Technical Report 68-59 January 1968

CONVERGENCE THEOREMS FOR GAUSS-SEIDEL AND OTHER MINIMIZATION ALGORITHMS

by

Richard Michael Elkin
Computer Science Center
University of Maryland

The completion of this report was in part supported by the National Aeronautics and Space Administration under Grant NsG 398 to the Computer Science Center of the University of Maryland.

ABSTRACT

Iterative minimization algorithms are studied for functionals defined on E^n and convergence theorems are obtained. First, a general review of convexity for functionals is given. Then, the two aspects of an iterative minimization algorithm --the direction in which the next iterate is sought and the step-size--are independently analyzed. In the analysis of direction emphasis is placed on the Gauss-Seidel and the block Gauss-Seidel methods rather than on gradient methods. A variety of step-size algorithms are studied including minimization, Curry, over-relaxed Curry, use of one Newton step and the methods of Altman, Armijo, and Goldstein. Finally, complete convergence theorems are given for representative algorithms.

ACKNOWLEDGMENTS

I would like to express my gratitude to my thesis advisor, Professor James M. Ortega, for his invaluable suggestions, patience and guidance. Also I would like to thank Professor Werner C. Rheinboldt for many helpful discussions, comments and criticisms. Further, I owe thanks to Mrs. Faye von Limbach for her help in the preparation of the manuscript.

Finally, I would like to thank the National Aeronautics and Space Administration for their generous financial support under Grant NsG-398 to the Computer Science Center of the University of Maryland.

TABLE OF CONTENTS

INTRODUCTION

In recent years iterative methods for finding the minimum of a (not necessarily quadratic) functional $g: D \subset E^n \to R$ have received considerable attention. In this paper we are concerned with convergence theorems for such iterative minimization algorithms and in particular for the Gauss-Seidel minimization algorithm.

Any iterative minimization algorithm which produces a sequence of iterates satisfying $u^{p+1} = u^p - t_p e^p$ has two tasks at each step: selecting the direction e^p in which the next approximation must lie, and choosing t_p the step-size or distance to the next iterate. Sometimes these decisions are made simultaneously, as with Newton's method, or are closely bound to each other as with conjugate gradient methods. But generally the decisions can be made independently, and we shall analyse them independently.

The directions we analyse fall into two categories. The first consists of directions which are related to the gradient direction. To minimize a functional g these methods select at u^p a direction e^p satisfying

$$g'(u^p)e^p \geq \delta_p \|g'(u^p)\|$$

where $\|e^p\| = 1$ and $\{\delta_p\}$ tends to zero only if $\{\|g'(u^p)\|\}$ tends to zero. Under suitable hypotheses, certain conjugate gradient methods and the methods of Newton, Jacobi, Gauss-Southwell and Seidel are gradient-related.

A great many papers deal with gradient or gradient-related minimization algorithms. Among the more recent are Altman [1], [2], Armijo [3], Goldstein [14],[15],[16],[17], Nashed [27], and Ostrowski [32]. The thrust of their work has been to extend the spaces in which the results hold, reduce the differentiability assumptions about the functional, to increase the number of ways in which step-size may be determined, and to generalize the relationship to the gradient.

In the second category are the Gauss-Seidel, Gauss-Seidel-Newton, block Gauss-Seidel, block Gauss-Seidel-Newton and Rosenbrock [36] algorithms, which are not gradient-related methods. In contrast to the many papers on gradient methods, only Schechter [38] has obtained non-local convergence results for Gauss-Seidel and Gauss-Seidel-Newton. The other methods have not previously been investigated at all for non-quadratic functionals (or non-linear equations).

This paper is organized in the following manner. Chapter I is background material and deals essentially with convexity for functionals. The relationships between convexity of a functional, monotonicity of its derivative, and positive definiteness of its second derivative are reviewed. We then consider pseduo-convexity and quasi-convexity and the relationships between the various kinds of convexity are examined. Finally we examine another kind of convexity and characterize it as precisely those quasi-convex functionals for which every local minimum is a global minimum.

Chapters II and III contain the basic results on convergence. In Chapter III a wide variety of methods for choosing step-size are examined, including, among others, all of the step-size algorithms given in [2], [3], [14],[16], [17] and [38]. We extend the applicability of some of these methods and study the principles underlying the proofs. The analysis is carried out in sufficient generality so that the results are independent of the direction algorithms. In Chapter II we analyse the direction algorithms and give conditions for $g'(u^p)$ to tend to zero and for the sequence of iterates to converge.

Chapter IV is devoted to applying the results of the previous chapters to specific algorithms. Because we can pair (in general) any distance choice with any direction choice there are a great many potential algorithms, most of them new but uninteresting. We therefore restrict ourselves to illustrating rather than exhausting the possibilities.

Chapter V is devoted to block methods generally, and block Gauss-Seidel and block successive over-relaxation in particular.

We can now summarize our results. The paper contains the first nonlocal convergence theorems for non-linear successive over-relaxation and the corresponding block methods. We also greatly extend the class of functions for which Gauss-Seidel is known to converge, eliminating in some cases convexity conditions on the functional. We also analyse for the first time the steepest descent-Newton method and the Jacobi iteration. Perhaps more important is the general theory which separates the questions of step-size and direction and analyses the essential factors in each.

CHAPTER I

CONVEXITY

This chapter contains primarily background material. Following some brief introductory material on derivatives the bulk of the chapter deals with various kinds of convexity for functionals.

These sections are intended as a survey of the relevant results in Mangasarian [25], Minty [26], Newman [29], Poljak [34], Ponstein [35], and Wilde [45]. We also obtain some new results which further characterize the different kinds of convexity and clarify the relationships between them.

<u>1.1 Preliminaries</u>. Let X and Y denote real Banach spaces, and $F: D \subset X \to Y$ a mapping defined on a subset D of X. If for some $x \in D$ and every $h \in X$

$$(1.1.1) \qquad VF(x;h) = \lim_{t \to 0} \frac{F(x+th) - F(x)}{t}$$

exists then we say that F is Gateaux differentiable at x. $VF(x;h)$ is called the Gateaux differential of F at x in the direction h. When $VF(x;h)$ is bounded and linear in h we denote it by $F'(x)h$ and $F'(x)$ is called the Gateaux derivative of F at x. If, in addition,

$$(1.1.2) \qquad \lim_{\|h\| \to 0} \frac{\|F(x+h) - F(x) - F'(x)h\|}{\|h\|} = 0$$

holds, then $F'(x)$ is called the Frechet derivative of F at x. If the Gateaux derivative exists in a neighborhood of x and is continuous at x then (1.1.2) holds and the Frechet derivative also exists at x. A complete discussion will be found in Vainberg [41].

If F maps a subset of X to Y, then F' is a mapping into $L(X,Y)$, the Banach space of all bounded linear operators from X to Y. The second (Gateaux or Frechet) derivative F'' is defined as above by $F'' = (F')'$. In particular, when F is a functional (that is $Y = R$, the real numbers), then $L(X,R)$ is the conjugate space X^* of X and whenever $F''(x)$ exists it is a bounded linear map from X to X^*, since, for $h \in X$, $F''(x)h \in X^*$, and is a bounded linear functional. $F''(x)$ may also be thought of as a bilinear mapping from $X \otimes X$ to R.

Let the open interval or line segment $\{u+t(v-u):0<t<1\}$ be denoted by (u,v) and the corresponding closed interval by $[u,v]$. We say that a map $G:D \subset X \to Y$ is continuous on line segments if for any closed interval $[u,v] \subset D$, and $0 \leqq t \leqq 1$, $G(u+t(v-u))$ is a continuous function of t. It follows immediately from the definition of a Gateaux derivative, (1.1.1) that if G is Gateaux differentiable in D then G is continuous on line segments in D.

For functionals we have a mean value theorem that is similar to and based on the corresponding theorem for functions of a single real variable. (See Vainberg [41].)

Let $g:D \subset X \to R$ have a Gateaux derivative that is continuous on line segments and let the closed interval $[u,v]$ lie in D. Then

$$g(u) - g(v) = g'(w)(u-v) \tag{1.1.3}$$

for some w in the open interval (u,v). There is also an integral form of the mean value theorem. Under the same assumptions as above

$$g(u) - g(v) = \int_0^1 g'(u+t(v-u))(v-u)\, dt. \tag{1.1.4}$$

A similar relation is obtained for $G:D \subset X \to X^*$. Let G have a Gateaux derivative that is continuous on line segments in D and suppose that the closed interval $[u,v]$ lies in D. Then

$$(G(u) - G(v))(u-v) = \int_0^1 [\, G'(u+t(v-u))(u-v)(u-v)]dt. \tag{1.1.5}$$

For further discussion of Gateaux and Frechet derivatives we refer the reader to Vainberg [41] and Dieudonne [8].

We will frequently use functions whose values approach zero only when their arguments do. We therefore define:

<u>Definition 1.1.1</u>. A function $d:[0,\infty) \to [0,\infty)$ <u>forces its argument to zero</u> if for any positive sequence $\{t_n\}$,

$$\lim_{n\to\infty} d(t_n) = 0 \quad \text{implies} \quad \lim_{n\to\infty} t_n = 0. \tag{1.1.6}$$

For brevity we simply say d is <u>forcing</u>. The following lemma, whose proof is obvious, characterizes forcing functions.

<u>Lemma 1.1.1</u>. A function $d:[0,\infty) \to [0,\infty)$ forces its argument to zero if and only if d is bounded away from zero in $[c,\infty)$ for any positive c.

