Technical Report TR-66-32 July 1966
NsG-398

Monotone Iterations for Nonlinear Equations With Application to Gauss-Seidel Methods

by
James M. Ortega
Research Associate Professor
and
Werner C. Rheinboldt
Research Professor

Computer Science Center
University of Maryland

CONTENTS

ABSTRACT

In this paper we study iterative processes of the form $y_{k+1} = y_k - B_k F y_k$ for approximating solutions of a system of nonlinear equations $Fy = 0$. We obtain monotonic behavior of the iterates y_k, in the sense of the natural partial ordering in real n-dimensional Euclidean space, if F satisfies a generalized convexity condition and the B_k are subinverses of $F'(y_k)$, i.e. $B_k F'(y_k) \leqq I$, $F'(y_k) B_k \leqq I$. Our results contain recent similar ones of Greenspan and Parter as well as a classical one of Kantorovich. We also study two-sided iterations as well as iterations defined by implicit processes such as the nonlinear Gauss-Seidel method. In addition, a class of iterative processes combining Newton's method with the Gauss-Seidel iteration is considered and an application is made to mildly nonlinear boundary-value problems.

Monotone Iterations for Nonlinear Equations With Application to Gauss-Seidel Methods

James M. Ortega and Werner C. Rheinboldt[1)]

1. Introduction

Recently Greenspan and Parter [7] have studied monotone iterative processes for solving discrete analogues of mildly nonlinear elliptic boundary value problems. In this paper we extend these results and incorporate them into a general theory for a broad class of monotone iterations. This class of iterations includes Newton's method as well as a family of methods, which we call Newton-Gauss-Seidel processes, that are obtained by using the Gauss-Seidel iteration on the linear systems of Newton's method. Our results also include the monotone iterations of Kantorovich [8] for obtaining fixed-points of isotone operators.

The theory is based upon generalized convexity conditions as well as the notion of a subinverse of a linear operator. The approach is related to the basic work of Baluev [2], [3] in which the Chaplygin method for differential equations

---

1) This work was supported in part under NASA grant NsG-398 and NSF grant PIVR06 to the University of Maryland.

(see e.g., [4]) is considered in abstract spaces. More recently, Slugin (see e.g., [14] - [17]) has extended Baluev's results in directions somewhat different from ours. Monotone sequences which provide upper and lower bounds for solutions of operator equations have also been considered by Albrecht, Collatz, Schmidt, and Schröder (see e.g., [1], [6], [11], [13]). But their work uses a basically different approach and does not appear to have a direct connection to the results discussed here.

For simplicity we have restricted our presentation to finite dimensional spaces. However, most of the discussion extends immediately to more general partially ordered linear topological spaces, provided suitable restrictions are placed on the connection between the topology and the partial ordering in order to assure convergence. For a discussion of these topological considerations in connection with Newton's method, see Vandergraft [18].

In Section 2 we define subinverses of linear operators and show the relation to the regular splittings of Varga [19]. Section 3 contains a discussion of various convexity properties of nonlinear operators; the material in this section was developed in connection with J. Vandergraft (see [18] for an extension of some of these results to more general spaces).

In Section 4 we present our main results and apply them to the special cases of convex as well as isotone operators. Then in Section 5 we consider the Newton-Gauss-Seidel processes and in Section 6 we apply our results to mildly nonlinear boundary value problems and show the relation to the results of [7]. Finally, in Section 7 we give a theorem for implicitly defined iterates and show its application to the non-linear Gauss-Seidel method studied by Bers [5] and Schechter [10].

2. Subinverses and Regular Splittings

Let R^n be the n-dimensional real coordinate space and $\mathfrak{M}^n$ the space of all real n x n matrices. For vectors $x,y \in R^n$ and matrices $A,B \in \mathfrak{M}^n$ we denote by $x \leqq y$ and $A \leqq B$ the usual componentwise partial orderings.

Definition 2.1: Let $A \in \mathfrak{M}^n$; then any $B \in \mathfrak{M}^n$ such that

(2.1) $$AB \leqq I \ , \quad BA \leqq I,$$

where I is the identity, is called a subinverse of A.

We note some obvious properties of subinverses: The null-matrix is a subinverse of any matrix. If A is a subinverse of B then B is a subinverse of A. If B and C are subinverses of A then so is $\lambda B + (1-\lambda)C$ for $0 \leqq \lambda \leqq 1$. If A^{-1} exists then it is a subinverse of A.

Varga [19] defines a decomposition $A = B - C$ to be a <u>regular splitting</u> of A if B is non-singular, $B^{-1} \geqq 0$, and $C \geqq 0$. There is a close connection between regular splittings and subinverses.

<u>Definition 2.2</u>: Let $A \in \mathfrak{M}^n$; then $A = B - C$ is a <u>weak regular splitting</u> of A if B is non-singular, $B^{-1} \geqq 0$, $B^{-1}C \geqq 0$, and $CB^{-1} \geqq 0$.

Clearly any regular splitting is also a weak regular splitting. The connection to subinverses is given by the following lemma:

<u>Lemma 2.1</u>: If $A = B - C$ is a weak regular splitting of A, then B^{-1} is a subinverse of A. Conversely, if $B \geqq 0$ is a non-singular subinverse of A, then $A = B^{-1} - (B^{-1}-A)$ is a weak regular splitting.

<u>Proof</u>: Let $A = B - C$ be a weak regular splitting; then

$$0 \leqq B^{-1}C = B^{-1}(B-A) = I - B^{-1}A,$$

and hence $B^{-1}A \leqq I$. Similarly, $AB^{-1} \leqq I$ follows from $CB^{-1} \geqq 0$. Conversely, if $B \geqq 0$ is a non-singular subinverse of A, then $0 \leqq I - BA = B(B^{-1} - A)$; similarly $(B^{-1} - A)B \geqq 0$, and hence $A = B^{-1} - (B^{-1}-A)$ is a weak regular splitting.

Weak regular splittings can be used to generate subinverses which appear in a natural way in the study of Gauss-Seidel type iterative processes (see Section 5).

<u>Lemma 2.2</u>: Let $A = B - C$ be a weak regular splitting and set $H = B^{-1}C$. Then for any $m \geqq 1$,

$$K_m = (I + H + \dots + H^{m-1})\, B^{-1} \tag{2.2}$$

is a subinverse of A.

<u>Proof</u>: Using the identity

$$(I + \dots + H^{m-1})(I-H) = (I-H)(I+\dots+H^{m-1}) = I - H^m$$

and $H \geqq 0$ we obtain

$$K_m A = (I+\dots+H^{m-1})B^{-1}(B-C) = I - H^m \leqq I. \tag{2.3}$$

Similarly, since $CB^{-1} \geqq 0$,

$$AK_m = (B-C)(I+\dots+H^{m-1})B^{-1} = B\,(I-H^m)B^{-1} = I - (CB^{-1})^m \leqq I.$$

It is of interest to know when K_m^{-1} exists and when $A = K_m^{-1} - (K_m^{-1} - A)$ is a weak regular splitting. For this we need an extension of a result of Varga [19] who showed that if $A = B - C$ is a regular splitting and $A^{-1} \geqq 0$, then $B^{-1}C$ is convergent, i.e., $B^{-1}C$ has spectral radius $\rho(B^{-1}C)$ less than one.

<u>Lemma 2.3</u>: Let $A = B - C$ be a weak regular splitting. Then $\rho(B^{-1}C) < 1$ if and only if A is non-singular and $A^{-1} \geqq 0$.

