CS 88

RELAXATION METHODS FOR CONVEX PROBLEMS

BY

SAMUEL SCHECHTER

TECHNICAL REPORT NO. CS 88
FEBRUARY 16, 1968

COMPUTER SC IENCE DEPARTMENT
School of Humanities and Sciences
STANFORD UNIVERS ITY

RELAXATION METHODS FOR CONVEX PROBLEMS*

by

Samuel Schechter[1]

* This work was supported in part by N.S.F. GP5962 and ONR 225(37) at Stanford University, and in part by U.S. Atomic Energy Commission Contract AT(30-1)-1480 at New York University.

1 Stanford Research Institute and Stanford University.

1. Introduction. In [10] it was shown that a freely ordered relaxation process or, in particular, a Gauss-Seidel type of successive "over-relaxation" method converges for certain nonlinear problems. We will show below that this process may be extended to group (or block) relaxation. In its extreme form this becomes a modified form of Newton's method in n dimensions.

We obtain, moreover, a less restrictive choice of the relaxation parameters than that given in [10]. It is also shown that the residually ordered processes given in [11] for linear equations can be extended to this class of nonlinear problems. Here one obtains an estimate for the error, as in the linear case. A special form of this method was outlined without proof by Householder [6, p. 134].

.A proof is also given for a cyclic process (sometimes referred to in the scalar case as "nonlinear overrelaxation" [11]) which is simpler than that given for the freely ordered.process.

Some related work is given in [8] and other results in the direction of finding asymptotic convergence rates may be found in [7]. These methods are usually applied to the solution of large systems arising from finite difference approximations of nonlinear elliptic equations as shown in [10]. Such applications go back at least ten years (see, for example, [4] and [5]). Some more recent applications are given in [1], [2],[3], and [9].

2. Definitions. Let $G(u) \epsilon C^2(R^n)$ be a real valued function, twice continuously differentiable over the whole Euclidean n space R^n. We seek a global minimum of $G(u)$, that is, a solution u^* of

$$(2.1) \qquad r(u) \equiv \text{grad } G(u) = 0$$

where $r(u) = (r_1(u), \ldots, r_n(u))^T$, $u = (u_1, \ldots, u_n)^T$, $r_i(u) = G_{u_i}(u) = \frac{\partial}{\partial u_i} G(u)$. Let $A(u) = (a_{ij}(u)) = (G_{u_i u_j}(u))$ denote the n by n Hessian matrix of G; $\lambda(A)$ and $\Lambda(A)$ will denote the minimum and maximum eignevalues of a symmetric matrix A, respectively. For a column vector u we write $|u|^2 = (u,u) = u^T u$ and let $\|r\|_D = \sup_{u \epsilon D} |r(u)|$. Write $A > 0$ (≥ 0) when A is a positive definite (semidefinite) matrix, and $A > \delta$ means $A - \delta I > 0$ for the identity I of order n.

Let $Z = (1,2,\ldots,n)$ and call $g = (i_1, i_2, \ldots, i_k)$ a multi-index of order $k \leq n$ if $1 \leq i_1 < i_2 < \ldots < i_k \leq n$. Let $g' = Z - g$ be the multi-index of order $n - k$ remaining in Z when g is removed. Denote the set of all multi-indices of order k by Q_{kn} and let $Q_n = \bigcup_{k=1}^n Q_{kn}$.

Any sequence $\{g_p\}_{p=0}^{\infty}$, $g_p \epsilon Q_n$ will be called an <u>ordering</u>. An ordering covers Z infinitely often if, for each $i \epsilon Z$, $i \epsilon g_p$ for infinitely many p; we then say that it is <u>freely ordered</u>.

We use the notation of [11] for subvectors and submatrices. That is, if $g \epsilon Q_{kn}$ then u_g is a subvector of u of dimension k: $(u_g)_\nu = u_{i_\nu}$ where $i_\nu \epsilon g$. Similarly, if $h \epsilon Q_{mn}$ then A_{gh} denotes the k by m submatrix of A whose (ν,μ) element is $a_{i_\nu j_\mu}$, $i_\nu \epsilon g$, $j_\mu \epsilon h$. If $g=h$ then A_{gg} is a principal submatrix of A, and let $\lambda_g = \lambda(A_{gg})$, $\Lambda_g = \Lambda(A_{gg})$ for any $g \epsilon Q_n$.