Clearly every non-decreasing function d such that $d(t) > 0$ for $t > 0$ is forcing. We shall call such functions <u>non-decreasing forcing functions</u> and use them throughout this chapter. Non-decreasing forcing functions include all polynomials with positive coefficients. More generally we have the following, whose proof is immediate by Lemma 1.1.1.

<u>Lemma 1.1.2</u>. The sum, product and composition of any two (non-decreasing) forcing functions is again a (non-decreasing) forcing function.

We also have

<u>Lemma 1.1.3</u>. If d is an integrable forcing function and ε is a non-negative constant then the function $\bar{d}$ defined by

$$\bar{d}(t) = \int_0^1 d(st+\varepsilon)\,ds$$

is forcing.

Proof: For any t define $t' = \frac{1}{2}t$. By Lemma 1.1.1 there is some $c' > 0$ such that $s \geqq t'$ implies $d(s) \geqq c'$. But then for $s \geqq t$,

$$\bar{d}(s) = \int_0^1 d(s\theta+\epsilon)\,d\theta = 1/s \int_\epsilon^{s+\epsilon} d(\theta)\,d\theta \geqq 1/s \int_{\frac{1}{2}s+\epsilon}^{s+\epsilon} c'\,d\theta \geqq \tfrac{1}{2}c'$$

and thus $\bar{d}$ is forcing.

We end this section by introducing a particular forcing function that we shall use frequently in later chapters. Let $g:D \subset X \to R$ be a functional with a uniformly continuous Frechet derivative g' in D and define

$$(1.1.7) \qquad \delta(t) = \inf\{\|u-v\|:u,v \in D;\|g'(u)-g'(v)\| \geqq t\}$$

for $0 \leqq t \leqq T = \sup\{\|g'(u)-g'(v)\|:u,v \in D\}$. For a given $\epsilon > 0$, the uniform continuity of g' implies that there is δ' such that

$$(1.1.8) \qquad \|u-v\| < \delta' \quad \text{implies} \quad \|g'(u)-g'(v)\| < \epsilon \ ;$$

hence, $\delta(\epsilon)$ as defined by (1.1.7) is the largest value of δ' such that (1.1.8) holds. Further δ is a non-decreasing function and by the uniform continuity of g', $\delta(t) > 0$ for $t > 0$. Thus δ is a non-decreasing forcing function. The function δ is essentially the inverse of the modulus of continuity of g' defined by

$$(1.1.9) \qquad \omega(t) = \sup\{\|g'(u)-g'(v)\|:u,v \in D, \|u-v\| \leqq t\}.$$

1.2 Convexity, Monotonicity and Positive Definiteness.

We next examine the relationship of convexity of a functional monotonicity of its derivative, and positive definiteness of its second derivative.

A functional $g:D \subset X \to R$ is <u>convex</u> in some convex set $D_0 \subset D$ if for all $u,v \in D_0$ and all $t \in [0,1]$,

(1.2.1) $$g(tu+(1-t)v) \leqq tg(u)+(1-t)g(v).$$

g <u>is strictly convex</u> if for $u \neq v$ and $t \in (0,1)$,

(1.2.2) $$g(tu+(1-t)v) < tg(u)+(1-t)g(v).$$

g is <u>uniformly convex</u> if for some non-decreasing forcing function d,

(1.2.3) $$g(tu+(1-t)v) \leqq tg(u)+(1-t)g(v) - \|u-v\| \max\begin{Bmatrix} td((1-t)\|u-v\|) \\ (1-t)d(t\|u-v\|) \end{Bmatrix}$$

The definitions of convexity and strict convexity are standard and classical; uniform convexity does not seem to have been defined until now although Poljak [34] defines a functional to be <u>strongly convex</u> when

(1.2.3') $$g((u+v)/2) \leqq \tfrac{1}{2}g(u) + \tfrac{1}{2}g(v) - \tfrac{1}{4}c\|u-v\|^2,$$

for $c > 0$.

We recall next the concept of a monotone mapping. If $g:D \subset X \to R$ has a Gateaux derivative g', then g' is <u>monotone</u> on D if for all $u,v \in D$

(1.2.4) $$(g'(u)-g'(v))(u-v) \geqq 0.$$

g' is <u>strictly monotone</u> if for $u \neq v$,

(1.2.5) $$(g'(u)-g'(v))(u-v) > 0.$$

g' is <u>uniformly monotone</u> if for some non-decreasing forcing function d,

(1.2.6) $$(g'(u)-g'(v))(u-v) \geqq \|u-v\| d(\|u-v\|).$$

Monotonicity has been defined in more generality for mappings $G: D \subset X \to X^*$ by Zarantonello [45] and exploited by various authors. Our definition of uniform monotonicity is a slight generalization of what is usually called <u>strong monotonicity</u>: $(d(t) = ct,\ c > 0)$

(1.2.6') $$(g'(u)-g(v))(u-v) \geqq c\|u-v\|^2.$$

We see immediately that uniform convexity implies strict convexity which in turn implies convexity. Likewise, uniform monotonicity implies strict monotonicity which implies monotonicity.

For functionals g with Gateaux derivatives which are continuous on line segments, we show below that convexity (respectively, strict convexity or uniform convexity) is equivalent to monotonicity (respectively, strict monotonicity or uniform monotonicity) of g'. These results, although elementary, appear to have been overlooked in the literature until recently. Minty [26] showed that convexity of g implies monotonicity of g'. This is done for a topological vector space and a generalization of the Frechet derivative. Poljak [34] states without proof the equivalence of (1.2.3') and (1.2.6'). For completeness we prove the following although only the equivalence of the "strict" statement seems to be entirely new.

<u>Theorem 1.2.1</u>. Let D be an open convex subset of a real Banach space X, and assume $g:D \subset X \to R$ has a Gateaux derivative that is continuous on line segments in D. Then:

(a) g' is monotone in D if and only if g is convex in D,

(b) g' is strictly monotone in D if and only if g is strictly convex in D,

(c) g' is uniformly monotone in D if and only if g is uniformly convex in D.

<u>Proof</u>: From the mean value theorem, (1.1.4)

$$g(u)-g(v) = \int_0^1 [g'(v+\theta(u-v))(u-v)]d\theta$$

and

$$g(v+\alpha(u-v))-g(v) = \int_0^1 [g'(v+\theta\alpha(u-v))(\alpha(u-v))]d\theta$$

so that

$$g(\alpha u+(1-\alpha)v)-\alpha g(u)-(1-\alpha)g(v) =$$

$$\alpha\int_0^1 [(g'(v+\theta\alpha(u-v))-g'(v+\theta(u-v))](u-v)d\theta =$$

$$\frac{\alpha}{\alpha-1}\int_0^1 \frac{[g'(v+\theta\alpha(u-v))-g'(v+\theta(u-v))]\ [v+\theta\alpha(u-v)-v+\theta(u-v)]}{\theta}d\theta$$

If g' is monotone the integrand is non-negative and g is convex. If g' is strictly monotone the integrand is positive and g is strictly convex. When g' satisfies (1.2.6) we have

$$g(\alpha\cdot u+(1-\alpha)v) - (\alpha g(u)+(1-\alpha)g(v)) \leqq \frac{\alpha}{1-\alpha}\int_0^1 (\|\theta(\alpha-1)(u-v)\|\cdot d(\|(\alpha-1)(u-v)\|)d\theta$$

and thus

$$\alpha g(u)+(1-\alpha)g(v) - g(\alpha u+(1-\alpha)v) \geqq \alpha\|u-v\|\bar{d}((1-\alpha)\|u-v\|)$$

where

$$\bar{d}(t) = \int_0^1 d(t\theta)d\theta.$$

If we set $\beta = (1-\alpha)$ and interchange the roles of u and v we have

$$\beta g(v)+(1-\beta)g(u) - g(\beta u+(1-\beta)v) \geqq \beta\|u-v\|\bar{d}((1-\beta)\|u-v\|).$$

Since d is non-decreasing and satisfies $d(t) > 0$ for $t > 0$, $\bar{d}$ has these same properties. Hence (1.2.3) holds and g is uniformly convex.

Conversely, if g is uniformly convex and $0 < \alpha < 1$, (1.2.3) yields

$$\frac{g(v+\alpha(u-v))-g(v)}{\alpha} \leqq g(u) - g(v) - \|u-v\|d((1-\alpha)\|u-v\|).$$

Now let α tend to zero. Since g is differentiable at v and d is a non-decreasing function, we conclude that

$$g'(v)(u-v) \leqq g(u) - g(v) - \|u-v\|d(\|u-v\|). \tag{1.2.7}$$

Interchanging the roles of u and v in (1.2.7) and adding the result to (1.2.7) we have

$$[g'(u)-g'(v)](u-v) \geqq 2\|u-v\|d(\|u-v\|).$$

Thus if g is uniformly convex, g' is uniformly monotone and with d set to zero this also shows that convexity of g implies monotonicity of g'. If g is strictly convex and $u \neq v$ we let $w = \frac{1}{2}(u+v)$ and obtain

$$g'(u)(v-u) = 2g'(u)(w-u)$$
$$\leqq 2[g(w)-g(u)]$$
$$< 2[\tfrac{1}{2}g(v) - \tfrac{1}{2}gu].$$

Thus

$$g'(u)(v-u) < g(v) - g(u)$$

and we conclude the strict monotonicity of g'.