<u>Proof</u>: Again set $H = B^{-1}C$; then using (2.3) we see that $A^{-1} \geqq 0$

implies $0 \leqq (I+\ldots+H^{m-1})B^{-1} \leqq A^{-1}$ for all m. Since $B^{-1} \geqq 0$ contains a non-zero element in each row and column it follows that $I+\ldots+H^{m-1} (\geqq 0)$ is bounded above for all m; hence $\lim_{m\to\infty} H^m = 0$, and $\rho(H) < 1$. Conversely, if $\rho(H) < 1$, then $(I-H)^{-1}$ exists and $(I-H)^{-1} \geqq 0$. Thus A^{-1} exists and $A^{-1} = (I-H)^{-1}B^{-1} \geqq 0$.

By means of Lemma 2.3 we now have

<u>Lemma 2.4</u>: Let $A = B - C$ be a weak regular splitting, set $H = B^{-1}C$ and, for any $m \geqq 1$, define K_m by (2.2). Suppose A is non-singular and $A^{-1} \geqq 0$. Then K_m^{-1} exists and $A = K_m^{-1} - (K_m^{-1} - A)$ is a weak regular splitting.

<u>Proof</u>: From Lemma 2.3 it follows that H is convergent. Hence $(I-H^m)^{-1}$ and, by (2.3), K_m^{-1} exist. By Lemma 2.2, K_m is a subinverse of A and, since $K_m \geqq 0$, the result is a direct consequence of Lemma 2.1.

It is easy to give examples of weak regular splittings that are not regular splittings. Moreover, even if $A = B - C$ is a regular splitting, the weak regular splittings of Lemma 2.4 are not necessarily also regular splittings as the following example shows[2]:

---

[2] We are indebted to R. Elkin for this example.

Let

$$A = \begin{pmatrix} 2 & -1 & 0 \\ 0 & 1 & -1 \\ -1 & 0 & 1 \end{pmatrix} \qquad B = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}.$$

Then

$$K_2^{-1} = \begin{pmatrix} 2 & -2 & 2 \\ 0 & 2 & -2 \\ -1 & 0 & 1 \end{pmatrix}$$

and clearly $K_2^{-1} - A$ is not nonnegative. Note that here we have used the usual Gauss-Seidel splitting and A is an M-matrix, i.e., $a_{ij} \leqq 0$ for $i \neq j$ and $A^{-1} \geqq 0$.

In the following sections, we shall frequently assume that a given matrix has a nonsingular, nonnegative subinverse. In most cases, it will be evident that such a subinverse can be found, but the general question of the existence of such subinverses is unresolved. A typical negative result is the following:

Suppose the kth row of $A = (a_{ij})$ is nonnegative and has at least two non-zero elements a_{kp} and a_{km}. The pth and mth rows of any nonnegative subinverse of A are zero. In particular, if A has any strictly positive row, then the only nonnegative subinverse of A is the null-matrix. A corresponding result holds for columns.

3. Convexity and Order-Convexity

<u>Definition 3.1</u>: Let $F: D \subset R^n \to R^m$ denote an operator defined

on some domain D in R^n. Then F is called <u>order-convex</u> on a convex subset $D_o \subset D$ if

$$F(\lambda x + (1-\lambda)y) \leqq \lambda Fx + (1-\lambda) Fy \tag{3.1}$$

whenever $x,y \in D_o$ are comparable (i.e., $x \leqq y$ or $y \leqq x$) and $0 \leqq \lambda \leqq 1$. If (3.1) holds for all $x,y \in D_o$ and $0 \leqq \lambda \leqq 1$, then F is said to be <u>convex</u>.

If we denote the components of Fx by $f_i(x)$, $i=1,\ldots,m$, then F is convex or order-convex if and only if each $f_i: D \subset R^n \to R^1$ has the same property.

As for real valued functions, it is possible to characterize convexity properties of F in terms of properties of the derivatives. We say that F is differentiable on $D_o \subset D$ if the Gateaux derivative $F'(x)$ exists for all $x \in D_o$, i.e., if

$$\lim_{t\to o} \frac{1}{t}[F(x+th) - Fx] = F'(x)h \quad , x \in D_o, \quad h \in R^n,$$

where $F'(x)$ is the m x n Jacobian matrix

$$F'(x) \equiv \left(\frac{\partial f_i}{\partial x_j}(x) \right) . \tag{3.2}$$

F is continuously differentiable on D_o if all elements of $F'(x)$ are continuous on D_o and we write in that case $F \in C^1(D_o)$. For $F \in C^1(D_o)$ we have the mean value theorem

$$f_k(y) - f_k(x) = \sum_{i=1}^{n} \int_o^1 \frac{\partial f_k}{\partial x_i}(x+t(y-x))(y_i-x_i)dt,$$

or

$$(3.3) \qquad Fy - Fx = \int_0^1 F'(x+t(y-x))(y-x)dt.$$

The following lemma provides a characterization of convexity and order-convexity in terms of the first derivative:

Lemma 3.1: Suppose that $F: D \subset R^n \to R^m$ is differentiable on a convex set $D_o \subset D$. Then F is order-convex on D_o if and only if

$$(3.4) \qquad F'(x)(y-x) \leqq Fy - Fx$$

for all comparable $x,y \in D_o$. F is convex on D_o if and only if (3.4) holds for all $x,y \in D_o$.

If $F \in C^1(D_o)$, then F is order convex if and only if

$$(3.5) \qquad F'(x)(y-x) \leqq F'(y)(y-x)$$

for all comparable $x,y \in D_o$. F is convex on D_o if and only if (3.5) holds for all $x,y \in D_o$.

Proof: Suppose (3.4) holds for all comparable $x,y \in D_o$. For given comparable x and y and $0 \leqq \lambda \leqq 1$ set $z = \lambda x + (1-\lambda)y$. Then $z \in D_o$ is comparable with x and y so that $Fy - Fz \geqq F'(z)(y-z)$ and $Fx - Fz \geqq F'(z)(x-z)$. Thus

$$\lambda Fx + (1-\lambda)Fy - Fz \geqq F'(z)[\lambda x + (1-\lambda)y - z] = 0 .$$

Conversely, if F is order convex on D_o, then for any $0 < t \leqq 1$

and any comparable $x,y \in D_o$,

$$\frac{1}{t}[F(x+t(y-x)) - Fx] \leq Fy - Fx,$$

and (3.4) follows as $t \to 0$.

If F is order convex, then (3.4) gives

$$F'(x)(y-x) \leq Fy - Fx \leq F'(y)(y-x)$$

for all comparable $x,y \in D_o$. Conversely, if $F \in C^1(D_o)$ and (3.5) holds for all comparable $x,y \in D_o$, then it follows from (3.3) that

$$Fy - Fx = \int_0^1 F'(x+t(y-x))(y-x)dt \geq \int_0^1 F'(x)(y-x)dt = F'(x)(y-x).$$

The proofs for the convex case proceed analogously.

Clearly, (3.5) is satisfied if F' is an isotone function of x, i.e., if $x \leq y$ implies that $F'(x) \leq F'(y)$. Thus if F' is continuous and isotone on D_o, then F is order convex. It also may be shown that an operator $F \in C^1(D_o)$ is order-convex on the convex set $D_o \subset D$ if (3.4) only holds for all $x,y \in D_o$ such that $x \leq y$ (or alternatively, such that $y \leq x$).