For any ordering we denote by $S = (h_1, \ldots, h_t)$ the set of different multi-indices that appear in the ordering and is called the minimal set of theordering. If the ordering covers Z then so does S.

3. Relaxation process. Given an ordering $\{g_p\}$ and an initial vector u^0 we may define, for a given sequence of numbers $\{\omega_p\}$, the iteration

$$(3.1) \qquad u_g^{p+1} = u_g^p + \omega_p d_p \ , \quad u_{g'}^{p+1} = u_{g'}^p,$$

where $d_p = -A_{gg}^{-1}(u^p) r_g(u^p)$ and $g = g_p$, providing the inverses exist. We call (3.1) a relaxation process with ordering $\{g_p\}$. The ω_p are called relaxation parameters.

This process is well known for linear problems, especially when the g_p are of order one, and has been studied extensively. It is sometimes called a group or block relaxation process, with the g_p indicating the "groups". For nonlinear problems, (3.1) was treated in [10] for freely ordered processes where each g_p was of order one (a scalar process). We will show here that for various orderings (3.1) will converge to a solution of (2.1) for a suitably restricted $G(u)$, $\{\omega_p\}$ and u^0. These conditions are found to be met by many nonlinear elliptic problems, as shown in [10].

4. Basic Lemmas. We assume henceforth that $G(u) \in C^2(R^n)$ and satisfies

$$(4.1) \qquad A(u) > 0 \quad \text{for all} \quad u \in R^n$$

so that (3.1) is defined. This also implies uniqueness of u^* as shown in [10]. For a given iterate u^p and index $g = g_p$ of (3.1) we define, for any $v \in R^n$:

$$\varphi_p(v) = (d_p, A_{gg}(v)d_p)/(d_p,d_p), \quad d_p \neq 0$$

(4.2)
$$D_p = \{u \mid G(u) \leq G(u^p)\}$$
$$\lambda_g^{(p)} = \min_{u \in D_p} \lambda(A_{gg}(u)), \quad \lambda^{(p)} = \lambda_Z^{(p)}$$
$$\rho_p = \|r_g\|_{D_p}, \quad \sigma_p = 2\rho_p/\lambda_g^{(p)}$$

whenever D_p is bounded.

For a given g let B_g be the closed unit ball $|v| \leq 1$ in the subspace R^g, which is the set of $v \in R^n$ such that $v_k = 0$, $k \in g'$. For $g = g_p$ let $D^p = u^p + 2|d_p|B_g = \{w \mid |w-u^p| \leq 2|d_p|, w_{g'} = u^p_{g'}\}$. When D_p is bounded we define

$$\Lambda_g^{(p)} = \|\Lambda(A_{gg}(u)\|_{D^p}, \quad \Lambda^{(p)} = \Lambda_Z^{(p)}$$

$$\lambda_S^{(p)} = \min \{\lambda(A_{hh}(u)) \mid u \in D_p, h \in S\}$$

(4.3)
$$\Lambda_S^{(p)} = \max \{\Lambda(A_{hh}(u)) \mid u \in D^p, h \in S\}$$

$$K_p = D_p + \sigma_p B_g$$

$$\|\varphi_p'\|_{D_p, g} = \max \{\varphi_p(v) \mid v \in D_p, v_{g'} = u^p_{g'}\}$$

and let

(4.4)
$$\gamma_p = \varphi_p(u^p)/\|\varphi_p\|_{D^p} \text{ if } d_p \neq 0,$$

but if $d_p = 0$, set $\gamma_p = 1$.

It then follows that

$$(4.5) \qquad \frac{\lambda_S^{(p)}}{\Lambda_S^{(p)}} \leq \frac{\lambda_g^{(p)}}{\Lambda_g^{(p)}} \leq \frac{\varphi_p(u^p)}{\|\varphi_p\|_{K_p}} \leq \gamma_p \leq 1.$$

For the special case when $g_0 = Z$ we write

$$\rho^* = \|r_Z\|_{D_0} = \|r\|_{D_0}, \quad \sigma^* = 2\rho^*/\lambda^{(0)}, \quad K^* = D_0 + \sigma^* B_Z$$

where now B_Z is the full unit ball in R^n. If $\{g_p\}$ is an arbitrary ordering (with g_0, in particular, any multi-index) we let Λ^* be the number obtained by replacing D^0 in the definition (4.3) of $\Lambda_S^{(0)}$, by K^*. Let $\gamma^* = \lambda_S^{(0)}/\Lambda^*$; then this constant depends only on u^0 and the minimal set of the ordering. From (4.5) it follows that $\gamma^* \leq \gamma_0$.