We note that in the course of the proof we have actually obtained the following basic differential inequalities which hold without any continuity assumptions on g':

(i) g is convex if and only if

(1.2.8) $\quad g'(v)(u-v) \leqq g(u) - g(v) \leqq g'(u)(u-v);$

(ii) g is strictly convex if and only if for $u \neq v$

(1.2.9) $\quad g'(v)(u-v) < g(u) - g(v) < g'(u)(u-v);$

(iii) g is uniformly convex if and only if for some non-decreasing forcing function d

(1.2.10) $\quad g'(v)(u-v) \leqq g(u) - g(v) - \|u-v\| d(\|u-v\|).$

Corresponding to convexity for the functional and monotonicity for the first derivative is positive definiteness of the second derivative. Again we have a three part definition.

Let Y be the collection of all bilinear mappings from $X \times X$ to R. A map $A: D \subset X \to Y$ is <u>positive definite</u> in D if for all $u \in D$ and $h \in X$

(1.2.11) $\quad A(u)hh \geqq 0.$

A is <u>strictly positive definite</u> if for $h \neq 0$

(1.2.12) $$A(u)hh > 0.$$

A is <u>uniformly positive definite</u> if for some $c > 0$,

(1.2.13) $$A(u)hh \geqq c\|h\|^2.$$

This nomenclature is slightly non-standard. The first concept is usually called non-negative definite or positive semi-definite and the second is called positive definite. Note that we cannot generalize uniform positive definiteness in a manner analogous to uniform monotonicity and uniform convexity by means of an arbitrary non-decreasing forcing function. $A(u)$ is by definition a bilinear operator and must act quadratically on h. Observe also that if D is compact and A is continuous then strict positive definiteness of A in D implies uniform positive definiteness of A in D. Since the converse is always true, strict and uniform positive definiteness are equivalent for continuous operators on compact domains. We next give the relationships between monotonicity of g' and positive definiteness of g''.

<u>Theorem 1.2.2</u>. Let $g:D \subset X \to R$ have a second Gateaux derivative that is continuous on line segments on D. Then:

(a) g'' is positive definite if and only if g' is monotone.

(b) If g'' is positive definite and strictly positive definite except perhaps on a set which contains no line segments then g' is strictly monotone.

(c) g'' is uniformly positive definite if and only if g' is uniformly monotone with a linear forcing function.

Proof: From the mean value theorem, (1.1.5),

$$(g'(u)-g'(v))(u-v) = \int_0^1 [\, g''(tv+(1-t)u)(u-v)(u-v)]dt.$$

If g'' is positive definite the integrand is non-negative and thus g' is monotone. If $u \neq v$ and g'' is positive definite and strictly positive definite except perhaps on a set containing no line segments, then g'' is strictly positive definite at some point of the closed interval $[u,v]$. But g'' is continuous on line segments and is therefore strictly positive definite on some sub-interval of $[u,v]$. Therefore the integrand is positive and g' is strictly monotone. If g'' is uniformly positive definite the integrand is greater than $c\|u-v\|^2$ for some $c > 0$, and g'' is uniformly positive definite.

Conversely if

$$(g'(u)-g'(v))(u-v) \geqq c\|u-v\|^2$$

then for $h \in X$, $u \in D$, and any $t \neq 0$,

$$\frac{(g'(u+th)-g'(u))th}{t^2} \geqq \frac{c\|th\|^2}{t^2}$$

and letting t approach zero we have by the differentiability of g',

$$g''(u)hh \geqq c\|h\|^2,$$

so that g'' is uniformly positive definite. If we set

$c = 0$ this shows that g'' is positive definite when g' is monotone.

The converse of (b) is not true as the example, $x_1^2 + x_2^4$, shows. This functional is strictly convex and thus has a strictly monotone derivative. But the second derivative is singular on the line $x_2 = 0$ and therefore is not strictly positive definite.

Part (a) of the theorem is a classical result. We note that the statement and proof of the entire theorem extend immediately to differentiable mappings $G:D \subset X \to X^*$, where G is not necessarily the gradient of a functional.

1.3 Pseudo-Convexity and Quasi-Convexity. A functional g which has a Gateaux derivative g' on an open subset D of a Banach space is pseudo-convex if for all $u,v \in D$

(1.3.1) $$g(u) < g(v) \quad \text{implies} \quad g'(v)(v-u) > 0.$$

g is strictly pseudo-convex if for $u \neq v$

(1.3.2) $$g(u) \leqq g(v) \quad \text{implies} \quad g'(v)(v-u) > 0.$$

g is uniformly pseudo-convex if for some non-decreasing forcing function d

(1.3.3) $$g(u) \leqq g(v) \text{ implies } g'(v)(v-u) \geqq \|u-v\| d(\|u-v\|).$$

The concepts of pseudo-convexity and strict pseudo-convexity are due to Mangasarian [25] while that of uniform pseudo-convexity is apparently new. Note that differentiability is necessary for the definition of pseudo-convexity (but see Nashed [28] for the notion of supportable convexity which seems to be a natural generalization of pseudo-convexity). On the other hand, it is not necessary that a pseudo-convex function be defined on a convex set.

An important property of pseudo-convex functionals is that critical points are minima. The following is essentially due to Mangasarian [25].

Theorem 1.3.1. Assume that the functional g is Gateaux differentiable on an open set $D \subset X$ and let H be a linear subspace of X. If g is (strictly) pseudo-convex on D and for some $x \in D$, $g'(x)h = 0$ for all $h \in H$, then g attains a (unique) minimum in $D \cap \hat{H}$ at x, where $\hat{H}$ is the affine subspace $\{x+h : h \in H\}$.

Proof: Suppose $z \in D \cap \hat{H}$ and $g(z) < g(x)$. Then pseudo-convexity of g implies that $g'(x)(x-z) > 0$ which, since $x-z \in H$, is a contradiction. Similarly, if $g(z) \leqq g(x)$ and $x \neq z$, we obtain the same contradiction if g is strictly pseudo-convex.

Corollary 1. Let $u \in D$, $e \neq 0$ and define $f(t) = g(u-te)$ for t such $u-te \in D$. If $g'(u-t_0e)e = 0$ and g is (strictly) pseudo-convex on D, then f takes on a (unique) minimum at t_0.

Corollary 2. If g is (strictly) pseudo-convex on D and $g'(x) = 0$, then g attains a (unique) minimum in D at x.

We show next that each form of pseudo-convexity is weaker than the corresponding form of convexity.

Theorem 1.3.2. Suppose g has a Gateaux derivative on an open convex subset D of a Banach space. If g is convex (respectively, strictly convex or uniformly convex) on D then g is pseudo-convex (respectively, strictly pseudo-convex or uniformly pseudo-convex) on D.

Proof: When g is convex, strictly convex or uniformly convex g' satisfies by (1.2.8),(1.2.9) or (1.2.10) respectively either

$$g'(v)(v-u) \geqq g(v) - g(u), \tag{1.3.4}$$

$$g'(v)(v-u) > g(v) - g(u), \tag{1.3.5}$$

or for some non-decreasing forcing function d

$$g'(v)(v-u) \geqq g(v) - g(u) + \|u-v\| d(\|u-v\|). \tag{1.3.6}$$

If $g(u) \leqq g(v)$ these relations immediately imply that g is pseudo-convex, strictly pseudo-convex or uniformly pseudo-convex respectively.

It is immediate that a convex functional has convex level sets, (that is, $\{x \in D : g(x) \leqq c\}$ is convex for any c) but the converse is not true. This leads to the well-known (see e.g. Fenchel [10]) definition of a quasi-convex functional as one whose level sets are all convex. Equivalently, we may define quasi-convexity as follows.

A functional g mapping a convex subset D of a Banach space X into R is <u>quasi-convex</u> if for all u,v in D and all w in the open interval (u,v)

$$g(u) \leqq g(v) \quad \text{implies} \quad g(w) \leqq g(v). \tag{1.3.7}$$

g is <u>strictly quasi-convex</u> if for $u \neq v$,

$$g(u) \leqq g(v) \quad \text{implies} \quad g(w) < g(v). \tag{1.3.8}$$

g is <u>uniformly quasi-convex</u> if for some non-decreasing forcing function d

$$g(u) \leqq g(v) \text{ implies } g(v) \geqq g(w) + \min \left\{ \begin{matrix} \|u-w\| d(\|u-w\|) \\ \|v-w\| d(\|v-w\|) \end{matrix} \right\}. \tag{1.3.9}$$

While the definition of quasi-convexity is standard we differ with Mangasarian [25] and Ponstein [35] on the definition of strict quasi-convexity. Their definition of strict quasi-convexity (which we will call semi-strict quasi-convexity in the next section) does not, as (1.3.8) does, imply the uniqueness of a minimum of g. Poljak [34] has also given definitions of strict and uniform quasi-convexity which are equivalent to ours when the functional is lower semi-continuous.

We have immediately, that uniform quasi-convexity implies strict quasi-convexity which implies quasi-convexity. Furthermore, each kind of quasi-convexity is a consequence of the corresponding kind of pseudo-convexity.

<u>Theorem 1.3.3</u>. Let g have a Gateaux derivative that is continuous on line segments in an open convex subset D of a Banach space. If g is pseudo-convex (respectively, strictly pseudo-convex or uniformly pseudo-convex) then g is quasi-convex (respectively, strictly quasi-convex or uniformly quasi-convex).