We proceed now to a characterization of convexity in terms of the second derivative. An operator $F: D \subset R^n \to R^m$ is called twice differentiable on $D_o \subset D$ if its second Gateaux derivative $F''(x)$ exists for all $x \in D_o$. In that case, all second partial derivatives of the components f_i exist on D_o. For each $x \in D_o$,

$F''(x)$ is a bilinear operator from $R^n \times R^n$ into R^m, and for $u,v \in R^n$, the kth component of $F''(x)uv$ is given by $u^T f_k''(x)v$ where $f_k''(x)$ is the n x n Hessian matrix

$$(3.6) \qquad f_k''(x) \equiv \left(\frac{\partial^2 f_k}{\partial x_i \partial x_j} (x) \right) .$$

F is twice continuously differentiable on D_o, $F \in C^2(D_o)$, if each $f_k''(x)$ is continuous in x on D_o. In this case, each of the matrices $f_k''(x)$ is symmetric; moreover, we have the mean value theorem

$$(3.7) \quad Fy - Fx - F'(x)(y-x) = \int_0^1 F''(x+t(y-x))(y-x)(y-x)dt .$$

<u>Lemma 3.2</u>: Let $F: D \subset R^n \to R^m$ be twice continuously differentiable in an open convex set $D_o \subset D$. Then F is order convex in D_o if and only if

$$(3.8) \qquad F''(x)\, h\, h \geqq 0$$

for all $x \in D_o$ and all $h \geqq 0$ in R^n. F is convex in D_o if and only if (3.8) holds for all $x \in D_o$ and all $h \in R^n$.

<u>Proof</u>: For order convex F and any $x \in D_o$, $h \geqq 0$, and sufficiently small $t \geqq 0$, we have by Lemma 3.1 that $F'(x+th)h \geqq F'(x)h$. Hence,

$$F''(x)hh = \lim_{t \to o} \frac{1}{t} [F'(x+th)h - F'(x)h] \geqq 0.$$

Conversely, suppose that (3.8) holds for all $x \in D_o$ and $h \geqq 0$.

Let $x,y \in D_o$ be such that $x \leqq y$ and set $h = y-x$. Then (3.7) implies that $Fy - Fx - F'(x)(y-x) \geqq 0$. If $y \leqq x$, then $h \leqq 0$ but the right hand side of (3.7) remains non-negative. Hence F is order convex.

The proof for the convex case proceeds analogously.

Under the conditions of Lemma 3.2, F is convex in D_o if and only if for each $x \in D_o$ the matrices $f_k''(x)$ of (3.6) are all positive semidefinite, i.e., if

$$(3.9) \qquad h^T f_k''(x)\, h \geqq 0 \ , \quad k=1,\ldots,m$$

for all $h \in R^n$ and $x \in D_o$. On the other hand, F is order convex if and only if (3.9) holds for all $x \in D_o$ and $h \geqq 0$. Thus a sufficient, but not necessary, condition that F be order convex is that $f_k''(x) \geqq 0$ for all $x \in D_o$ and $k=1,\ldots,m$. In this case we write $F''(x) \geqq 0$, $x \in D_o$. Note that for $F \in C^2(D_o)$ the condition $F''(x) \geqq 0$ for $x \in D_o$ implies that

$$F'(y) - F'(x) = \int_0^1 F''(x+t(y-x))(y-x)dt \geqq 0$$

whenever $x,y \in D_o$, $x \leqq y$, i.e., that F' is isotone on D_o.

These results also provide a simple example of an order convex but not convex operator F. In fact, the quadratic form $F: R^n \to R^1$, $Fx = x^TAx$, where $A \geqq 0$ but A is not positive semi-definite, has this property.

We end this section with a result of a different kind:

<u>Lemma 3.3</u>: Suppose $F: D \subset R^n \to R^m$ is convex on D and $A: R^p \to R^n$ is a linear operator. Then the composite function $Gx = FAx$ is convex on $\hat{D} = \{x \in R^p \mid Ax \in D\}$. If F is order convex and $A \geqq 0$ then G is order convex on $\hat{D}$.

<u>Proof</u>: Let F be convex and $x, y \in \hat{D}$. Then for $0 \leqq \lambda \leqq 1$

$$(3.10) \qquad G(\lambda x + (1-\lambda)y) = F(\lambda Ax + (1-\lambda)Ay)$$

$$\leqq \lambda FAx + (1-\lambda)FAy = \lambda Gx + (1-\lambda)Gy \; .$$

If x and y are comparable and $A \geqq 0$ then Ax and Ay are also comparable. Hence (3.10) still holds if F is order convex.

4. <u>Convergence Theorems</u>

We consider now the construction of sequences which converge monotonically to a solution of $Fx = 0$. For any points $x_o \leqq y_o$, $[x_o, y_o]$ denotes the interval $\{x \in R^n \mid x_o \leqq x \leqq y_o\}$, and $y_k \downarrow y^*$ shall mean that $y_o \geqq y_1 \geqq \dots \geqq y_k \geqq y_{k+1} \geqq y^*$ and $\lim_{k \to \infty} y_k = y^*$. The main result is given by the following.

<u>Theorem 4.1</u>: Let $F: D \subset R^n \to R^n$ and suppose there exist points $x_o, y_o \in D$ such that

$$(4.1) \qquad x_o \leqq y_o, \; [x_o, y_o] \subset D, \quad Fx_o \leqq 0 \leqq Fy_o.$$

Assume there is a mapping $A: [x_o, y_o] \to \mathfrak{M}^n$ such that

$$(4.2) \qquad Fy - Fx \leqq A(y)(y-x) \quad , \; x_o \leqq x \leqq y \leqq y_o.$$

Then the sequence

(4.3) $$y_{k+1} = y_k - B_k Fy_k \quad , \; k=0,1,\ldots ,$$

where B_k is any non-negative subinverse of $A(y_k)$, is well-defined and there exists a $y^* \in [x_o, y_o]$, such that

(4.4) $$y_k \downarrow y^* \text{ for } k \to \infty .$$

Moreover, any solution of $Fx = 0$ in $[x_o, y_o]$ is contained in $[x_o, y^*]$, and if F is continuous at y^* and there exists a non-singular matrix $B \geqq 0$ such that

(4.5) $$\liminf_{k \to \infty} B_k \geqq B,$$

then $Fy^* = 0$.

<u>Proof</u>: From $B_o \geqq 0$ and $Fy_o \geqq 0$ it follows that $y_1 \leqq y_o$. Using (4.1) - (4.3), together with the fact that $B_o \geqq 0$ is a subinverse of $A(y_o)$, we find that for any $x \in [x_o, y_o]$

(4.6) $$x - B_oFx = y_1 - (y_o - x) + B_o(Fy_o - Fx) \leqq y_1 - [I - B_oA(y_o)](y_o - x) \leqq y_1 .$$

Hence, in particular, $x_o \leqq x_o - B_oFx_o \leqq y_1$. Similarly, we obtain

$$Fy_1 \geqq Fy_o + A(y_o)(y_1 - y_o) = [I - A(y_o)B_o]\, Fy_o \geqq 0.$$

Proceeding in the same manner we see by induction that

(4.7) $$y_{k-1} \geqq y_k \geqq x_o \quad , \; Fy_k \geqq 0 \; , \; k=1,2,\ldots .$$

Hence, as a monotone decreasing sequence that is bounded below,

$\{y_k\}$ has a limit $y^* \geqq x_o$.