We will show that for a suitable choice of u^0 and ω_p the relaxation process will be well defined and the $G(u^p)$ (and the D_p) will be nonincreasing as $p \to \infty$.

<u>Lemma</u> 4.1. <u>Let</u> $u^0 \epsilon R^n$ <u>be such that</u>

$$(4.6) \qquad D_0 \text{ is bounded.}$$

<u>Let</u> $g = g_0$ <u>be any multi-index in</u> Q_n <u>and let</u> γ <u>be a constant such that</u> $0 < \gamma < \gamma^* \leq \gamma_0 \leq 1$. If ω_0 <u>is chosen in the interval</u>

$$(4.7) \qquad 0 < \gamma \leq \omega_0 \leq 2\gamma_0 - \gamma < 2$$

<u>then, for</u> u^1 <u>defined by</u> (3.1),

$$(4.8) \qquad -\Delta G_0 \equiv G(u^0) - G(u^1) \geq \epsilon_0 |r_g(u^0)|^2 > 0$$

where $$\epsilon_o = \omega_0(\zeta_0 + \tfrac{1}{2}\gamma)/\Lambda_g^{(0)} \geq 0,$$

$$\zeta_0 = 1-(||\varphi_0||_{D_\theta,g})/(||\varphi_0||_{D^0}) \geq 0$$

and $$u^1 \epsilon D_0 \supset D_1,\ D^1 \subset K^*,\ \gamma < \gamma^* \leq \gamma_1.$$

Proof. Let $d_0 \neq 0$ and let $I(u^\theta, u^1)$ denote the open line segment joining u^θ and u^1. Then Taylor's theorem in n dimensions gives us

$$G(u^1) - G(u^0) = (r(u^0), u^1 - u^0) + \tfrac{1}{2}(u^1 - u^0, A(z)(u^1 - u^0)) \tag{4.9}$$

for some $z \epsilon I(u^0, u^1)$. From (3.1), (4.2), and (4.7) we get that

$$|u^1 - u^0| \leq 2|d_0| \leq 2\rho_0/\lambda_g^{(0)} = \sigma_0.$$

Since u^1 and u^θ differ by a vector in Rg, $u^1 \epsilon D^0$ and therefore $z\epsilon D^\theta$.

From (4.9) we get

$$-\Delta G_0 = \omega_0((A_{gg}(u^0)d_0, d_0) - \tfrac{1}{2}\omega_0(d_o, A_{gg}(z)d_0))$$

$$= \tfrac{1}{2}\omega_0(d_0, d_0)(2\varphi(u^0) - \omega_0\varphi_0(z)).$$

Since $z_{g'} = u^\theta_{g'}$, we get from (4.3) and (4.7) that

$$-\Delta G_0 \geq \tfrac{1}{2}\gamma\omega_0(d_0, d_0)||\varphi_0||_{D^0} \geq 0.$$

Thus $u^1 \epsilon D_0$ and also $z\epsilon D_0$. We may then estimate further from

$$\omega_o\varphi_o(z) \leq (2\gamma_0 - \gamma)||\varphi_0||_{D_0,g} = 2\varphi_0(u^0)(1 - \zeta_0) - \gamma||\varphi_0||_{D_0,g}$$

$$-\Delta G_0 \geq \omega_0(\zeta_0\varphi_0(u^0) + \tfrac{1}{2}\gamma||\varphi_0||_{D_0,g})(d_0, A_{gg}(u^0)d_0)/\varphi_0(u^0)$$

$$\geq \omega_0(\zeta_0 + \tfrac{1}{2}\gamma)(r_g(u^0), r_g(u^0))/\Lambda_g^{(0)} = \epsilon_0 |r_g(u^0)|^2$$

$$\geq \tfrac{1}{2}\gamma^2 |r_g(u^0)|^2/\Lambda^*.$$

If $d_0 = 0$ then $r_g(u^0) = 0$ and the lemma is valid. Thus from $D_0 \supset D_1$, we get that $\lambda_{g_1}^{(1)} \geq \lambda_{g_1}^{(0)} \geq \lambda_S^{(0)} \geq \lambda^{(0)}$ and that $\rho_1 \leq \rho^*$ or $\sigma_1 \leq \sigma^*$. This implies that $D^1 \subset K^*$ and $\Lambda_{g_1}^{(1)} \leq \Lambda^*$. From (4.5) it follows that $\gamma_1 \geq \gamma^*$ which completes the proof.