<u>Proof</u>: Assume for $u \neq v$ that $g(u) \leqq g(v)$. Let w belong to the open interval (u,v) and let x be the minimum of g on the closed line segment $[u,v]$. (g is continuous and $[u,v]$ is compact so that the minimum is attained.) Choose z as u or v so that w belongs to the line segment (z,x). By the mean value theorem, (1.1.4),

$$g(z) - g(w) = \int_0^1 g'(w+t(z-w))(z-w)\,dt$$

and thus, since $g(z) \leqq \max\{g(u), g(v)\} = g(v)$,

$$g(v) - g(w) \geqq \int_0^1 g'(w+t(z-w))(z-w)\,dt. \tag{1.3.10}$$

But $z-w$ is a positive multiple of $w+t(z-w)-x$ and thus the sign of $g'(w+t(z-w))(z-w)$ is the same as that of $g'(w+t(z-w))(w+t(z-w)-x)$. Since $g(x) \leqq g(w+t(z-w))$, pseudo-convexity implies that the integrand of (1.3.10) is non-negative and hence g is quasi-convex; likewise strict pseudo-convexity implies the integrand of (1.3.10) is positive and hence g is strictly quasi-convex. If g is uniformly pseudo-convex the integrand satisfies

$$\begin{aligned} g'(w+t(z-w))(z-w) &= \frac{g'(w+t(z-w))(w+t(z-w)-x)(\|z-w\|)}{\|w+t(z-w)-x\|} \\ &\geqq \frac{d(\|w+t(z-w)-x\|)\|w+t(z-w)-x\|\cdot\|z-w\|}{\|w+t(z-w)-x\|} \\ &= d(\|t(z-w)+w-x\|)\|z-w\| \\ &\geqq d(\|t(z-w)\|+\|w-x\|)\|z-w\|, \end{aligned}$$

so that

$$g(v)-g(w) \geqq \bar{d}(\|z-w\|)\|z-w\|,$$

where $\bar{d}(t) \equiv \int_0^1 d(t\theta)\,d\theta$. As z may be either u or v we have

$$g(v)-g(w) \geqq \min \begin{Bmatrix} \bar{d}(\|u-w\|)\|u-w\| \\ \bar{d}(\|v-w\|)\|v-w\| \end{Bmatrix}$$

and as d is a non-decreasing forcing function this implies $\bar{d}$ is a non-decreasing forcing function. Hence g is uniformly quasi-convex.

The relationships derived in Theorems 1.3.2 and 1.3.3 are summarized in Figure 1.3.1.

$$\begin{array}{ccccc} \text{UC} & \Rightarrow & \text{UPC} & \Rightarrow & \text{UQC} \\ \Downarrow & & \Downarrow & & \Downarrow \\ \text{SC} & \Rightarrow & \text{SPC} & \Rightarrow & \text{SQC} \\ \Downarrow & & \Downarrow & & \Downarrow \\ \text{C} & \Rightarrow & \text{PC} & \Rightarrow & \text{QC} \end{array}$$

Figure 1.3.1 Relationships between various kinds of convexity.

There are a number of other formulations of quasi-convexity. Poljak [34] uses the mid-point formulation: For $u, v \in D$

$$(1.3.11) \qquad g(u) \leqq g(v) \text{ implies } g\left(\frac{u+v}{2}\right) \leqq g(v).$$

Newman [29] and Wilde [44] each generalize a concept of Keifer [23] to n dimensions and call it unimodality and linear unimodality respectively. Their definition is equivalent to what we call strict unimodality below and turns out (Theorem 1.3.5) to be equivalent to strict quasi-convexity.

A functional $g: D \subset X \to R$, defined on a convex domain D is <u>unimodal</u> if, for any closed interval $[u,v] \subset D$ we have $g(u) \geqq g(w) \geqq g(x^*)$ where x^* is the minimum of g on $[u,v]$ and w is any point in the open interval (u,x^*). g is <u>strictly unimodal</u> if $g(u) > g(w) > g(x^*)$ when $u \neq x^*$.

We have already noted that quasi-convexity is equivalent to convexity of level sets. Other equivalences are given in the next theorem. We recall that a lower semi-continuous functional satisfies $\liminf_{x \to x^*} g(x) \geqq g(x^*)$, and takes on a minimum on any compact set.

<u>Theorem 1.3.4</u>. Let g be a lower semi-continuous functional defined on a convex subset D of a Banach space. Then the following are equivalent:

(a) g is quasi-convex;

(b) g is unimodal;

(c) for any $u,v \in D$, $g(u) \leqq g(v)$ implies $g(\frac{u+v}{2}) \leqq g(v)$;

(d) for every real c the set $S_c = \{u:g(u)<c\}$ is convex.

If g is continuously differentiable the following are also euqivalent to quasi-convexity:

(e) for any $u,v \in D$
$g(u) \leqq g(v)$ implies $g'(v)(v-u) \geqq 0$;

(f) for any $u,v \in D$
$g(u) < g(v)$ implies $g'(v)(v-u) \geqq 0$.

Proof: (a)$\Rightarrow$(b). Let x^* be the minimum of g on $[u,v]$ and $w \in (u,x^*)$. Then $g(x^*) \leqq g(w)$ and $g(x^*) \leqq g(u)$. The last inequality and quasi-convexity imply $g(w) \leqq g(u)$. (b)$\Rightarrow$(c). Assume $g(u) \leqq g(v)$. Let x^* be the minimum of g on $[u,v]$. If $x^* = \frac{u+v}{2}$ we are done. If not, $\frac{u+v}{2}$ belongs to (u,x^*) or (v,x^*) and thus $g(\frac{u+v}{2})$ is less than $g(u)$ or $g(v)$ respectively. In either case $g(\frac{u+v}{2}) \leqq g(v)$. (c)$\Rightarrow$(d). If $u,v \in S_c$, then $\max\{g(u),g(v)\} < c$. We prove by induction on i that the points $w_{i,j} = j2^{-i}u+(1-j2^{-i})v, (0\leqq j\leqq 2^i; i=0,1,..)$ satisfy $g(w_{i,j}) \leqq \max\{g(u),g(v)\}$. When $i = 0$ the result is clear. Now assume $g(w_{i',j}) \leqq \max\{g(u),g(v)\}$ for $i' \leqq i$; there then are two cases to consider for i. When j is even, $w_{i,j}$ is also the point $w_{i-1,\frac{1}{2}}$, and by the induction hypothesis $g(w_{i,j}) \leqq \max\{g(u),g(v)\}$. When j is odd, $w_{i,j} = (w_{i-1,\frac{1}{2}j-1} + w_{i-1,\frac{1}{2}j+1})/2$ and the induction follows from (c).

The set $G = \{w_{i,j}: 0 \leqq j \leqq 2^i;\ i = 0,1,..\}$ is dense in (u,v) and any point w of (u,v) is therefore the limit of some sequence of points in G. But the value of g at each point in G is less than or equal to $\max\{g(u),g(v)\}$ and by the lower semi-continuity of g, $g(w) \leqq \max\{g(u),g(v)\} < c$.

(d)$\Rightarrow$(a). Let $u,v \in D$, $g(u) \leqq g(v)$, $w \in (u,v)$ and assume that $g(w) > g(v)$. Then u and v both belong to $S_{g(w)}$. Convexity of $S_{g(w)}$ implies $w \in S_{g(w)}$ which is a contradiction. Thus $g(w) \leqq g(v)$. The equivalence of (a), (d), (e) and (f) was proved by Ponstein [35] in finite dimensions under the condition that g is continuously differentiable. Since his proof of the equivalence of (a), (e) and (f) goes over verbatim to a Banach space we do not repeat it here.

Poljak[34] states equivalence of (c) and (a) for lower semi-continuous functionals. The recognition that unimodality and quasi-convexity are related seems to have gone unnoticed until now.

We next consider equivalent formulations of strict quasi-convexity. As we might expect strict quasi-convexity is equivalent to strict unimodality and strict mid-point quasi-convexity when the functional is lower semi-continuous.

<u>Theorem 1.3.5</u>. Let g be a lower semi-continuous functional defined on a convex subset D of a Banach space.

Then the following are equivalent:

(a) g is strictly quasi-convex;

(b) g is strictly unimodal;

(c) for $u, v \in D$, $g(u) \leqq g(v)$ implies $g(\frac{u+v}{2}) < g(v)$.

Proof: We note that (a), (b) and (c) each imply g is quasi-convex. For quasi-convex functionals if there is some w in the open interval (u,v) such that $g(u) \leqq g(w) = g(v)$ then g is constant either on the interval $[u,w]$ or on the interval $[w,v]$. For if g were not constant in either we could find x' in $[u,w]$ and x'' in $[w,v]$ such that $g(x') < g(w)$ and $g(x'') < g(w)$. Hence, $g(w)$ is strictly greater than the maximum of $g(x')$, $g(x'')$ even though $w \in [x', x'']$. This contradicts quasi-convexity. Now assume (a) holds. To prove (b) assume that x is the minimum of g on the interval $[u,v]$ and that x is not equal to u. g is quasi-convex, hence unimodal, and for any w in the open interval (u,x), $g(u) \geqq g(w) \geqq g(x)$. We need only show each of these inequalities is strict. If $g(u) = g(w)$ then as noted above g is constant on a subinterval. But a strictly quasi-convex function cannot be constant on a subinterval. If $g(w) = g(x)$, then by quasi-convexity any w' in (w,x) satisfies $g(w') < g(x)$ which contradicts the definition of x. Now assume (b), and for $u \neq v$ set $w = \frac{u+v}{2}$. By the quasi-convexity of g

we know that $g(w) \leqq \max\{g(u),g(v)\}$; therefore we wish only to show that the equality is strict. But if we had equality we could conclude g was constant on some line-segment and thus contradict strict unimodality. As similar argument shows (c) implies (a). If we had equality, i.e., $g(w) = \max\{g(u),g(v)\}$, for some w in the open interval (u,v), $u \neq x$ then we would have g constant on a line segment and this contradicts (c). This completes the proof.