If z is any solution of $Fx = 0$ in $[x_o, y_o]$, then (4.6) implies that $z = z - B_oFz \leqq y_1$ and by induction that $z \leqq y_k$ for all k. Hence $z \leqq y^*$. Finally, if F is continuous at y^*, it follows from (4.7) that $Fy^* \geqq 0$. If, in addition, (4.5) holds, then

$$0 = \liminf\,(y_{k+1} - y_k + B_kFy_k) = (\liminf B_k)\; Fy^* \geqq B\; Fy^* \geqq 0$$

and $BFy^* = 0$. Therefore, since B is nonsingular, $Fy^* = 0$. This completes the proof.

We note that the existence condition (4.5) can be replaced by other conditions which guarantee that the B_k are bounded away from singularity; for example

$$\liminf_k \|B_kx\| \geqq \alpha\,\|x\| \quad , \alpha > 0 \; , \; x \in R^n \; .$$

Also, there are other versions of Theorem 4.1 corresponding to different sign configurations. We indicate these for reference in Table 1 where the first column represents the theorem as stated.

$x_o \leqq y_o$	$x_o \leqq y_o$	$x_o \geqq y_o$	$x_o \geqq y_o$
$Fx_o \leqq 0 \leqq Fy_o$	$Fx_o \geqq 0 \geqq Fy_o$	$Fx_o \geqq 0 \geqq Fy_o$	$Fx_o \leqq 0 \leqq Fy_o$
$Fy-Fx \leqq A(y)(y-x)$	$Fy-Fx \geqq A(y)(y-x)$	$Fy-Fx \leqq A(y)(y-x)$	$Fy-Fx \geqq A(y)(y-x)$
$B_k \geqq 0$	$B_k \leqq 0$	$B_k \leqq 0$	$B_k \geqq 0$
$y_{k+1} \leqq y_k$	$y_{k+1} \leqq y_k$	$y_{k+1} \geqq y_k$	$y_{k+1} \geqq y_k$

Table 1

As a first corollary, we consider the construction of an additional monotonically increasing sequence starting from x_o.

Corollary 4.1: Assume that - except for (4.5) - the conditions of Theorem 4.1 are satisfied and, in particular, that the sequence $\{y_k\}$ is defined by (4.3). Suppose, in addition, that A is isotone, i.e.,

$$(4.8) \qquad A(x) \leq A(y) \qquad \text{whenever } x_o \leq x \leq y \leq y_o .$$

Then the sequence

$$(4.9) \qquad x_{k+1} = x_k - C_k Fx_k , \quad k=0,1,\ldots ,$$

where C_k is any non-negative subinverse of $A(y_k)$, is well-defined and there exists an $x^* \in [x_o, y^*]$ such that

$$(4.10) \qquad x_k \uparrow x^* \qquad \text{for } k \to \infty.$$

Moreover, the interval $[x^*, y^*]$ contains all solutions of $Fx = 0$ in $[x_o, y_o]$, and if F is continuous at x^* and there exists a non-singular $C \geq 0$ such that

$$(4.11) \qquad \liminf_k C_k \geq C$$

then $Fx^* = 0$.

Proof: Clearly $Fx_o \leq 0$ and $C_o \geq 0$ imply that $x_1 \geq x_o$, and, in a manner similar to the proof of (4.6), we see that

$$y_o \geq y_o - C_o F y_o = x_1 + (y_o - x_o) + C_o (Fx_o - Fy_o)$$
$$\geq x_1 + [I - C_o A(y_o)](y_o - x_o) \geq x_1 .$$

Now using (4.2) and the isotonicity of A it follows that

$$Fx_1 \leqq Fx_o + A(x_1)(x_1-x_o) \leqq [I-A(y_o)C_o]\, Fx_o \leqq 0$$

and hence, from (4.6), $x_1 \leqq x_1 - B_oFx_1 \leqq y_1$. By induction, we then see that

$$x_{k-1} \leqq x_k \leqq y_k \quad , \quad Fx_k \leqq 0 \ , \ k=1,2,\ldots \ ,$$

and all conclusions of the corollary follow in a manner analogous to Theorem 4.1.

The solutions x^* and y^* are called the <u>minimal</u> and <u>maximal</u> solutions of $Fx = 0$ in $[x_o,y_o]$. The case of most interest is when $x^* = y^*$, because then the sequences $\{y_k\}$ and $\{x_k\}$ constitute upper and lower bounds for the unique solution y^* of $Fx = 0$ in $[x_o,y_o]$. In this connection, the following uniqueness result is of interest:

<u>Lemma 4.1</u>: Let $F: D \subset R^n \to R^n$ and suppose the points $x_o,y_o \in D$ satisfy (4.1). In addition, assume that there is a mapping $C: [x_o,y_o] \to \mathcal{M}^n$ such that

(4.12) $\quad Fy - Fx \geqq C(x)(y-x) \quad , \ x_o \leqq x \leqq y \leqq y_o,$

where for all $x \in [x_o,y_o]$, $C(x)$ is non-singular and $[C(x)]^{-1} \geqq 0$. If $Fx = 0$ has either a minimal or maximal solution in $[x_o,y_o]$, then there are no other solutions in that interval.

<u>Proof</u>: Suppose $x^* \in [x_o,y_o]$ is a minimal solution and $z^* \in [x^*,y_o]$

is any other solution of $Fx = 0$. Then

$0 = F z^* - F x^* \geqq C(x^*)(z^* - x^*)$ and, because $[C(x^*)]^{-1} \geqq 0$, $z^* - x^* \leqq 0$. Hence $z^* = x^*$. The proof is similar if a maximal solution exists.

Theorem 4.1 and Corollary 4.1 are related to results of Baluev [2], [3] who essentially used, instead of (4.2), a two-sided estimate of the form

(4.14) $Fy+A_1(x,y)(z-y) \leqq Fz \leqq Fx+A_2(x,y)(z-x) \quad , x_o \leqq x \leqq z \leqq y \leqq y_o,$

and considered the iterations (4.3) and (4.9) with

$$B_k \equiv [A_1(x_k,y_k)]^{-1} \quad , \quad C_k \equiv [A_2(x_k,y_k)]^{-1} .$$

Here B_k and C_k are again assumed to be non-negative. Note that in Theorem 4.1 only the one-sided estimate (4.2) is required in order to obtain the monotonicity of the sequence $\{y_k\}$. Note also that in Baluev's setting we have to assume that $[A(y_k)]^{-1} \geqq 0$ while our use of subinverses of $A(y_k)$ is considerably more general. A related subinverse condition has also been used previously by Slugin [16].

There is also a close connection to a basic result of Kantorovich [8]; we give this result as a corollary:

<u>Corollary 4.2</u>: Let $G: D \subset R^n \to R^n$, set $Fx = x - Gx$ and suppose there exist points $x_o,y_o \in D$ which satisfy (4.1). If G is continuous and isotone on $[x_o,y_o]$, then the sequences

$$x_{k+1} = Gx_k \quad , \quad y_{k+1} = Gy_k \; , \; k=0,1,..$$

satisfy $x_k \uparrow x^*$, and $y_k \downarrow y^*$ for $k \to \infty$, where $x^* \leqq y^*$ are the minimal and maximal fixed points of G in $[x_o, y_o]$.

Proof: Since G is isotone it follows that

$$Fy - Fx = y - x-(Gy - Gx) \leqq y - x, \; x_o \leqq x \leqq y \leqq y_o \; .$$

Hence all conditions of Theorem 4.1 and Corollary 4.1 are satisfied if we take $B_k \equiv C_k \equiv B \equiv C \equiv I \equiv A(x)$.