Lemma 4.2. Let an ordering $\{g_p\}$ be given and let $u^0, \gamma, \gamma^*, \gamma_0$ satisfy the hypotheses of Lemma 4.1. Then there exist $\{\omega_p\}$ satisfying

$$0 < \gamma \leq \omega_p \leq 2\gamma_p - \gamma, \quad \gamma < \gamma^* \leq \gamma_p \tag{4.10}$$

such that the iterates $\{u^p\}$ of the relaxation process (3.1) satisfy

$$-\Delta G_p \equiv G(u^p) - G(u^{p+1}) \geq \epsilon_p |r_g(u^p)|^2 \tag{4.11}$$

where $g = g_p$;

$$\epsilon_p = \omega_p(\zeta_p + \tfrac{1}{2}\gamma)/\Lambda_g^{(p)} \geq \tfrac{1}{2}\gamma^2/\Lambda^* \equiv \epsilon^* > 0$$

$$\zeta_p = 1 - (\|\varphi_p\|_{D_p,g})/(\|\varphi_p\|_{D^p}) \geq 0,$$

for $p = 0, 1, 2, \ldots$.

Proof. The proof follows by induction by using Lemma 4.1 as the initial and inductive step.

Corollary 4.1. Under the hypotheses of Lemmas 4.1 and 4.2 it follows that for any ordering $\{g_p\}$, $r_{g_p}(u^p) \to 0$ as $p \to \infty$.

Proof. This follows from the fact that all the iterates u^p lie in D_0 so that $\{G(u^p)\}$ is a sequence bounded from below. Since these.

are monotone nonincreasing with p, $G(u^p) \to G_\infty$, which implies that $\Delta G_p \to 0$. The result then follows from Lemma 4.2.

Remarks. We note from the proofs that Lemmas 4.1, 4.2, and the corollary are valid even if we only assume $A(u) \geq 0$ but require that $A_{gg}(u) > 0$ for all u and all $g = g_p \epsilon S$, and replace $\lambda^{(0)}$ by $\lambda_S^{(0)}$ in σ^*.

For the scalar case we get a simple form for γ_p:

$$\gamma_p = a_{ii}(u^p)/\|a_{ii}\|_{D^p}, \quad i=i_p,$$

where

$$\|a_{ii}\|_{D^p} = \max \{a_{ii}(u_1^p, u_2^p, \ldots, u_{i-1}^p, u_i, u_{i+1}^p, \ldots, u_n^p) | u_i \epsilon I_p\}$$

$$I_p = \{u_i | |u_i - u_i^p| \leq 2|d_p|\}.$$

In [10] it was shown that for a free ordering with scalar indices the relaxation process converged for a choice of γ_p which was some fixed constant less than γ^*. We will show below that the relaxation process converges in the more general case of (4.7) for a free ordering. Since the cyclic orderings are more important and easier to prove, we give first a proof of their convergence.

5. Cyclic Orderings.. We assume that a finite set of t multi-indices $S = \{h_i\}_{i=1}^t$, $h_i \epsilon Q_n$, is given such that $\bigcup_{i=1}^t h_i \supset Z$. If a sequence $\{g_p\}$ runs through the list S in a cyclic fashion, i.e.,

$$g_p = h_{p(\text{mod}\, t)+1}, \quad p = 0,1,2,\ldots \tag{5.1}$$

then we say that the ordering is cyclic with S as minimal set.

Theorem 5.1. Let $G(u)$ and u^0 satisfy (4.1) and (4.6) and let $\{g_p\}$ be a cyclic ordering with minimal set S'. Then, if the $\{\omega_p\}$ satisfy (4.10), the $\{u^p\}$ (3.1) converge to the solution u^* of (2.1).