Strictly quasi-convex functionals have the property that if a minimum exists, it is unique. In the next section we give a geometric interpretation of strict quasi-convexity.

<u>1.4 Strict and Semi-strict Quasi-convexity</u>. We noted in the last section that quasi-convexity was equivalent to assuming all level sets are convex. Our next objective is to give geometric conditions, in terms of level sets, for a functional to be strictly quasi-convex, and to discuss another kind of convexity which is intermediate between quasi-convexity and strict quasi-convexity. We show that functionals with this property, which we call semi-strict quasi-convexity, are characterized among continuous quasi-convex functionals by the fact that every local minimum is a global minimum.

We begin with the following lemma:

Lemma 1.4.1. Let C be a convex subset of a Banach space.

(a) If $u \in C$ and v is an interior point of C then every point w of the open interval (u,v) is an interior point of C.

(b) If $u,v \in C$ then the open interval (u,v) contains only interior points or only boundary points.

Proof: (a). If v is an interior point of C then there is a sphere S of radius ϵ about v that is contained in C. Let w be any point of the open interval (u,v) and α be that positive number such that $\alpha(u-v) = (u-w)$. Then the sphere of radius $\alpha\epsilon$ about w is in the convex hull of $S \cup \{u\}$ and thus in C. Therefore w is an interior point of C. (b) follows easily from (a).

We next introduce some notation and terminology. If $g:D \subset X \to R$, set $L_c = \{u \in D: g(u) \leqq c\}$, $E_c = \{u \in D: g(u) = c\}$ and B_c equal to the boundary points of L_c.

Definition 1.4.1. The functional g has property S if for any c, E_c contains no line segments, (i.e., if $u,v \in E_c$ then there exists $t_0 \in (0,1)$ such that $t_0 u + (1-t_0)v \notin E_c$.)

Definition 1.4.2. A set S is strictly convex if for any two points $u \neq v$ on its boundary the open interval (u,v) contains only interior points of S.

In view of Lemma 1.4.1, a set S is strictly convex if and only if for any two points $u \neq v$ in its closure the open interval (u,v) contains only interior points of S.

Theorem 1.4.1. Let g be a continuous functional defined on a Banach space X. Then the following are equivalent:

(a) g is strictly quasi-convex;

(b) g is quasi-convex and has property S;

(c) the level sets L_c are strictly convex and $E_c \subset B_c$ for all c.

Proof: Assume (a); g is clearly quasi-convex so we need only prove g has property S. If E_c contained a line segment, g would have constant value c on the segment. But this is impossible for strictly quasi-convex functionals. Thus g has property S. Now assume (b), and let $u,v \in B_c$. L_c is closed and therefore u and v belong to L_c. By Lemma 1.4.1, if (u,v) contained any boundary point it would contain only boundary points. But this interval of boundary points would by continuity lie in E_c and this contradicts property S. Hence L_c is strictly convex. To prove E_c contains only boundary points of L_c assume initially that g is not constant on L_c and thus there is some $u \in L_c$ such that $g(u) < c$. If E_c had an element w that was an interior point of L_c then the line segment $[u,w]$ would have a proper extension $[u,v]$ lying in L_c. But $g(u) < g(w)$, $g(w) \geqq g(v)$ and quasi-convexity of g would then imply g is constant on the line segment $[w,v]$. This however contradicts property S, and E_c can therefore contain no interior points of L_c. In the case where g is constant on L_c, we have $E_c = L_c$. If L_c contained two distinct points $u \neq v$ then by the convexity of L_c the entire interval $[u,v]$ would belong to L_c and hence to E_c. Again this contradicts property S. Finally assume (c); if $u \neq v$, and $g(u) \leqq g(v)$ and $w \in (u,v)$ then by the strict convexity of $L_{g(v)}$ we have $g(w) \leqq g(v)$ and w is an interior point of $L_{g(v)}$. But $E_{g(v)}$ contains no interior points and thus $g(w) < g(v)$.

We note that although the theorem was proved only for functionals defined on the entire space the result still holds for functionals defined on a convex set D provided that all topological notions are interpreted in the relative topology for D. In particular the notion of a boundary used in the definitions of B_c and of a strictly convex set must be relative to D.

<u>Corollary 1</u>. If g is pseudo-convex and has property S then g is strictly pseudo-convex.

<u>Proof</u>: By Theorem 1.3.3 pseudo-convexity implies

quasi-convexity and by the previous theorem this and property S imply strict quasi-convexity of g. If $g(u) \leqq g(v)$ then by the strict quasi-convexity, for any w in (u,v), $g(w) < g(v)$. This and pseudo-convexity imply $g'(v)(v-w) > 0$. Multiplying by $\|v-u\|/\|v-w\|$ gives $g'(v)(v-u) > 0$, and thus g is strictly pseudo-convex.

Mangasarian [25] and Ponstein [35] have considered the condition: for $u,v \in D$, D convex, and $w \in (u,v)$

$$(1.4.1) \qquad g(u) < g(v) \text{ implies } g(w) < g(v),$$

and they call this strict quasi-convexity. This condition is clearly weaker than what we have called strict quasi-convexity and implies quasi-convexity when g is lower semi-continuous. We therefore define:

Definition 1.4.3. A function g defined on a convex set D is semi-strictly quasi-convex if for $u \neq v$, $u,v \in D$ and $w \in (u,v)$, (1.4.1) holds.

Theorem 1.4.2. Assume g is semi-strictly quasi-convex and lower semi-continuous in a convex set D. Then g is quasi-convex in D.

Proof: From (1.4.1) and (1.3.7) we need only show that for $u,v \in D$ and $w \in (u,v)$, $g(u) = g(v)$ implies $g(w) \leqq g(v)$. Assume, to the contrary, that $g(w) > g(v)$ and let $w' \neq w$ belong to (u,v). Then (1.4.1) implies $g(w') < g(w)$. In fact, $g(w') \leqq g(v)$, for if $g(w') > g(v) = g(u)$ and w was in, say, (u,w') then (1.4.1) would imply $g(w) < g(w')$ which is contrary to what we have just shown. Thus $g(w') \leqq g(v)$ on the entire closed interval $[u,v]$ with the exception of the point w. But g is lower semi-continuous and thus $g(w) \leqq g(v)$. Hence g is quasi-convex.

In Theorem 1.4.1 we showed that strict quasi-convexity was equivalent to assuming, for all c, that

(1.4.2) $$E_c \subset B_c ,$$

and that the level sets L_c are strictly convex. We show next that semi-strict quasi-convexity is equivalent to assuming that the level sets L_c are convex and that (1.4.2) holds for $c \neq \min\{g(u):u \in D\}$. As with the previous theorem we will prove the result only for functionals defined on the entire space and note that an extension is possible for functionals defined on a convex domain.

<u>Theorem 1.4.3</u>. A continuous functional g is semi-strictly quasi-convex if and only if for all c, L_c is convex and either $E_c \subset B_c$ or $E_c = L_c$.

<u>Proof</u>: Assume g is semi-strictly quasi-convex. By Theorem 1.4.2 g is quasi-convex and thus has convex level sets. We will show that if E_c contains an interior point of L_c then g is constant on L_c and therefore $L_c = E_c$.

Let w be an interior point of L_c such that $g(w) = c$, and let u be any other point of L_c. The line segment $[u,w]$ has a proper extension $[u,w] \subset [u,v] \subset L_c$ for some $v \neq w$ in L_c and if $g(u) \neq g(v)$, say $g(u) < g(v)$, then semi-strict quasi-convexity implies $g(w) < g(v) \leqq c$. But this is a contradiction and thus $g(u) = g(v)$. The quasi-convexity of g then implies $c = g(w) \leqq g(u) \leqq c$. Thus $g(u) = c$ and since u was arbitrary, g must be constant on L_c.

Conversely if every level set is convex then g is quasi-convex and for all u,v and $w \in (u,v)$

$$g(u) < g(v) \text{ implies } g(w) \leqq g(v).$$

To complete the proof we need only show that $g(w) \neq g(v)$. If $g(u) < g(v)$, then g is not constant on $L_{g(v)}$ and thus $E_{g(v)}$ is equal to the set of boundary points of $L_{g(v)}$. If w is any point of the interval (u,v) then Lemma 1.4.1 shows that w is an interior point. Thus w does not belong to $E_{g(v)}$ and therefore $g(w) \neq g(v)$.

Ponstein [35] has shown that a local minimum of a semi-strictly quasi-convex functional is at the same time a global minimum. We would like to show that in some sense this characterizes them. The next theorem says that among the class of continuous quasi-convex functionals the set of semi-strict quasi-convex functionals are precisely those for which every local minimum is a global minimum.* Again we prove the result only for g defined on the entire space but note that an extension to g defined only on a convex set is possible.

