DEMONSTRATIO MATHEMATICA
Vol. XIV No 2 1981

Suhrit K. Dey

GENERALIZED CONVERGENCE ANALYSIS OF PERTURBED JACOBI ITERATIONS

1. Introduction

In [1] a functional iterative scheme to solve numerically nonlinear systems of equations has been developed by perturbing nonlinear Jacobi iterations. The same concept was broadened and discussed from various aspects, both mathematical and computational, in [4], [5], [6]. The advantages of such perturbed iterative schemes over other existing methods, for example, Newton's method, nonlinear Jacobi and Gauss-Seidel methods were also discussed in these articles [1], [4], [5], [6] with appropriate examples. These discussions are not repeated here. However, an analysis of convergence properties of these methods seem to be dealing with special cases of perturbed functional iterations. In this article, these properties have been derived from one uniform principle and as such theorems on convergence derived before become simply particular cases of the convergence theorem derived here. Thus, within the framework of this context, this article is of a theoretical nature. It may be worth mentioning that theorems on convergence of functional iterations given in [3] are also particular cases of the principle developed here. As an example, we considered here nonlinear Jacobi iterations.

In [2], stability and convergence properties of nonlinear partial difference equations were analyzed. This analysis deals with a sequence of matrices, having variable elements, which damp out instabilities and errors and thereby cause convergence. These matrices, called decaying matrices or D-matrices, have been found to be applicable to the convergence analysis of nonlinear functional iterative schemes. This study has been conducted in this article.

2. Mathematical preliminaries

Let R^n be a real n-dimensional space and $x=(x_1,x_2,\dots,x_n)^T \in Q \subset R^n$. Let A be a square matrix $(n \times n)$ whose elements are $a_{ij}(x) : Q \to Q$. If there exists a constant L such that for every $x \in Q$ $|a_{ij}(x)| \leq L$, $i = 1,2,\dots,n$; $j = 1,2,\dots,n$, then A is a bounded matrix on Q. Limits of matrices with variable elements have been discussed quite thoroughly in [7].

Let $x^k = (x_1^k,x_2^k,\dots,x_n^k)^T \in Q$. We consider a recurrsive matrix equation

$$(1) \qquad x^k = A_k x^{k-1} \quad (k = 1,2,\dots),$$

where A_k's are commuting, bounded square matrices $(n \times n)$ having elements $a_{ij}^k = a_{ij}(x^{k-1},b)$, where $b \in R^n$ remains invariant.

If there exist certain conditions such that as $k \to \infty$, $x^k \to 0$ the motion given by (1) is called a motion of decadence of x on Q. A study of these conditions will now be done.

Let us assume that $\{A_k\}$ forms a sequence of bounded square matrices of the same type. Let these matrices have certain properties such that the sequence: $A_1 \cdot A_2$, $A_1 \cdot A_2 \cdot A_3, \dots,$ $A_1 \cdot A_2 \dots \cdot A_k, \dots$ tend to Φ as $k \to \infty$. Then we have the following:

Definition 1. The sequence $\{A_k\}$ is called a decaying sequence of matrices and each matrix A_k is called a decaying matrix or a D-matrix on Q.

Since $\|A\| = 0$ iff $A = \Phi$, A_k s are D-matrices iff

$$\lim_{k\to\infty} \|A_1 A_2 \dots A_k\| = 0. \tag{2}$$

Several properties of D-matrices have been discussed in [2]. A few of them may be mentioned here.

Theorem 1. Equation (1) gives a motion of decayness iff for every k and $x^0 \in Q$, A_k is a D-matrix on Q.

Proof. $x^k = A_k x^{k-1} = A_k A_{k-1} x^{k-2} = \dots = A_k A_{k-1} \dots$ $\dots A_1 x^0$.

Since for every k A_k is a D-matrix, the matrices A_k are bounded and commuting. Thus, $\lim_{k\to\infty} A_1 A_2 \dots A_k = \Phi$ which proves the theorem.

Theorem 2. A sufficient condition for A_k to be a D-matrix for every $k \geqslant K$ is $\|A_k\|_q \leqslant \alpha < 1$ for some q-norm.

Proof. Since $\|A_1 A_2 \dots A_n\|_q \leqslant \|A_1\|_q \dots$ $\dots \|A_K\|_q \|A_{K+1}\|_q \dots \|A_n\|_q$ and for every $k \geqslant K$, $\|A_k\|_q \leqslant \alpha < 1$ hence $\lim_{n\to\infty} \|A_1 A_2 \dots A_n\|_q = 0$, which proves the theorem.