Proof. From Corollary 4.1 we get that $r_{g_p}(u^p) \to 0$ as $p \to \infty$ and that $G(u^p) \to G_\infty$. It follows from Lemma 4.2 that for any $p < q$

$$(5.2) \qquad G(u^p) - G(u^q) = -\Sigma_{\nu=p}^{q-1}\Delta G_\nu \geq \epsilon^* \Sigma_{\nu=p}^{q-1} |r_{g_\nu}(u^\nu)|^2 .$$

This implies that for all p,q such that $|p - q| \leq t$,

$$\Sigma_{\nu=p}^{q-1} |r_{g_\nu}(u^\nu)| \to 0.$$

Furthermore, since $G \epsilon C^2(D_0)$ there exists a constant M depending only on, u^0 and G such that for any i, $1 \leq i \leq t$,

$$|r_{h_i}(u^{\nu+1}) - r_{h_i}(u^\nu)| \leq M|u^{\nu+1} - u^\nu| = M|u^{\nu+1}_{g_\nu} - u^\nu_{g_\nu}|$$

$$\leq M_1 |r_{g_\nu}(u^\nu)|$$

where $M_1 = 2M/\lambda_S^{(0)}$. This implies that the left side of

$$|r_{h_i}(u^q) - r_{h_i}(u^p)| \leq \Sigma_{\nu=p}^{q-1} |r_{h_i}(u^{\nu+1}) - r_{h_i}(u^\nu)|$$

$$\leq M_1 \Sigma_{\nu=p}^{q-1} |r_{g_\nu}(u^\nu)|$$

goes to zero for $|p-q| \leq t$ as p and $q \to \infty$. For i fixed and any $p > 0$ set $q = [\frac{p}{t}]t + i - 1$ (where $[\frac{p}{t}]$ is the greatest integer contained in p/t), then $|p-q| \leq t$ while $g_q = h_i$. Thus $r_{h_i}(u^q) = r_{g_q}(u^q)$ which goes to zero as $p \to \infty$, whence $r_{h_i}(u^p) \to 0$ and

$r(u^p) \to 0$ as $p \to \infty$.

This implies that every limit point of $\{u^p\}$ is a stationary point of $G(u)$, and since D_0 is bounded there is at least one limit point. It follows, however, as in [10], that there is at most one stationary point u^*, so that $u^p \to u^*$ and the proof is complete.

Corollary 5.1. Let $G(u)$ and u^0 satisfy (4.1) and (4.6), ω_p satisfy (4.10), then a modified Newton's method:

$$(5.2) \qquad u^{p+1} = u^p - \omega_p A^{-1}(u^p) r(u^p)$$

converges to the solution u^* of (2.1).

Proof. This follows from Theorem 5.1 by taking $t = 1$ and S to consist of the set Z.

We will see in the next section that we can estimate the convergence rate of (5.2).

6. Residually Ordered Processes. We will show that the basic lemmas of Section 4 may be used to obtain an extension of Theorem 1 of [11]. A residually ordered process (r.o.p.) may be defined in the same way as in [11], as follows:

Let $\pi_p = (g_1^{(p)}, \dots, g_{N_p}^{(p)}), N_p \leq N \leq n, g_\kappa^{(p)} \in Q_n$ be a given sequence of coverings of Z and $\{\|*\|_p\}$ a given sequence of norms on R^n. Assume further that there exist positive constants η_p, τ_p, τ that satisfy, for any $w \in R^n$,

$$\tau_p \eta_p |w|^2 \leq \|w\|_p \leq \eta_p |w|^2, \quad p = 0,1,2,\dots,$$

$$0 < \tau_p \eta_p \leq \eta_p, \quad 0 < \tau \leq \tau_p \leq 1.$$

A relaxation process whose ordering $\{g_p\}$ is given by the multi-index g_p such that

$$\|r_{g_p}\|_p = \max_{h\in\pi_p} \|r_h\|_p$$

is called an r.o.p. For this process we prove

Theorem 6.1. Let $G(u)$ and u^0 satisfy (4.1) and (4.6), then if the $\{\omega_p\}$ satisfy (4.10), the r.o.p. converges to the solution u^* of (2.1). The iterates converge like a geometric series; that is, there exist positive constants θ, α such that

$$(6.1)\qquad |u^p - u^*|^2 \le \theta\alpha^p |u^0 - u^*|^2, \quad 0 \le \alpha < 1.$$