Corollary 4.2 provides one way of obtaining the mapping A needed in (4.2). A more interesting possibility arises when F is order convex.

Corollary 4.3: For $F\colon D \subset R^n \to R^n$ let $x_o, y_o \in D$ satisfy (4.1) and suppose that F is differentiable and order convex on the interval $[x_o, y_o]$. If the matrices B_k in (4.3) are non-negative subinverses of $F'(y_k)$, then (4.4) holds. Moreover, if F' is isotone on $[x_o, y_o]$ and the matrices C_k in (4.9) are non-negative subinverses of $F'(y_k)$, then (4.10) holds.

The result follows immediately from Theorem 4.1 and Corollary 4.1 because Lemma 3.1 implies that (4.2) is satisfied with $A(x) \equiv F'(x)$. Additional assumptions such as (4.5) are again needed to insure that the limit elements are solutions of $Fx = 0$.

As a special case of Corollary 4.3 we obtain a generally known result for Newton's method which dates back at least to Baluev [2].

<u>Corollary 4.4</u>: Let $F: D \subset R^n \to R^n$ and suppose that $x_o, y_o \in D$ satisfy (4.1). Assume that $F \in C^1([x_o, y_o])$, F' is isotone on $[x_o, y_o]$, and for all $x \in [x_o, y_o]$, $F'(x)$ is nonsingular and $[F'(x)]^{-1} \geqq 0$. Then the sequences

$$(4.15) \quad y_{k+1} = y_k - (F'(y_k))^{-1} Fy_k \, , \; x_{k+1} = x_k - [F'(y_k)]^{-1} Fx_k , \quad k=0,1,\ldots$$

satisfy $x_k \uparrow y^*$, $y_k \downarrow y^*$ for $k \to \infty$, where y^* is the unique solution of $Fx = 0$ in $[x_o, y_o]$.

The proof follows immediately from Corollary 4.3 and Lemma 4.1 by making the following identifications:

$$B_k \equiv C_k \equiv (F'(y_k))^{-1} \, , \; B \equiv C \equiv (F'(y_o))^{-1} \, , \; C(x) \equiv F'(x).$$

We note that in the construction of the subsidiary sequence $\{x_k\}$, the choice of the particular subinverses $[F'(y_k)]^{-1}$ is beneficial for two reasons. First of all, if Gaussian elimination is used to solve the linear systems implied by (4.15), it requires very little additional work to obtain both x_{k+1} and y_{k+1} at the same time. Secondly, it is easy to show that if $F''(y^*)$ exists then the convergence of the interval $[x_k, y_k]$ is quadratic, i.e., $\|x_{k+1} - y_{k+1}\| \leqq c\|x_k - y_k\|^2$ under any norm on R^n. Finally we note that Vandergraft [18] has obtained a result similar to Corollary 4.4 even when $F'(x)$ is not invertible.

The following result is useful for the comparison of different iterative processes:

<u>Corollary 4.5</u>: Assume that - except for (4.5) - the conditions of Theorem 4.1 are satisfied. In addition to the sequence $\{y_k\}$ defined by (4.3) consider another sequence

$$y'_{k+1} = y'_k - B'_k\, Fy'_k \;, \; k=0,1,\ldots \;, \; y'_o = y_o,$$

where B'_k is any subinverse of $A(y'_k)$ which satisfies $B_k \geqq B'_k \geqq 0$ for all k. Then $y_k \leqq y'_k$ for $k=1,2,\ldots,$.

The proof follows by induction from

$$y'_{k+1} - y_{k+1} = y'_k - y_k + (B_k - B'_k)Fy_k - B'_k(Fy'_k - Fy_k)$$
$$\geqq [I - B'_k A(y'_k)]\;(y'_k - y_k) \geqq 0.$$

We end this section with two simple lemmas concerning the crucial condition (4.1). These results are not completely satisfying, especially in connection with the methods discussed in the next section; in general, it is a non-trivial problem to satisfy (4.1) in a simple way. For other results of this type see Section 6 and Schmidt [12].

<u>Lemma 4.2</u>: Let $F: D \subset R^n \rightarrow R^n$ be convex and differentiable on D. Assume that for some $x \in D$, $[F'(x)]^{-1}$ exists and that

$y_o = x - (F'(x))^{-1}Fx \in D$. Then $Fy_o \geqq 0$.

The proof follows immediately from $Fy_o \geqq Fx + F'(x)(y_o-x) = 0$ which in turn is a consequence of Lemma 3.1. Note that if $D = R^n$ and $[F'(x)]^{-1} \geqq 0$ for all $x \in R^n$, Lemma 4.2 together with Corollary 4.4 gives a global convergence theorem for Newton's method.

<u>Lemma 4.3</u>: Let $F: D \subset R^n \to R^n$ be order convex and differentiable in D and suppose there exists a non-negative matrix C, such that $F'(x)\, C \geqq I$ for all $x \in D$. If $Fy_o \geqq 0$ and $x_o = y_o - CFy_o \in D$, then $Fx_o \leqq 0$. Conversely, if $Fx_o \leqq 0$ and $y_o = x_o - CFx_o \in D$, then $Fy_o \geqq 0$.

<u>Proof</u>: Assume $Fy_o \geqq 0$ and $x_o \in D$; then $x_o \leqq y_o$ and by Lemma 3.1

$$Fx_o \leqq Fy_o + F'(x_o)(x_o-y_o) = [I - F'(x_o)C]\, Fy_o \leqq 0 \,.$$

Conversely, if $Fx_o \leqq 0$ and $y_o \in D$, then $y_o \geqq x_o$ and

$$Fy_o \geqq Fx_o + F'(x_o)(y_o-x_o) = [I - F'(x_o)C]\, Fx_o \geqq 0 \,.$$

5. <u>Newton-Gauss-Seidel Methods</u>

Assume that $F: D \subset R^n \to R^n$ is differentiable on D and that for each $x \in D$

$$F'(x) = D(x) - L(x) - U(x) \tag{5.1}$$

is a decomposition of the Jacobian into block-diagonal, strictly lower - , and strictly upper - block triangular matrices. We assume further that $D(x)$ is non-singular and for real ω define

$$(5.2)\quad H_\omega(x) = (D(x)-\omega L(x))^{-1}((1-\omega)D(x) + \omega U(x)) , \; x \in D.$$

For a given sequence of integers $m_k \geqq 1$, $k=0,1,...,$ define the matrix functions

$$(5.3)\quad B_k(x) = \omega(I+ \dots +H_\omega^{m_k-1}(x)) (D(x) - \omega L(x))^{-1} .$$

Then we call the iteration

$$(5.4)\quad y_{k+1} = y_k - B_k(y_k) Fy_k , \quad k=0,1,... ,$$

a <u>Newton-Gauss-Seidel process</u>. Note that this is just the formal representation of taking m_k block Gauss-Seidel iterations toward the solution of the linear system

$$F'(y_k) y = F'(y_k) y_k - Fy_k .$$

The indices m_k may be given a priori or determined a posteriori by a convergence criterion on the inner Gauss-Seidel iteration.