<u>Theorem 1.4.4</u>. Suppose that g is continuous on the Banach space X. If g is semi-strictly quasi-convex then any local minimum is a global minimum. Conversely, if g is quasi-convex and any local minimum is a global minimum then g is semi-strictly quasi-convex.

<u>Proof</u>: Suppose u is a local minimum of g but there is some v such that $g(v) < g(u)$. Then any point w on the line segment (u,v) satisfies $g(w) < g(u)$ and u is not a local minimum. Conversely, by Theorem 1.4.3, it is sufficient to show that whenever E_c contains an interior point of L_c then $E_c = L_c$, i.e., g is constant on L_c. Therefore we assume u is an interior point of L_c and $g(u) = c$. Set $S_c = \{x|g(x)<c\}$ and assume that S_c is not empty. Then by continuity, S_c is open and since $u \notin S_c$ the separation theorem for convex sets in a linear topological space (see, e.g., Dunford - Schwartz [9]) gives a continuous linear functional f such that $f(x) < a$ for $x \in S_c$ and $f(u) = a$. Moreover, since u is an interior point of L_c the line segment $[v,u]$ has a proper extension $[v,u']$ in L_c i.e., $[v,u] \subset [v,u'] \subset$

*B. Martos has independently obtained essentially this result working in finite dimensions (private communications).

L_c $u' \neq u$, such that $f(u') > a$ and by Lemma 1.4.1, u' may be chosen as an interior point of L_c. Since f is continuous there is a neighborhood $N \subset L_c$ of u' such that $f(w) > a$ for $w \in N$. Therefore N contains no points of S_c and N must be a subset of E_c. This implies u' is an interior point of E_c and is thus a local minimum of g. By the hypothesis u' is a global minimum of g and hence g is constant on L_c.

CHAPTER II

DIRECTION ALGORITHMS

2.1 Introduction. The purpose of this chapter and the next is to provide a unified convergence theory for iterative minimization of (generally non-quadratic) functionals. Simply stated the problem is when will an algorithm produce a sequence $\{u^p\}$, satisfying

$$(2.1.1) \qquad u^{p+1} = u^p - t_p e^p, \; t_p \geqq 0, \; \|e^p\| = 1, \; p = 0,1,\dots ,$$

which converges to a minimum of a given functional g or a solution of $g'(x) = 0$. (By a minimum of g we mean, of course, a point at which g attains its minimum value.) We approach this through the related problem of proving that $g'(u^p) \to 0$, and hence, at the outset we discuss when $g'(u^p) \to 0$ implies that the sequence $\{u^p\}$ converges, either to a solution of $g'(u) = 0$, or to a minimum of g. Then, in two steps we shall study when $g'(u^p)$ tends to zero. These two steps reflect a natural division in the problem, since any minimization algorithm can be thought of having two tasks--picking the next direction e^p and choosing the distance (or step-size) t_p. In the next chapter we will analyse a number of step-size algorithms with the object of showing that a suitable choice of step-size implies $g'(u^p)e^p \to 0$. In this chapter, under the assumption that $g'(u^p)e^p$ tends to zero, we consider various methods of choosing the sequence of directions

e^p and demonstrate that they yield iterations for which $g'(u^p)$ tends to zero.

One obvious advantage of separating the analysis of step-size and direction is that any suitable step-size may then be combined with a suitable direction and our results will therefore apply to a great many algorithms.

The underlying space used in this chapter will be the real n-dimensional Euclidean Space E^n with $\|x\| = (\Sigma x_i^2)^{\frac{1}{2}}$, although the results extend easily to other norms, and in many instances (particularly for the gradient-related directions of section 2.3) extensions are possible to an arbitrary Banach space.

Before we begin the analysis let us consider two examples of iterative minimization algorithms. Perhaps the best known algorithm is the method of steepest descent, first proposed by Cauchy [5] in 1847. Suppose $g: E^n \rightarrow R$ is a Gateaux differentiable map. The method of steepest descent uses the gradient direction

(2.1.2) $$e^p = [g'(u^p)]^T / \|g'(u^p)\|,$$

and a step-size t_p satisfying

(2.1.3) $$g(u^p - t_p e^p) = \min_{t \geqq 0} g(u^p - te^p),$$

to obtain

$$u^{p+1} = u^p - t_p e^p .$$

Intuitively, we choose the direction of local maximum decrease of the functional and then from among all points in this direction select one for which the value of the functional is least.

Another minimization algorithm is the Gauss-Seidel method. Again suppose $g: E^n \to R$ is Gateaux differentiable and let $e_0, e_1, \ldots, e_{n-1}$ be the n orthonormal coordinate vectors of E^n; then the Gauss-Seidel directions are given by

$$(2.1.4) \qquad e^p = \operatorname{sgn}(g'(u^p)e_i)e_i \;, \qquad i = p \pmod{n}.$$

Thus the direction sequence consists of the n coordinate vectors repeated cyclically. (The signs are chosen so that $g'(u^p)e^p \geqq 0$ and thus t_p is positive.) The step-size may be determined, for example, by letting t_p be the smallest non-negative solution if it exists, of

$$(2.1.5) \qquad g'(u^p - t_p e^p)e^p = 0,$$

and the next iterate is given by $u^{p+1} = u^p - t_p e^p$.

In these algorithms, the choice of step-size and direction are independent. Thus a step-size t_p defined by (2.1.5) could be used with the directions satisfying (2.1.2). Indeed, this combination was proposed and analysed by Curry [7]. Moreover, if t_p is defined by (2.1.3) and e^p by (2.1.4) we have a variant of the Gauss-Seidel algorithm. Note that t_p given by (2.1.3) is also a non-negative solution of $g'(u^p - te^p)e^p = 0$ and if g has a strictly positive definite second derivative both distance choices coincide.

The name Gauss-Seidel is usually associated with a method of solving a system of simultaneous equations in which one treats the equations cyclically, solving the i^{th} equation for the i^{th} unknown with the remaining variables fixed and immediately substitute the solution for the old estimate of that coordinate. In applying the Gauss-Seidel algorithm to a functional g, we are carrying out precisely this procedure on the equation $g'(x) = 0$.

2.2 Convergence of the Iterates. Even if we know that $g'(u^p) \to 0$ the problem of proving convergence of the sequence $\{u^p\}$ produced by a minimization algorithm is still an open question for arbitrary g. (For some partial results see Ostrowski [33].) However, under the assumptions that $g'(u) = 0$ has a unique solution, that $\{u^p\}$ lies in a compact subset of D, and that $g: D \subset E^n \to R$ is continuously differentiable, it is easy to show that the sequence $\{u^p\}$ converges. For, from the compactness, the sequence will have limit points in D; from the continuity of g' these limit points will be solutions of $g'(u) = 0$; and if there is a unique solution, then the sequence converges. Nonetheless, the requirement that $g'(u) = 0$ have a unique solution is restrictive. A stronger result, due to Ostrowski, is possible when $\|u^p - u^{p+1}\| \to 0$. For completeness we reproduce the proof here.

Theorem 2.2.1. Let $g:D \subset E^n \to R$ have a continuous derivative on an open set D; suppose the sequence $\{u^p\}$ is contained in a compact set, $D_0 \subset D$, and assume that $g'(u^p) \to 0$, $\|u^p - u^{p+1}\| \to 0$, and the set $\{u:g'(u) = 0\}$ consists only of isolated points. Then the sequence $\{u^p\}$ converges to a limit x and $g'(x) = 0$.

Proof: The essential fact is that $\|u^p - u^{p+1}\| \to 0$ implies that the set of limit points of the sequence $\{u^p\}$ is connected. To see this, assume that the set of limit points consisted of two separated sets. We could then find disjoint closed neighborhoods A and B of each of these sets and A and B would have a positive distance ϵ from each other. But the sequence $\{u^p\}$ must eventually lie in $A \cup B$ and $\|u^p - u^{p+1}\| < \epsilon$ for p sufficiently large. Therefore, the sequence will eventually lie entirely in either A or B. This implies that the other neighborhood could contain no limit points of $\{u^p\}$ and the set of limit points of $\{u^p\}$ must be connected. But the set of limit points of $\{u^p\}$ is a subset of the set of solutions of $g'(u) = 0$ and the latter contains only isolated points. Thus the set of limit points contains only isolated points; and, since it is connected, it has precisely one point. Therefore the sequence $\{u^p\}$ converges.

<u>2.3 Gradient-related Directions</u>. In the next chapter we will analyse a variety of methods for choosing step-size with the objective of showing that these algorithms imply $g'(u^p)e^p \to 0$, $p \to \infty$. In the rest of this chapter we will analyse when $g'(u^p) \to 0$ under the assumption that $g'(u^p)e^p \to 0$.

The direction algorithms we consider fall into two classes--those that we think of as generalizations of steepest descent, and those that are generalizations of the Gauss-Seidel directions. For steepest descent itself it is immediate that $g'(u^p)e^p \to 0$ implies $g'(u^p) \to 0$ since $e^p = g'(u^p)^T/\|g'(u^p)\|$ and this $g'(u^p)e^p = \|g'(u^p)\|$. The simplicity of this argument extends to the following class of directions.

We say that a set of directions $\{e^p\}$ is <u>gradient-related</u> if there is a forcing function d such that

$$g'(u^p)e^p \geqq d(\|g'(u^p)\|). \tag{2.3.1}$$

(recall Definition 1.1.1, d is forcing if $d(t) \geqq 0$ and $d(t_n)$ implies $t_n \to 0$.) Clearly if the sequence $\{e^p\}$ is gradient-related and $g'(u^p)e^p \to 0$, then $\|g'(u^p)\| \to 0$.