Proof. From Lemmas 4.1 and 4.2 we obtain

$$(6.2)\qquad \begin{aligned} -\Delta G_p &\ge \epsilon_p |r_{g_p}(u^p)|^2 \ge \epsilon_p \|r_{g_p}(u^p)\|_p^2/\eta_p \\ &\ge (\epsilon_p/N_p\eta_p)\Sigma_{h\epsilon\pi_p} \|r_h(u^p)\|_p^2 \\ &\ge (\epsilon_p\tau_p/N_p)\,|r(u^p)|^2 \ge (\epsilon^*\tau/N)\,|r(u^p)|^2 \end{aligned}$$

Thus $r(u^p) \to 0$ as $p \to \infty$ and, as in the cyclic case, it follows that $u^p \to u^*$.

To show that (6.1) holds, we set $e^p = u^p - u^*$. From (4.9)

$$V_p \equiv G(u^p) - G(u^*) = \tfrac{1}{2}(e^p, A(v)e^v), \quad v\epsilon I(u^p, u^*)$$

On the other hand, there is a $z\epsilon I(u^p, u^*)$ such that

$$(r(u^p), e^p) - V_p = \tfrac{1}{2}(e^p, A(z)e^p).$$

If $e_p \ne 0$ we set

$$\begin{aligned} \mu_p &= 1 + \min_{y,w\epsilon D_p}\left((e^p, A(y)e^p)/(e^p, A(w)e^p) \right) \\ &\ge 1 + \lambda^{(0)}/\Lambda^* = \mu \end{aligned}$$

and then

$$\begin{aligned} |r(u^p)|\,|e^p| &\ge \mu_p V_p, \\ |r(u^p)|^2 &\ge \tfrac{1}{2}\mu_p\lambda^{(p)} V_p \end{aligned}$$

If $e_p = 0$, set $\mu_p = 2$. From (6.2) we get that

$$-\Delta G_p = V_p - V_{p+1} \geq \beta_p V_p \geq \beta V_p$$

where $\beta_p = A^{(p)} \mu_p^2 \epsilon_p \tau_p / 2N_p$, $\beta = \lambda^{(0)} \mu^2 \epsilon^* \tau / 2N \leq \beta_p$,

so that $V_{p+1} \leq (1 - \beta_p) V_p \leq (1 - \beta) V_p$.

Since $\zeta_p \leq 1 - \gamma_p$, then $\epsilon_p \leq \omega_p (2 - \omega_p)/2\Lambda_{g_p}^{(P)}$ and

$$0 < \beta \leq \beta_p \leq \omega_p (2 - \omega_p) \tau_p \lambda^{(P)} / N_p \Lambda_{g_p}^{(P)} < 1$$

if $N_p > 1$. If $N_p = 1$, then $\beta_p \leq 1$.

Setting $\alpha = 1 - \beta$, $\theta = \Lambda^* / \lambda^{(0)}$ we get

$$V_p \leq \alpha^p V_0$$

or that $$|e^p|^2 < \theta \alpha^p |e^0|^2$$

which proves the theorem.

Corollary 6.1. Under the hypotheses of Corollary 5.1 the modified Newton's Method (5.2) converges like a geometric series.

Proof. This follows from Theorem 6.1 since for all p, π_p consists of the single multi-index Z and is automatically an r.o.p.

7. -Free Orderings. In [10] it was shown that for the scalar case, convergence is obtained for free orderings, that is, where a sequence $\{i_p\}$ is arbitrary but all indices of Z appear infinitely often. On the other hand, this was proved for group relaxation for linear problems in [11]. We will now combine these two results into one, in which the less stringent condition on ω_p as given by (4.10) is used.

Theorem 7.1 Let $G(u)$ and u^0 satisfy (4.1) and (4.6). Let $\{g_p\}$ be freely ordered; then if $\{\omega_p\}$ satisfy (4.7), the relaxation process (3.1) converges to the solution u^* of (2.1).

Proof. The idea of the proof is similar to that used in Theorem 3.1 of [10]. From Lemma 4.2 and Corollary 4.1 we get that $r_{g_p}(u^p) \to 0$ as $p \to \infty$.