The special case $m_k \equiv 1$, $\omega = 1$, has been considered recently by Greenspan and Parter [7] in a particular context (see Section 6), and the following result represents an extension of their Theorem 4.3:

<u>Theorem 5.1</u>: Assume that $F: D \subset R^n \to R^n$ is continuously differentiable and order convex on $[x_o, y_o] \subset D$, where x_o and y_o

satisfy (4.1). Suppose further that for each $x \in [x_o, y_o]$, $F'(x)$ is an M-matrix. Then $Fx = 0$ has a unique solution y^* in $[x_o, y_o]$ and the sequence $\{y_k\}$, defined by (5.2) - (5.4) with $0 < \omega \leqq 1$ and an arbitrary sequence of indices $m_k \geqq 1$, satisfies

$$y_k \downarrow y^* \text{ for } k \to \infty. \tag{5.5}$$

Proof: Since for any $x \in [x_o, y_o]$, $F'(x)$ is an M-matrix, it follows that $D(x)$ is also an M-matrix; hence $[D(x)]^{-1} \geqq 0$ and

$$H_\omega(x) = [I - \omega D^{-1}(x)L(x)]^{-1}[(1-\omega)I + \omega D^{-1}(x)U(x)] \geqq 0.$$

Therefore $B_k(x) \geqq 0$, and because

$$F'(x) = \frac{1}{\omega}[D(x) - \omega L(x)] - \frac{1}{\omega}[(1-\omega)D(x) + \omega U(x)]$$

is a weak regular splitting of $F'(x)$ we have by Lemma 2.2 that $B_k(x)$ is a non-negative subinverse of $F'(x)$. Corollary 4.3 then assures that (5.5) holds. To conclude that $Fy^* = 0$, we note that the continuity of $F'(x)$ implies that $(F'(x))^{-1}$ is continuous; hence the matrix B in (4.5) can be taken equal to $[F'(y^*)]^{-1}$. The uniqueness of y^* follows from Lemma 4.1., with $C(x) \equiv F'(x)$.

From Corollary 4.3 we also obtain a result for the subsidiary sequence defined by

$$x_{k+1} = x_k - B_k(y_k)Fx_k \quad , \quad k=0,1,\ldots . \tag{5.6}$$

Corollary 5.1: Let $F: D \subset R^n \to R^n$ and suppose that $F \in C^1([x_o, y_o])$

where x_o, y_o satisfy (4.1). Assume further that F' is isotone on $[x_o, y_o]$ and that $F'(x)$ is an M-matrix for all $x \in [x_o, y_o]$. Then $x_k \uparrow y^*$ for $k \to \infty$.

Also of interest is a comparison result between different processes of the form (5.4).

<u>Corollary 5.2</u>: Assume that the conditions of Corollary 5.1 hold. Let $\{y_k'\}$ be another sequence defined by the process (5.2) - (5.4) with $0 < \omega' \leqq \omega \leqq 1$, $m_k' \leqq m_k$, $k=0,1,\ldots$, and $y_o = y_o'$. Then $y_k \leqq y_k'$ for all k.

<u>Proof</u>: Let $B_k'(x)$ be the matrix defined by (5.3) with m_k' and ω' instead of m_k and ω. Then it is easily shown that $B_k'(x) \leqq B_k(x)$ for all $x \in [x_o, y_o]$. Moreover, since F' is isotone, i.e.,

$$D(x) - L(x) - U(x) \leqq D(y) - L(y) - U(y)$$

whenever $x_o \leqq x \leqq y \leqq y_o$, it follows that $D^{-1}(y)L(y) \leqq D^{-1}(x)L(x)$ and hence

$$[I - \omega D^{-1}(y)L(y)]^{-1} \leqq [I - \omega D^{-1}(x)L(x)]^{-1}.$$

Therefore $H_\omega(y) \leqq H_\omega(x)$ and $B_k(y) \leqq B_k(x)$. Altogether then

$$B_k'(y) \leqq B_k(y) \leqq B_k(x) \qquad \text{whenever } x_o \leqq x \leqq y \leqq y_o,$$

and the result follows from Corollary 4.5, with $A(x) \equiv F'(x)$.

For the limiting case $m_k = \infty$ (k=0,1,..) we find that the

Newton iterates can be no slower than any Newton-Gauss-Seidel sequence:

Corollary 5.3: Under the conditions of Theorem 5.1 and the additional assumption that F' is isotone on $[x_o, y_o]$, the Newton iterates

$$\hat{y}_{k+1} = \hat{y}_k - (F'(\hat{y}_k))^{-1} F\hat{y}_k \quad , \quad k=0,1,\ldots$$

satisfy

$$\hat{y}_k \leqq y_k \quad , \quad k=0,1,\ldots$$

where $\{y_k\}$ is any sequence defined by (5.2) - (5.4) with $y_o = \hat{y}_o$ and $0 < \omega \leqq 1$.

The proof again follows from Corollary 4.5.

6. Application to Mildly Non-linear Equations

Let $G: D \subset R^n \to R^n$ and consider the equation

$$(6.1) \qquad Fx \equiv Ax + Gx = 0$$

where, throughout this section, A is assumed to be an M-matrix. Greenspan and Parter [7] have recently studied a special class of equations of this kind arising as discrete analogues of mildly nonlinear elliptic boundary value problems of the form

$$(6.2) \qquad \Delta u(s,t) = f(u(s,t)) \quad , \quad (s,t) \in \Omega \; , \; u = \varphi \text{ on } \partial\Omega \, .$$

We show in this section how some of the results of [7] relate to the general theory of the previous sections, and also give some extensions. The following lemmas, concerning the existence of points x_o, y_o for which $Fx_o \leqq 0 \leqq Fy_o$, are essentially contained in [7].

Lemma 6.1: Suppose there exists an $a \geqq 0$ such that $-a \leqq Gx \leqq a$ for $x \in R^n$. Set $y_o = A^{-1}a$ and $x_o = -y_o$. Then $Fx_o \leqq 0 \leqq Fy_o$.

Proof: $Fx_o = Ax_o + Gx_o \leqq -a + Gx_o \leqq 0 \leqq a + Gy_o = Fy_o$.

Lemma 6.2: Suppose $G(0) \leqq 0$ and $Gx \geqq G(0)$ for $x \geqq 0$. Set $y_o = -A^{-1}G(0)$. Then $Fy_o \geqq 0$.

Proof: $Fy_o = Ay_o + Gy_o = -G(0) + Gy_o \geqq 0$.

Lemma 6.3: Suppose $G(0)$ exists and set $y_o = A^{-1}|G(0)|$, $x_o = -y_o$. Assume that G is defined and isotone on $[x_o, y_o]$. Then $Fx_o \leqq 0 \leqq Fy_o$.

Proof: $Fx_o = -|G(0)| + Gx_o \leqq -|G(0)| + G(0) \leqq 0 \leqq |G(0)| + G(0) \leqq$
$$\leqq |G(0)| + Gy_o = Fy_o .$$

The following theorem, together with the preceeding three lemmas, contains Theorems 3.1, 3.2, and 3.3 of [7].

Theorem 6.1: Let $F: D \subset R^n \to R^n$ be defined by (6.1) and suppose there exist $x_o, y_o \in D$ such that (4.1) holds. Assume further that on $[x_o, y_o]$, G is continuous and satisfies

$$(6.3) \qquad Gy - Gx \leqq k(y-x) \ , \ x_o \leqq x \leqq y \leqq y_o,$$

with some scalar $k \geqq 0$. Finally, let C be any nonnegative nonsingular subinverse of $A+kI$. Then $Fx = 0$ has maximal and minimal solutions $y^* \geqq x^*$ in $[x_o, y_o]$ and the sequences

$$(6.4)\quad x_{k+1} = x_k - CFx_k \quad , \quad y_{k+1} = y_k - CFy_k \quad , \quad k=0,1,\ldots$$

satisfy $x_k \uparrow x^*$, $y_k \downarrow y^*$ for $k \to \infty$.