3. Perturbed Jacobi iterations for coupled nonlinear systems

Let us consider a coupled nonlinear system

$$x^i = G^i(x^1, x^2, \dots, x^n), \tag{3}$$

where

$$x^i = (x_1^i x_2^i \dots x_n^i)^T \in Q^i \subset R^n,$$

$$G^i = (G_1^i G_2^i \dots G_n^i)^T \in Q^i \subset R^n, \quad i=1,2,\dots,n.$$

Thus, each $G^i : Q^1 \times Q^2 \times \dots \times Q^n \subset R^n \times R^n \times \dots \times R^n \to Q^i$.

Let $Q = Q^1 \times Q^2 \times \dots \times Q^n$ and $R = R^n \times R^n \times \dots \times R^n$. We assume that (3) has a solution in Q, given by

$$x^{i,*} = \left(x_1^{i,*} \ x_2^{i,*} \dots x^{i,*}\right)^T \in Q^i.$$

Therefore, for every i

(4) $$x^{i,*} = G^i(x^{1,*}\ x^{2,*}\ \ldots\ x^{n,*}).$$

The nonlinear Jacobi iterations may be expressed as

(5) $$x^{i,k} = G^{i,k-1},$$

where $x^{i,k}$ is the value of x^i at the kth iteration and,

(6) $$G^{i,k-1} = G^i(x^{1,k-1}\ x^{2,k-1}\ \ldots\ x^{n,k-1}).$$

A perturbed Jacobi iteration, in the element form, may be expressed as

(7) $$x_j^{i,k} = w_j^{i,k} + G_j^{i,k-1}$$

$(i = 1,2,\ldots n;\ j = 1,2,\ldots n)$,
where $w_j^{i,k}$ are the perturbation parameters which are yet to be computed. To do this, we assume that:

(i) these parameters are small with respect to the other terms in (7) and the squares of them may be neglected,

(ii) for every $x^i \in Q^i$ and all i,j we have $\partial G_j^i/\partial x_j^i \neq 1$,

(iii) for every $x^i \in Q^i$ and all i,j, $\partial^2 G_j^i/\partial x_j^{i2}$ is bounded.

Assuming convergence after (k-1) iterations, we have for all i,j

$$x^{i,*} = x_j^{i,k} = x_j^{i,k-1} = G_j^{i,k-1}.$$

Hence from (7) we obtain

(8) $$w_j^{i,k} + G_j^{i,k-1} = G_j^i(x^{1,k-1}x^{2,k-1}\ldots x^{i-1,k-1};$$
$$x_1^{i,k-1}\ldots x_{j-1}^{i,k-1},\ w_j^{i,k} + G_j^{i,k-1},\ x_{j+1}^{i,k-1}\ldots x_n^{i,k-1};$$
$$x^{i+1,k-1}\ldots x^{n,k-1}).$$

If the right side of (8) is expanded in a Taylor's series and the assumptions (i), (ii) and (iii) are applied, one may get

$$(9)\qquad w_j^{i,k} = \left(\hat{G}_j^{i,k-1} - G_j^{i,k-1}\right)\Big/\left(1 - \partial_j\, \hat{G}_j^{i,k-1}\right)$$

$$(i = 1,2,\dots n;\ j = 1,2,\dots n;\ k = 1,2,\dots),$$

where,

$$(10)\qquad \hat{G}_j^{i,k-1} = G_j^i(x^{1,k-1},\dots x^{i-1,k-1};\ x^{i,k-1}\dots$$

$$\dots x_{j-1}^{i,k-1}, G_j^{i,k-1},\ x_{j+1}^{i,k-1},\dots x_n^{i,k-1};\ x^{i+1,k-1},\dots$$

$$\dots, x^{n,k-1})$$

and

$$(11)\qquad \partial_j \hat{G}_j^{i,k-1} = \left[\frac{\partial G_j^i}{\partial x_j^i}\right]_{(x^{1,k-1},\dots,x^{i-1,k-1};x_1^{i,k-1},\dots, x_{j-1}^{i,k-1}, G_j^{i,k-1}, x_{j+1}^{i,k-1},\dots, x_n^{i,k-1}; x^{i+1,k-1},\dots x^{n,k-1})}.$$

Thus the algorithm of a perturbed Jacobi iteration to solve coupled nonlinear system numerically consists of the following steps:

1. Make an initial guess $x^{i,0} = (x^{1,0}\dots x^{n,0})^T \in Q^i$ for each $i = 1,2,\dots n$.

Then, at some kth iteration

2. Compute $w_j^{i,k}$ using (9).
3. Compute $x_j^{i,k}$ using (7).

The convergence criterion is

$$(12)\qquad \max_{i,j} |w_j^{i,k}| < \varepsilon,$$

where ε is positive and arbitrarily small.