One technique for obtaining a sequence $\{e^p\}$ satisfying (2.3.1) is to define

$$e^p = A_p[g'(u^p)]^T/\|A_p g'(u^p)^T\|$$

where $\{A_p\}$ is a sequence of symmetric matrices satisfying

$$M\|h\|^2 \geqq h^T A_p h \geqq m\|h\|^2, \quad m > 0, \quad p = 0,1,\ldots$$

Then, since $\|A_p h\| \leqq M\|h\|$ we have

$$g'(u^p)e^p \geqq (m/M)\|g'(u^p)\|$$

and the sequence $\{e^p\}$ is gradient-related. This approach has been exploited either explicitly or implicitly by several authors (see e.g., Nashed [27]). In particular if g has a continuous bounded, uniformly positive definite second Frechet derivative, i.e.,

$$M\|h\|^2 \geqq g''(u)hh \geqq m\|h\|^2, \quad M \geqq m > 0,$$

then Newton's method

$$u^{p+1} = u^p - [g''(u^p)]^{-1}[g'(u^p)]^T$$

produces a gradient-related sequence of directions.

For gradient-related algorithms, therefore, it is immediate that $g'(u^p)e^p \to 0$ implies $g'(u^p) \to 0$ and the effort is directed towards showing they are gradient-related. In Chapter 4 we will give other examples.

2.4 Uniformly Linearly Independent Directions. In contrast to steepest descent and Newton's method the directions of the Gauss-Seidel algorithm are not gradient-related and we therefore need another approach. The significant feature of the Gauss-Seidel directions is that n successive directions are orthogonal. In seeking to generalize this we might consider requiring only that among m successive vectors there are n that are linearly independent. This approach leads to difficulties if, as $p \to \infty$, the directions become "almost" dependent. We therefore define:

Definition 2.4.1. A sequence of vectors $\{e^p\}$, with $\|e^p\| = 1$, is uniformly linearly independent if there is some $m \geqq n$ and $c > 0$ such that for any p' and any $x \in E^n$

$$\max_{p'+1 \leqq p \leqq p'+m} |x^T e^p| \geqq c\|x\|. \tag{2.4.1}$$

It is easy to verify that uniform linear independence of a sequence $\{e^p\}$ is equivalent to the requirement that there is some $c' > 0$ such that from every m successive vectors we may choose n of them which satisfy

$$|\det(e^{p_1},\ldots\ldots,e^{p_n})| \geqq c'.$$

The following theorems give sufficient conditions that, for a uniformly linearly independent sequence of directions, $g'(u^p) \to 0$ whenever $g'(u^p)e^p \to 0$.

Theorem 2.4.1. Let $g: D \subset E^n \to R$ have a continuous Frechet derivative on a compact set $D_0 \subset D$. If, for a sequence $\{u^p\} \subset D_0$ satisfying $u^{p+1} = u^p - t_p e^p$, the sequence of vectors $\{e^p\}$ is uniformly linearly independent, $g'(u^p)e^p \to 0$,

$p \to \infty$, and $t_p \to 0$, $p \to \infty$, then $g'(u^p) \to 0$, $p \to \infty$.

Proof: Let $\epsilon > 0$ be given. Since g' is uniformly continuous on D_0 the function δ defined by (1.1.7), i.e.,

$$\delta(t) = \inf \{\|u-v\| : u,v \in D_0; \|g'(u)-g'(v)\| \geq t\},$$

satisfies $\delta(t) > 0$ for $t > 0$. Therefore, because $\|u^p-u^{p+1}\| \to 0$ and $g'(u^p)e^p \to 0$, we can find a K sufficiently large that

$$\|u^p - u^{p+1}\| \leq \delta(\tfrac{1}{2}\epsilon c)/m, \qquad p \geq K, \tag{2.4.2}$$

and

$$g'(u^p)e^p \leq \tfrac{1}{2}\epsilon c, \qquad p \geq K, \tag{2.4.3}$$

where m and c are the constants of (2.4.1) in the definition of uniform linear independence. From (2.4.2) and the triangle inequality it follows that

$$\|u^p - u^{p+i}\| \leq \delta(\tfrac{1}{2}\epsilon c) \qquad 1 \leq i \leq m, \tag{2.4.4}$$

and then the definition of δ implies that

$$\|g'(u^p)-g'(u^{p+i})\| \leq \tfrac{1}{2}\epsilon c, \quad 1 \leq i \leq m.$$

Therefore for any vector e of norm unity, we have

$$\tfrac{1}{2}\epsilon c \geq |[g'(u^p)-g'(u^{p+i})]e| \geq |g'(u^p)e| - |g'(u^{p+i})e|,$$

and thus

$$\tfrac{1}{2}\epsilon c + |g'(u^{p+i})e| \geq |g'(u^p)e|, \quad 1 \leq i \leq m.$$

It then follows from (2.4.3) that

$$\epsilon c \geq |g'(u^p)e^{p+i}|, \quad 1 \leq i \leq m, \ p \geq K$$

which, with the uniform linear independence of $\{e^p\}$, implies that

$$\varepsilon c \geqq \max_{1\leqq i\leqq m} \{|g'(u^p)e^{p+i}|\} \geqq c\|g'(u^p)\|,$$

and therefore for $p \geqq K$ we have $\|g'(u^p)\| \leqq \varepsilon$. But ε was arbitrary and hence $g'(u^p) \to 0$.

One of the hypotheses of Theorem 2.4.1 was the assumption $\|u^p - u^{p+1}\| \to 0$. For several of the step-size algorithms we will discuss in the next chapter, $\|u^p - u^{p+1}\| \to 0$ follows directly from their definition and $g'(u^p)e^p \to 0$. On the other hand, this is not true for other algorithms without some additional assumptions about the functional, or about both the functional and the algorithm. One such assumption on the algorithm is that for some positive $c \leqq 1$,

(2.4.5) $$g(u^p) \geqq g(tu^p + (1-t)u^{p+1}) \geqq g(u^{p+1}), \ 0 \leqq t \leqq c.$$

If t_p is defined as the smallest non-negative solution of $g'(u^p - te^p)e^p = 0$, for example, the mean value theorem implies that (2.4.5) holds with $c = 1$. We next show that if g has property S and (2.4.5) holds then $\|u^p - u^{p+1}\| \to 0$, and we may apply Theorem 2.4.1. Recall that a functional g has property S (Definition 1.4.1) if for any u, v such that $g(u) = g(v)$ there is a $t_0 \in (0,1)$ such that $g(t_0 u + (1-t_0)v) \neq g(u)$, and that strict convexity, strict quasi-convexity and strict pseudo-convexity of g all imply property S.

<u>Theorem 2.4.2</u>. If a sequence $\{u^p\}$ lying in a compact set satisfies (2.4.5) for continuous g with property S,

then $\|u^p - u^{p+1}\| \to 0$.

Proof: If $\|u^p - u^{p+1}\| \not\to 0$ we can find an $\epsilon > 0$ and a subsequence $\{u^{p_j}\}$ such that $\|u^{p_j} - u^{p_j+1}\| \geqq \epsilon$. Since the sequence $\{u^{p_j}\}$ lies in a compact set we can find a possible finer subsequence $\{u^{p_j'}\}$ such that $u^{p_j'} \to x^*$ and $u^{p_j'+1} \to x^{**}$ and clearly then, $\|x^{**} - x^*\| \geqq \epsilon$. But since $g(u^{p+1}) \leqq g(u^p)$ we must have $g(x^*) = g(x^{**})$ and so by (2.4.5) and the continuity of g,

$$g(x^*) = g(tx^{**} + (1-t)x^*) = g(x^{**}), \quad 0 \leqq t \leqq c.$$

This, however, contradicts property S and hence $\|u^p - u^{p+1}\| \to 0$.

For a number of the step-size algorithms discussed in Chapter 3, (2.4.5) does not hold. However, the assumption of strict pseudo-convexity about the functional allows another approach.

Theorem 2.4.3. Suppose g is a continuously differentiable strictly pseudo-convex functional defined on an open set D and the sequence $\{u^p\}$ lies in a compact set $D_0 \subset D$ and satisfies $g(u^p) \geqq g(u^{p+1})$ and $g'(u^p)e^p \to 0$. Then $\|u^p - u^{p+1}\| \to 0$.

Proof: As in the proof of Theorem 2.4.2, if $\|u^p - u^{p+1}\| \not\to 0$, then we can find a subsequence $\{u^{p_i}\}$ and points $x^* \neq x^{**}$ such that $u^{p_i} \to x^*$, $u^{p_i+1} \to x^{**}$, and $g(x^*) = g(x^{**})$. The strict pseudo-convexity of g then implies that

$g'(x^*)(x^*-x^{**}) > 0$. But $g'(u^p)(u^p-u^{p+1}) = g'(u^p)e^p\|u^p-u^{p+1}\|$ and $\|u^p-u^{p+1}\|$ is bounded while $g'(u^p)e^p$ tends to zero; this and the continuity of g' imply that $g'(x^*)(x^*-x^{**}) = 0$ which is a contradiction. Therefore $\|u^p- u^{p+1}\|$ must tend to zero.