The proof follows immediately from Theorem 4.1 since

$$Fy - Fx \leqq (A+kI)(y-x) \quad , \quad x_o \leqq x \leqq y \leqq y_o \; .$$

The proof also follows directly from the Kantorovich-lemma (Corollary 4.2) since the function $x - CFx$ is isotone on $[x_o, y_o]$.

Since A is an M-matrix, we note that $A+kI$ also is an M-matrix. Hence the inverse of any matrix obtained from $A+kI$ by setting off-diagonal elements to zero represents a permissible C. The special choice $C = (A+kI)^{-1}$ was used in [7].

In the special case that (6.1) is a discretization of the form

$$(6.5)\quad \sum_{j=1}^{n} a_{ij}\xi_j + h^2[f(\xi_i) + b_i] = 0, \; i=1,\ldots,n \; , \; x=(\xi_1,\ldots,\xi_n),$$

of the boundary value problem (6.2), it follows that

$$(6.6)\qquad g_i(x) \equiv g_i(\xi_i) \quad , \; i=1,\ldots,n,$$

i.e., the ith component of G depends only on the ith variable. Then the condition (6.3) will be satisfied if we assume that

$$(6.7)\qquad |f(u) - f(v)| \leqq k(c)\; |u - v|$$

whenever $|u-v| \leqq c$. This is the basic assumption of [7]. In particular, (6.7) is satisfied whenever f is continuously differentiable. Under the condition (6.7), Lemma 6.1 together with Theorem 6.1 provides an existence result for the system (6.5) when f is bounded. Lemmas 6.2 or 6.3, together with the theorem, provide existence results when f(u) is monotone for $u \geqq 0$ (e.g., $f(u) = u^{2k}$) or monotone for all u, respectively.

Finally we note that extensions of Theorem 6.1 are possible. In particular, assume that instead of (6.3) the more general estimate

$$Gy - Gx \leqq B(y-x)$$

is satisfied. Then the theorem remains valid if we can find a non-negative non-singular subinverse C of A+B. In this case, A need not be an M-matrix. However, this leads to the unresolved question of the existence of a nonsingular, non-negative subinverse of a given matrix.

Next, we consider the application of the results of Section 5 to (6.1). We make the following assumptions:

(a) The basic interval $[x_o,y_o] \subset D$ of (4.1) exists.

(b) $G \in C^1([x_o,y_o])$ and $G'(x)$ is a non-negative diagonal matrix.

(c) G is order convex on $[x_o,y_o]$.

These conditions, together with the fact that A is an

M-matrix, imply that $F'(x)$ is an M-matrix for each $x \in [x_o, y_o]$ and that F is order convex on $[x_o, y_o]$. Hence Theorem 5.1 applies. In the context of (6.2) and (6.5), assumptions (b) and (c) are satisfied if $f'(u) \geqq 0$ and $f''(u) \geqq 0$. In this setting and for the special case $\omega = 1$ and $m_k = 1$, our result is then equivalent with Theorem 4.3 of [7].

It is possible to extend these results to boundary value problems of the form

$$(6.7) \qquad \Delta u = f(u, u_s, u_t) \quad , \quad (s,t) \in \Omega \ , \ u = \varphi \quad \text{on } \partial\Omega \ .$$

Assume that $f = f(u,p,q)$ is a convex differentiable function defined on all of R^3, and that

$$(6.8) \qquad f_u \geqq 0 \ , \ |f_p|, \ |f_q| \leqq m.$$

For simplicity, assume further that $\Omega = [0,1] \times [0,1]$ and that the left side of (6.7) is discretized by means of the usual five-point formula while the right side is discretized by

$$f(u(s,t) \ , \ \frac{u(s+h,t)-u(s-h,t)}{2h} \ , \ \frac{u(s,t+h) - u(s,t-h)}{2h}) \ .$$

Then the ith component of the operator F of (6.1) has the general form

$$(6.9) \quad f_i(x) = \sum_{j=1}^{n} a_{ij}\xi_j + h^2 \cdot f(\xi_i, \frac{1}{h} \sum_{j=1}^{n} \alpha_{ij}\xi_j, \frac{1}{h} \sum_{j=1}^{n} \beta_{ij}\xi_j) + h^2 b_i,$$

where the α_{ij} and β_{ij} are $\pm 1/2$ and $\alpha_{ij} = \beta_{ij} = 0$ unless $a_{ij} \neq 0$.

Therefore, we have

$$\frac{\partial f_i}{\partial \xi_i} = a_{ii} + h^2 f_u$$

$$\frac{\partial f_i}{\partial \xi_j} = a_{ij} + h\,\alpha_{ij} f_p + h\,\beta_{ij} f_q \quad , i \neq j$$

and from (6.8) it follows that for sufficiently small h, F'(x) has non-positive off-diagonal elements and positive diagonal elements. Moreover, because of cancellations in forming the sum $\sum_{j=1}^{n} \frac{\partial f_i}{\partial \xi_j}$, it is easily seen that F'(x) inherits from A the property of being irreducibly diagonally dominant. Hence, for each $x \in [x_o, y_o]$, F'(x) is an M-matrix. Finally, because of the convexity of f, Lemma 3.3 implies that F is also convex. Theorem 5.1 now applies as well as Corollary 5.1.

We note that from the computational viewpoint the condition $0 < \omega \leqq 1$ represents a severe restriction for the Newton-Gauss-Seidel methods, especially when these methods are applied to systems of the form (6.5) or (6.9). For similar methods, it is shown in [9] that the optimum ω for (6.5) is, roughly speaking, about that of the corresponding linear problem. Hence $\omega > 1$ will in general be necessary for faster convergence and in this case the results of this section do not apply. However, it still is possible that monotone convergence may be preserved in the initial stages of the iteration, but a convergence theory will, of course, require a different approach than used here.

7. An Implicit Theorem and the Nonlinear Gauss-Seidel Method

In conclusion we discuss a modification of Theorem 4.1 for implicit iterations of the form

$$(7.1) \qquad G(y_{k+1}, y_k) = 0 \quad , k=0,1,\ldots .$$

Theorem 7.1: Let $G: D \times D \subset R^n \times R^n \to R^n$ and suppose that $x_o, y_o \in D$ are such that $x_o \leq y_o$, $[x_o, y_o] \subset D$, and

$$(7.2) \qquad G(x_o, x_o) \leq 0 \leq G(y_o, y_o).$$

Assume there exist mappings $A: D \times D \to \mathfrak{M}^n$ and $B: D \to \mathfrak{M}^n$ for which

$$(7.3) \quad [A(x,y)]^{-1} \geq 0 \quad , B(x) \leq 0 \quad \text{for all } x,y \in D,$$

$$(7.4) \quad G(x,y) - G(z,y) \geq A(z,y)(x-z) \quad \text{for all } x,y,z \in D,$$

and

$$(7.5) \quad G(x,x) - G(x,y) \geq B(y)(x-y) \quad \text{for } x_o \leq x \leq y \leq y_o .$$

Suppose finally that the sequence $\{y_k\} \in D$ satisfies (7.1). Then $y_k \downarrow y^* \in [x_o, y_o]$, and if G is continuous at (y^*, y^*), then $G(y^*, y^*) = 0$.

