This restricted Newton interpretation explains the ultimate necessity for a full (and nonstationary) R matrix to ensure convergence; also why it may be expensive to find suitable values for R, and why (since $R^{-1}$ is not the true value of $F^{(s-1)}$) it is still possible to construct convergent first order iterations when the Newton method is divergent. Multiparameter extrapolations therefore represent a retreat towards Newton methods which may sacrifice the computational advantages of other techniques (e.g. second order iterations, which continue to exploit sparseness without derivative evaluations) without securing guaranteed convergence in exchange.

REFERENCES

1. BROWN, K.M. and GEARHART, W.B. Deflation Techniques for the Calculation of Further Solutions of a Nonlinear System, Num. Math., 16 (1971), 334-42.

2. GANTMACHER, F.R. Applications of the Theory of Matrices, Interscience Publishers Co., New York, 1959.

3. HADJIDIMOS, A. The optimal solution of the extrapolation problem of a first order scheme. Intern. J. Computer Math., 15 (1983), 153-68.

4. HAGEMAN, L.A. and YOUNG, D.M. Applied Iterative Methods, Academic Press, New York, 1981.

5. HMSO. Report of the Committee on Policy Optimisation, Cmnd 7148, HM Stationary Office, London, 1978.

6. HUGHES HALLETT, A.J. Some extensions and comparisons in the theory of Gauss-Seidel iterative techniques for solving large equation systems, In Proceedings of the 1979 European Econometric Society Meeting: Essays in Honour of Stefan Valavanis (Ed. E.G. Charatsis) North-Holland, Amsterdam, 1981.

7. HUGHES HALLETT, A.J. Alternative techniques for solving systems of nonlinear equations. J. of Comp. and Appl. Math., 8 (1982), 35-48.

8. HUGHES HALLETT, A.J. Second order iterations with guaranteed convergence, J. of Comp. and Appl. Math., 10 (1984), 285-91.

9. HUGHES HALLETT, A.J. Simple and optimal extrapolations for first order iterations, Intern. J. Computer Math., 16 (1984).

10. HUGHES HALLETT, A.J. Techniques which Accelerate the Convergence of First Order Iterations Automatically, Lin. Alg. and Applics., (forthcoming, 1985).

11. OSTRWOSKI, A.M. Solutions to Equations and Systems of Equations (second edition), Academic Press, New York, 1966.

12. RAO, C.R. Linear Statistical Inference and its Applications, John Wiley and Sons, New York, 1973.

13. SALMON, M.H. On the detection of successive solutions to nonlinear econometric models. Research report, University of Warwick, Coventry, England, 1983.

14. VARGA, R.S. Matrix Iterative Analysis, Prentice-Hall, Englewood Cliffs, NJ, 1962.

15. WONHAM, W.M. Linear Multivariate Control: A Geometric Approach, Springer Verlag, Berlin and New York, 1974.

16. YOUNG, D.M. Iterative Solution of Large Linear Systems. Academic Press, New York, 1971.