In [1] we have recorded that condition (12) may be both necessary and sufficient for convergence. Here, we will apply a more general analysis of this property.

4. Analysis of convergence

Let $x = (x^1\ x^2\ \dots\ x^n)^T$. Then $x \in Q$ and (3) may be expressed as $x = G(x)$ where $G : Q \longrightarrow Q$.

By convergence of iterations it is meant

$$\lim_{k\to\infty} x_j^{i,k} = x_j^{i,*} \quad \text{for all } i,j.$$

This implies

(13) $$\lim_{k\to\infty} x^k = x^*.$$

Obviously, (7) may be expressed as

(14) $$x^k = w^k + G^{k-1}.$$

Since convergence of iterations according to (4) also implies $\lim_{k\to\infty} G^{k-1} = G^* = x^*$, a necessary condition for convergence is that

(15) $$\lim_{k\to\infty} \| w^k \| = 0$$

for some norm. This equation implies (12). We need to prove now that (15) may also be a sufficient condition for convergence.

Let $|x| = (|x_1||x_2| \dots |x_n|)^T$. Let $x^k = H(x^{k-1})$ be a recursive relation such that $x^k \in Q$. Furthermore, let for $\alpha \in Q$

(16) $$|H(x^k) - H(\alpha)| = A_k\, |x^k - \alpha|,$$

where A_k is an isotone matrix (and obviously for nonlinear cases, $A_k = A_k(x^k, \alpha)$).

D e f i n i t i o n 2. If for every k A_k is a D-matrix on Q, H is called a D-mapping on Q.

T h e o r e m 3. If G is a D-mapping on Q and

$$\lim_{k \to \infty} |w^k| = \emptyset \tag{17}$$

the iterative scheme (14) will converge to x^* and x^* is the fixed point of G in Q. Furthermore, if $(I - A_k)^{-1}$ exists for all $x^k \in Q$, x^* is the unique fixed point of G.

P r o o f . Because we have

$$\begin{aligned} |x^k - x| &\leqslant |w^k| + |G(x^{k-1}) - G(x^*)| \leqslant \\ &\leqslant |w^k| + A_k|x^{k-1} - x^*| \leqslant \dots \\ &\leqslant \sum_{j=1}^{k} A_k A_{k-1} \dots A_{j+1} |w^j| + A_k A_{k-1} \dots A_1 |x^0 - x^*| \leqslant \\ &\leqslant \sum_{j=1}^{k_0} (A_k A_{k-1} \dots A_{k_0+1})(A_{k_0} A_{k_0-1} \dots A_{j+1}) |w^j| + \\ &+ \sum_{j=k_0+1}^{k} (A_k A_{k-1} \dots A_{j+1}) |w^j| + A_k A_{k-1} \dots \\ &\dots A_1 |x^0 - x^*|, \end{aligned} \tag{18}$$

and equation (17) implies that for some $k \geqslant k_0 + 1$, $|w^k| < \varepsilon$, hence from (18) we get

$$\begin{aligned} |x^k - x^*| &\leqslant (A_k A_{k-1} \dots A_{k_0+1}) \sum_{j=1}^{k_0} A_{k_0} A_{k_0-1} \dots A_{j+1} |w^j| + \\ &+ \sum_{j=k_0+1}^{k} (A_k A_{k-1} \dots A_{j+1}) \varepsilon + A_k A_{k-1} \dots A_1 |x^0 - x^*| . \end{aligned} \tag{19}$$

Since $\| |x| \| = \|x\|$, we have for any norm

$$(20)\quad \|x^k - x^*\| \leq \|A_k A_{k-1} \dots A_{k_0+1}\| \; \|\sum_{j=1}^{k_0} (A_{k_0} A_{k_0-1} \dots A_{j+1}) |w^j| \| +$$

$$+ \|\sum_{j=k_0+1}^{k} A_k A_{k-1} \dots A_{j+1}\| \cdot \|\varepsilon\| + \|A_k A_{k-1} \dots A_1\| \cdot \|x^0 - x^*\|.$$

Since A_k is a D-matrix, $\lim_{k\to\infty} A_1 A_2 \dots A_k = \emptyset$, giving $\lim_{k\to\infty} \|A_1 A_2 \dots A_k\| = 0$ and $\lim_{k\to\infty} \|A_{k_0+1} A_{k_0+2} \dots A_k\| = 0$. Hence, ε being arbitrarily small, $\lim_{k\to\infty} \|x^k - x^*\| = 0$ which establishes convergence.