Let x be a limit point of the sequence $\{u^p\}$. We may assume that $r(x) \neq 0$, otherwise we get convergences as before. Let S be the minimal set of the ordering and set

$$\rho = \min\{|r_g(x)| \, | \, r_g(x) \neq 0, g \in S\}.$$

Let ν be the maximal order of the multi-indices of S and let λ, Λ be positive constants such that

$$\lambda(w,w) \leq (w,A(u)w) \leq \Lambda(w,w)$$

for all $u \in D_0$, and all $w \in R^n$.

Define U to be the neighborhood of x such that

$$|u - x| < \delta, \quad \delta = \gamma\rho/2\Lambda\sqrt{\nu}$$

and let N be sufficiently large that for all $p > N$

$$-\Delta G_p < \frac{1}{4}\epsilon^* \lambda^2 \delta^2.$$

We get from (4.11) and (3.1) that

$$-\Delta G_p \geq \epsilon^*(\lambda/\omega_p)^2 |u^p - u^{p+1}|^2$$

so that

$$|u^p - u^{p+1}| < 6, \ p > N.$$

If for all $u^p \epsilon U$, $p > N$, $r_{g_p}(x) = \theta$, then $(u^{p+1} - u^p, r(x)) = 0$. By the same argument used in Theorem 3.1 of [10], all the u^p, $p > N$ will have to be in U from some point on. If, say, $r,(x) \neq \theta$ for some index $\varkappa$, $1 \leq \varkappa \leq n$, then $\varkappa$ can appear at most a finite number of times among the g_p in the ordering. This contradicts the hypothesis on the infinite covering of Z.

If, on the other hand, there is for some $p > N$ a $u^p \epsilon U$ such that $r_{g_p}(x) \neq 0$, then for each $\varkappa \epsilon g_p$ there is a $w \epsilon \mathfrak{W}(u^p, x)$ such that

$$|r_\varkappa(u^p) - r_\varkappa(x)| \leq |A(w)(u^p - x)| < \Lambda\delta < \rho/2\sqrt{\nu}.$$

Thus for $g = g_p$, $|r_g(u^p) - r_g(x)| < \rho/2$ or $|r_g(u^p)| > \rho/2$.
Since

$$\omega_p |r_g(u^p)| = |A_{gg}(u^p)(u_g^{p+1} - u_g^p)| \leq \Lambda |u^{p+1} - u^p|$$

$$|r_g(u^p)| < \delta\Lambda/\gamma = \rho/2\sqrt{\nu} \leq \tfrac{1}{2}\rho,$$

we get a contradiction and the proof is complete.

8. Remarks. i) It follows from the proof given above that instead of the requirements on G(u) to prevail on the whole of R^n we could simply assume them only in some domain containing K^*.

ii) Another condition which is sufficient for convergence is as follows: Assume that $G(u) \epsilon C^2(R^n)$ and $A(u) > 0$ for all u. Let there exist a point u^* such that

(a) $G(u) \geq G(u^*)$ for all $u \in R^n$,

(b) $A(u^*) > 0$, and

(c) $A_{gg}(u^p) > 0$ for $g = g_p$, $p = 0,1,2,\ldots$.

Then the relaxation processes described above in Sections 5 and 6 will converge for any starting u^0.

Thus we must show that for each z the set $D_z = \{u | G(u) < G(z)\}$ is bounded. We may without loss assume that $u^* = 0$ and assume D_z is unbounded for some z. Then there exists a ray tv, $t \geq 0$, for some fixed v, which lies in D_z. Setting $\varphi(t) = G(tv)$, then $\varphi(t)$ is convex in t and $\varphi'(0) = 0$, $\varphi''(0) > 0$. Thus there exists a $t_0 > 0$ such that $\varphi'(t_0) > 0$. Let $\{t_p\}$ be a sequence of increasing numbers, such that $t_p > t_o$, $p > 0$, $t_p \to \infty$. Since $\varphi'(t_p) \geq \varphi'(t_0) > 0$ and

$$G(z) - G(0) \geq \varphi(2t_p) - \varphi(t_p) \geq \varphi'(t_p)t_p,$$

we get that $\varphi'(t_p) \to 0$, which is impossible.