Proof: By (7.2), (7.4) and (7.1) we have

$$0 \geq G(y_1, y_o) - G(y_o, y_o) \geq A(y_o, y_o)(y_1 - y_o)$$

and hence, by (7.4), $y_1 - y_o \leq 0$. Similarly, using (7.2), (7.3), and (7.5) and then (7.4),

$$0 \geqq G(x_o,x_o) \geqq G(x_o,y_o) + B(y_o)(x_o-y_o) \geqq G(x_o,y_o)$$
$$\geqq G(y_1,y_o) + A(y_1,y_o)(x_o-y_1) = A(y_1,y_o)(x_o-y_1),$$

so that by (7.3), $x_o - y_1 \leqq 0$. Finally, (7.1) and (7.3) imply that

$$G(y_1,y_1) \geqq G(y_1,y_o) + B(y_o)(y_1-y_o) \geqq 0.$$

The conclusions of the theorem now follow by induction.

Theorem 7.1 has immediate application to explicit iterative processes of the form $y_{k+1} = Hy_k$, where $H: D \subset R^n \to R^n$ is some nonlinear operator. In this case we can take $A \equiv I$. It is more interesting, however, when $G(x,y)$ is nonlinear in x as well as y.

Assume that $F: R^n \to R^n$ has components f_i which are defined on the entire space R^n. We define the components g_i of G by

$$g_i(x,y) \equiv f_i(q_i(x,y)) \quad , \; i=1,\dots,n, \quad x,y \in R^n,$$

where the mappings $q_i: R^n \times R^n \to R^n$ are given in terms of the components ξ_i of x and η_i of y by

$$q_i(x,y) = (\xi_1,\dots,\xi_i, \eta_{i+1},\dots,\eta_n) \quad , \; i=1,\dots,n.$$

Then (7.1) is the nonlinear Gauss-Seidel process studied by Bers [5] and Schechter [10]. (See also [9].)

To apply Theorem 7.1, we make the following assumptions about F:

(a) $F'(x)$ exists and is an M-matrix for each $x \in R^n$.

(b) F is continuous and convex on R^n.

(c) F' is isotone on R^n.

(d) For each $y \in R^n$ there exists an $x \in R^n$ such that $G(x,y) = 0$.

For example, in the case of the system (6.5) belonging to the boundary value problem (6.2) all these conditions are satisfied if the matrix (a_{ij}) is an M-matrix and if $f'(t) \geqq 0$, $f''(t) \geqq 0$ for $-\infty < t < +\infty$. We also assume, as usual, that (4.1) is satisifed, i.e., that there exist $x_o, y_o \in R^n$ for which $x_o \leqq y_o$ and $Fx_o \leqq 0 \leqq Fy_o$; this implies that (7.2) holds.

Next we introduce the n x n matrices

$$G_x(x,y) \equiv \left(\frac{\partial g_i}{\partial \xi_j}(x,y) \right) , \quad G_y(x,y) \equiv \left(\frac{\partial g_i}{\partial \eta_j}(x,y) \right).$$

Then it is easy to verify that

$$\frac{\partial g_i}{\partial \xi_j}(x,y) = \begin{cases} \frac{\partial f_i}{\partial \xi_j}(q_i(x,y)) & \text{for } i \geqq j , \\ 0 & \text{for } i < j , \end{cases}$$

and

$$\frac{\partial g_i}{\partial \eta_j}(x,y) = \begin{cases} 0 & \text{for } i \geqq j , \\ \frac{\partial f_i}{\partial \xi_j}(q_i(x,y)) & \text{for } i < j . \end{cases}$$

Hence it follows from (a) that $[G_x(x,y)]^{-1} \geqq 0$ and $G_y(x,y) \leqq 0$ for all $x,y \in R^n$. Moreover, using (b) and Lemma 3.1, an easy computation shows that

$$G(x,y) - G(z,y) \geqq G_x(z,y)(x-z) \quad \text{for all } x,y,z \in R^n,$$

and similarly, using (c), that

$$G(x,x) - G(x,y) \geqq G_y(y,y)\ (x-y) \quad \text{for } x_o \leqq x \leqq y \leqq y_o.$$

Thus conditions (7.3) - (7.5) are all satisfied with $A(x,y) \equiv G_x(x,y)$ and $B(x) \equiv G_y(x,x)$. The assumption (d) assures that the sequence $\{y_k\}$ of (7.1) exists; hence Theorem 7.1 applies and we have $y_k \downarrow y^* \in [x_o, y_o]$. The continuity of F implies that of G and therefore $Fy^* = G(y^*,y^*) = 0$.

References

[1] J. Albrecht, Fehlerschranken und Konvergenzbeschleunigung bei einer monotonen oder alternierenden Iterationsfolge, Numer. Math., 4 (1962), pp. 196-208.

[2] A. Baluev, On the method of Chaplygin, Dokl. Akad. Nauk SSSR, 83 (1952), pp. 781-784.

[3] A. Baluev, On the method of Chaplygin for functional equations, Vestnik Leningrad Univ. Ser. Mat. Meh. Astron., 13 (1956), pp. 27-42.

[4] I. Berezin and N. Zhidkov, Computing Methods, Fizmatgiz, Moscow, 1959, translated by O. Blunn and A. Booth, Pergamon Press, New York, 1965.

[5] L. Bers, On mildly non-linear partial difference equations of elliptic type, J. Res. Nat. Bur. Standards, 51 (1953), pp. 229-236.

[6] L. Collatz, Funktionalanalysis und Numerische Mathematik, Springer-Verlag, Berlin, 1964.

[7] D. Greenspan and S. Parter, Mildly nonlinear elliptic partial differential equations and their numerical solution II, Numer. Math., 7 (1965), pp. 129-147.

[8] L. Kantorovich, The method of successive approximations for functional equations, Acta Math., 71 (1939), pp. 63-97.

[9] J. Ortega and M. Rockoff, Nonlinear difference equations and Gauss-Seidel type iterative methods, SIAM J. Num. Anal., (to appear).

[10] S. Schechter, Iteration methods for nonlinear problems, Trans. Amer. Math. Soc., 104 (1962), pp. 179-189.

[11] J. Schmidt, Konvergenzbeschleunigung bei monotonen Vektorfolgen, Acta Math. Acad. Sc. Hungaricae, 1/2, 16 (1965), pp. 221-229.

[12] J. Schmidt, Ausgangsvektoren fuer monotone Iterationen, Numer. Math., 6 (1964), pp. 78-88.

[13] J. Schroeder, Anwendung von Fixpunktsaetzen bei der numerischen Behandlung nichtlinearer Gleichungen in halbgeordneten Raeumen, Arch. Rat. Mech. Anal., 5 (1959), pp. 177-192.

[14] S. Slugin, Approximate solution of operator equations by the method of Chaplygin, Dokl. Akad. Nauk SSSR, 103 (1955), pp. 565-568.

[15] S. Slugin, A modification of the abstract analogue to Chaplygin's method, Dokl. Akad. Nauk SSSR, 120 (1958), pp. 256-258.

[16] S. Slugin, On the theory of Newton's method and Chaplygin's method, Dokl. Akad. Nauk SSSR, 120 (1958), pp. 472-474.

[17] S. Slugin, Some applications of methods of two-sided approximations, Izv. Vyss. Ucebn. Zaved Mat., 6(7) (1958), pp. 244-256.

[18] J. Vandergraft, Partially ordered spaces and Newton's method for convex operators, Computer Science Center, University of Maryland, Technical Report, TR-66-29, May 1966.

[19] R. Varga, Matrix Iterative Analysis, Prentice Hall, Englewood Cliffs, New Jersey, 1962.