To prove uniqueness, we assume that $x = y^*$ is a second root. Then

$$|x^* - y^*| = |G(x^*) - G(y^*)| = A_*|x^* - y^*|.$$

Thus,

$$(I - A_*)|x^* - y^*| = \emptyset.$$

Since, $(I - A_k)$ is nonsingular for all k, we have $|x^* - y^*| = \emptyset$ or, $x^* = y^*$. For the ordinary nonlinear Jacobi iterations,

$$(21)\quad x^k = G(x^{k-1}),$$

and we can prove the following theorem.

T h e o r e m 4. The iterations (21) converge to x^* iff G is a D-mapping on Q.

P r o o f . Since $x^* \in Q$, we have

$$|x^k - x^*| = |G(x^{k-1}) - G(x^*)| = A_k |x^{k-1} - x^*| =$$

$$= \dots = A_k A_{k-1} \dots A_1 |x^0 - x^*|.$$

Thus convergence follows for all $|x^0 - x^*|$ iff A_k is a D-matrix.

We may also prove that

T h e o r e m 5. If G is a D-mapping and for every k $(I - A_k)^{-1}$ exists, then $x = x^*$ is the unique fixed point of G on Q.

The proof follows from Theorems 3 and 4.

These theorems reveal the apparent differences of convergence properties of perturbed and ordinary Jacobi iterations for nonlinear equations.

5. Discussions

Practical applications of D-matrices may be found in [2]. The basic concept of such a matrix is certainly an extension of that of a convergent matrix. One main drawback of convergent matrices is that product of two convergent matrices is not necessarily a convergent matrix. For example if

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}$$

then $p(A) = p(B) = 0$, where $p(A)$ is a spectral radius of A, but

$$AB = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}, \quad p(AB) = 1.$$

Thus, although A and B are convergent matrices, AB is not a convergent matrix. If for every k, $\| A_k \| < 1$ for some norm, A_k is a convergent matrix but may not be a D-matrix. But if for every k, $\| A_k \| \leqslant \alpha < 1$, A_k is both a convergent matrix as well as a D-matrix. However, if $\{A_k\}$ form a sequence of D-matrices their product $\rightarrow \emptyset$, whereas if $\{A_k\}$ form a sequence of convergent matrices, their product may not tend to be a null matrix. There are many other very subtle differences between these two groups of matrices some of which have been discussed in [2]. We may still note that

Theorem 6. If for every k, $A_k = A$, then A is a D-matrix iff A is a convergent matrix. The proof is rather simple [2].

By introducing the concept of D-matrices, convergence theorems in [1], [3] for nonlinear perturbed and ordinary Jacobi iterations become particular cases of Theorems 3, 4 and 5. These theorems may be extended for a generalized convergence analysis of perturbed Gauss-Seidel iteration [5]. The primary advantages of these perturbed functional iterations are:

(i) they have simpler algorithms and

(ii) they have quadratic rates of convergence in the vicinity of the root.

Also, it has been found computationally that these methods displayed global convergence properties. Detailed discussions on these and their applications may be found in [1], [4], [6].

Acknowledgment. This work was partially funded by a grant from the Council on Faculty Research of Eastern Illinois University.

REFERENCES

[1] S.K. Dey: Numerical solutions of a system of nonlinear equations by perturbations, Demonstratio Math. 9 (1976) 691-705.

[2] S.K. Dey: Numerical instabilities of nonlinear partial differential equations, CFT 7800/SKD/IE, von Karman Institute for Fluid Dynamics, Brussels, Belgium 1978.

[3] J.M. Ortega, W.C. Rheinboldt: Iterative solution of nonlinear equations in several variables. New York 1970.

[4] S.K. Dey: Numerical solution of nonlinear equations with real and complex variables, Franklin Institute J. Vol 307, No. 1, (1979) 21-30.

[5] S.K. D e y : Convergence properties of perturbed iterative scheme for n-coupled nonlinear systems. (to appear in the Internat. Math. and Math. Sci.).

[6] S.K. D e y : Perturbed iterative solution of coupled nonlinear systems with applications to fluid dynamics, J. Comput. Appl. Math. 4 (1978) 117-128.

[7] D.K. F a d d e e v , V.N. F a d d e e v a : Computational methods of linear algebra. San Francisco 1963.

[8] A. O s t r o w s k i : Solution of equations and systems of equations. New York 1960.

DEPARTMENT OF MATHEMATICS, EASTERN ILLINOIS UNIVERSITY, CHARLESTON, IL. 61920, U.S.A.
Received January 3, 1980.