This argument may be used to show that the minimum u^* is unique, which then guarantees convergence.

iii) A single condition which assures convergence for any initial guess is the existence of a constant μ such that $A(u) > \mu > 0$ for all $u \in R^n$. This occurs in the case-of certain uniformly elliptic problems, as shown in [10].

iv) In [10], it was required that a uniform'upper bound be available for the $a_{ii}(u)$ for the scalar processes. That is, a $\varkappa$ was sought such that $a_{ii}(u) < \varkappa$ for all u and i. If such a bound is available, then an allowable choice of ω_p would be $\omega_p = a_{ii}(u^p)/\varkappa \leq \gamma_p$. This implies that the iteration

$$u_i^{p+1} = u_i^p - r_i(u^p)/\varkappa, \; i = i_p, \; u_{i'}^{p+1} = u_{i'}^p,$$

would converge if any of the sufficient conditions for convergence were satisfied. In the case of a discrete Plateau problem, it was shown in [10] that $a_{ii}(u) \leq 4$ for all u and i. It was also shown there that $a_{ii}(u^p) \geq 4h^6/G(u^0)^3$, where h is the mesh size of the net. If γ is a positive number $< h^6/G(u^0)^3$, then, for example, a choice of $\omega_p = \frac{1}{2} a_{ii}(u^p) - \gamma, \; i = i_p$ would yield convergence for any starting u^0. This represents a considerable improvement over the allowable choice of ω^p given in [10].

v) If a system of equations is given by $r(u) = 0$, $r_i(u) \in C'(R^n)$ and if the Jacobian matrix A(u) of this system is symmetric for all u, then there is a G(u) such that $r(u) = \text{grad}G(u)$. If $A(u) > 0$ for all u, one can check the other sufficient conditions for convergence. An example of this is given by $r(u) \equiv Cu + f(u)$ where C is a constant symmetric matrix such that $C > 0$ and f(u) has a symmetric Jacobian matrix $f'(u) \geq -\varkappa \geq -\lambda(C)$. In this case $A(u) > \mu = \lambda(C) - \varkappa > 0$, so that any starting guess will yield convergence for the relaxation processes described above. This example is realized in the approximate solution of semilinear elliptic boundary problems, when f'(u) is often a diagonal matrix. Thus if one is to solve the usual discrete form of $-\Delta\omega + g(\omega) = 0$ with, say, Dirichlet boundary data, and $g'(\omega) \geq 0$, then the relaxation methods given above will converge from any starting guess. To determine, say, γ_0, one needs an upper bound on $g'(u_i)$ for u in D^0. At times an a priori bound on the solution u^* may be used to bound g! A similar situation is obtained if $-\Delta\omega$ is replaced by a uniformly elliptic self-adjoint, but possibly nonlinear, operator.

REFERENCES

[1] W. F. Ames, Nonlinear partial differential equations in engineering, Academic Press, New York, 1965.

[2] C. W. Cryer, On the numerical solution of a quasi-linear elliptic equation, J. Assoc. Comput. Mach. 14 (1967), 363-375.

[3] P. Concus, Numerical solution of the minimal surface equation, Math. Comp. 21 (1967), 340-350.

[4] H. J. Greenberg, W. S. Dorn, and E. H. Wetherell, A comparison of flow and deformation theories in plastic torsion of a square cylinder, Plasticity: Proc. Symp. Naval Structural Mechanics, 2nd, Brown University, E. H. Lee and P. S. Symonds, Eds., Pergamon Press, New York, 1960.

[5] H. J. Greenberg, Solving structural mechanics problems on digital computers, AEC Computing and Applied Math. Center, Courant Inst. of Math. Sciences, New York University, NYO-8670, 1958.

[6] A. S. Householder, Principles of numerical analysis, McGraw-Hill Book Co., Inc., New York, 1953.

[7] J. M. Ortega and M. L. Rockoff, Nonlinear difference equations and Gauss-Seidel type iterative methods, SIAM J. Numer. Anal. 3 (1966), 497-513.

[8] A. M. Ostrowski, Solution of equations and systems of equations, 2nd Ed., Academic Press, New York, 1966.

[9] D. J. Prager and M. L. Rasmussen, The flow of a rarefied plasma past a sphere, Institute for Plasma Research, Stanford University, SU-IPR Report No. 132 (or SU-DAAR No. 299), 1967.

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

[11] __________, Relaxation methods for linear equations, Comm. Pure Appl. Math. 12 (1959), 313-335.