2.5 Free-steering Methods. The classical Gauss-Seidel method uses the n orthonormal coordinate vectors cyclically. The so-called free-steering methods of Ostrowski [31], and Schechter [37], [38], allow the coordinate vectors to appear in any order, requiring only that each appear infinitely often. We abstract the essential elements of this approach in the following definition.

Definition 2.5.1. A sequence of vectors $\{e^p\}$, with $\|e^p\| = 1$, is *free-steering* if the sequence contains only a finite number of distinct elements and, for any N, the set $\{e^p : p \geqq N\}$ spans E^n.

The assumption that a sequence of directions is free-steering is relatively weak, and must be balanced in our next theorem by the strongest assumption yet--uniform pseudo-convexity--about the functional. The theorem is a generalization of a result of Schechter [38], who required a uniformly positive definite second derivative, the coordinate directions for e^p and a particular choice of step-size.

Theorem 2.5.1. Let $g:D \subset E^n \to R$ be uniformly pseudoconvex and have a continuous derivative on an open convex set D. Suppose that the sequence $\{e^p\}$ is free-steering, the sequence $u^{p+1} = u^p - t_p e^p$ remains in a compact subset $D_0 \subset D$, $g(u^{p+1}) \leqq g(u^p)$, $p = 0,1, \ldots$ and $g'(u^p)e^p \to 0$. Then $g'(u^p) \to 0$ and the sequence $\{u^p\}$ converges to the unique minimum of g in D.

Proof: Let x be a limit point of the sequence $\{u^p\}$ and suppose that $g'(x) \neq 0$. We will show that this leads to the contradictory statement $g'(x) = 0$.

Let a be the least positive element of

$$\{|g'(x)e^p| : p = 0,1, \ldots ,\}.$$

Since by the definition of a free-steering sequence, there are only finitely many distinct e^p's and they span E^n, a is well-defined. Let $\delta(t)$ be defined by (1.1.7), i.e.,

$$\delta(t) = \inf\{\|u-v\| : u,v \in D; \|g'(u)-g'(v)\| \geqq t\}$$

and set $r = \delta(\frac{a}{2})$; the uniform continuity of g' ensures that $r > 0$. Let S be an open sphere of radius r about x and $K = \{u : g'(x)(x-u) > \frac{1}{2}rd(r)\}$ where d is the nondecreasing function in the definition of uniform pseudoconvexity:

$$g(u) \leqq g(v) \text{ implies } g'(v)(v-u) \geqq d(\|u-v\|)\|u-v\|.$$

By the definition of a forcing function, $d(r) > 0$ and hence

K is an open half plane such that $x \notin \bar{K}$.

Now suppose $u \in L = \{v : g(v) \leqq g(x)\}$ and $u \notin S$. Then $\|u-x\| \geqq r$ and since $g(u) \leqq g(x)$, the uniform pseudo-convexity of g shows that

$$g'(x)(x-u) \geqq d(\|x-u\|)\|x-u\| \geqq rd(r) > \tfrac{1}{2}rd(r)$$

and thus u belongs to K. Therefore $L \subset S \cup K$. It is easy to see, since $g(u^{p+1}) \leqq g(u^p)$, that the value of g is the same for all limit points of $\{u^p\}$ and thus every limit point of $\{u^p\}$ lies in L. But $S \cup K$ is an open neighborhood of L and, since $\{u^p\}$ is bounded, we can choose an N so large that $p \geqq N$ implies $u^p \in S \cup K$. As $g'(u^p)e^p \to 0$, we may also require that

$$(2.5.1) \qquad |g'(u^p)e^p| < \tfrac{1}{2}a, \qquad p \geqq N.$$

Now let C be the set of points in S that are not in K. We will show next that if $u^p \in C$ for $p \geqq N$, then $g'(x)e^p = 0$ and $u^{p+1} \in C$. For if $\|u^p - x\| < \delta(\tfrac{1}{2}a)$ then the definition of δ implies that $\|g'(u^p) - g'(x)\| \leqq \tfrac{1}{2}a$ which with (2.5.1) gives $|g'(x)e^p| < a$. But a is the least positive element of $\{|g'(x)e^p| : p = 0, 1, \ldots\}$ and thus $g'(x)e^p = 0$. Moreover, since $u^{p+1} = u^p - te^p$, $g'(x)(u^p - u^{p+1}) = 0$ which implies $g'(x)(u^p - x) = g'(x)(u^{p+1} - x)$. Thus, since $u^p \notin K$, it follows that $u^{p+1} \notin K$. Since $u^{p+1} \in S \cup K$, u^{p+1} must belong to C.

Because x is a limit point of $\{u^p\}$ and $x \notin \bar{K}$ the sequence $\{u^p\}$ is in C infinitely often. However, once $u^{p'}$ belongs to C for some $p' \geqq N$, we have just shown that all subsequent u^p also belong to C. But then $g'(x)e^p = 0$, $p \geqq p'$ and the assumption that $\{e^p : p \geqq p'\}$ spans E^n implies that $g'(x) = 0$.

Thus, any limit point x of $\{u^p\}$ must satisfy $g'(x) = 0$ and since, for a uniformly pseudo-convex functional, by Theorem 1.3.1, the unique minimum of g is the only solution of $g'(x) = 0$, the sequence $\{u^p\}$ converges. This completes the proof.

CHAPTER III

STEP-SIZE ALGORITHMS

3.1 Basic Lemmas. In the last chapter we examined the choice of "directions" e^p for a sequence $\{u^p\}$ satisfying

$$(3.1.1) \qquad u^{p+1} = u^p - t_p e^p$$

and obtained convergence results under the crucial assumption that

$$(3.1.2) \qquad \lim_{p\to\infty} g'(u^p)e^p = 0.$$

In this chapter we examine a variety of methods for choosing the step-size t_p and concentrate on proving that $g'(u^p)e^p \to 0$. For given step-size algorithms, we also conclude that the sequence $\{u^p\}$ is well defined. The only assumptions now about the directions will be the normalization conditions

$$(3.1.3) \qquad \|e^p\| = 1, \qquad g'(u^p)e^p \geqq 0.$$

The second condition is merely a convenience which enables us to take t_p as non-negative. We stress that because our analysis is concerned solely with step-size the results will apply to both gradient-related and Gauss-Seidel methods.

Among the step-size algorithms we investigate are ones discussed by Altman [2], Armijo [3], Goldstein [14], [15], [17], Ostrowski [32], and Schechter [38]. In these papers analysis of step-size is interwoven with the discussion of the questions we studied in Chapter 2, and therefore results we attribute to these authors may only be implicitly contained in more complete theorems.

As well as (3.1.3) it will be assumed throughout this chapter, without further explicit mention, that D is an open subset of E^n and the function $g:D \subset E^n \to R$ has a uniformly continuous Frechet derivative in D. We also assume that on the component L_0 of the level set $\{u:g(u) \leqq g(u^\circ)\}$ to which the initial point u° belongs, g is bounded below and that L_0 itself is closed. If the functional is defined in the entire space then L_0 is necessarily closed by the continuity of g. But because g is not assumed to be defined everywhere this explicit assumption is needed.

We also note that all results of this chapter extend trivially to an arbitrary Banach space since once the direction e^p has been selected the task of choosing step-size becomes a one dimensional problem, involving only points on the half-line $\{u^p - te^p : t \geqq 0\}$.

At first glance, one might consider using any step-size algorithm which decreases the value of g. As simple examples even in one dimension show, however, the condition

$$g(u^{p+1}) < g(u^p), \qquad p = 0,1, \ldots$$

does not imply that $g'(u^p)e^p \to 0$. On the other hand, we have the following result.

<u>Lemma 3.1.1</u>. Suppose $\{u^p\} \subset L_0$ and

$$g(u^p) - g(u^{p+1}) \geqq d(g'(u^p)e^p), \quad p = 0,1, \ldots \tag{3.1.4}$$

for some forcing function d. Then $g'(u^p)e^p \to 0$.

<u>Proof</u>: Recall (Definition 1.1.1) that $d:[0,\infty) \to [0,\infty)$ and $d(t_p) \to 0$ only if $t_p \to 0$; hence the sequence $\{g(u^p)\}$ is non-increasing. But $\{u^p\} \subset L_0$ implies that $\{g(u^p)\}$ is bounded below and therefore converges. It follows that $g(u^p) - g(u^{p+1}) \to 0$ and $d(g'(u^p)e^p) \to 0$, which, since d is a forcing function, shows that $g'(u^p)e^p \to 0$.

Showing that (3.1.4) holds, which might be termed the principle of sufficient decrease, is the underlying theme of this chapter. For every step-size algorithm we study we will obtain a relation of the form of (3.1.4) with an appropriate forcing function d.

One method of obtaining estimates like (3.1.4) is what we have called the <u>comparison principle</u>. Suppose, for example, we have already analysed some algorithm and shown it produces a sequence of iterates satisfying (3.1.4). Denote by $\bar{u}^p$ the iterate produced by the algorithm at u^p. To prove that a second algorithm produces iterates satisfying (3.1.4) it is sufficient to show that $g(u^{p+1}) \leqq g(\bar{u}^p)$, $p = 0,1, \ldots$; for then we have

$$g(u^p) - g(u^{p+1}) \geqq g(u^p) - g(\bar{u}^p) \geqq d(g'(u^p)e^p),$$

and the second iteration satisfies (3.1.4).

The sequences $\{u^p\}$ will be defined inductively and it is always necessary to prove that the u^p belong to the domain of definition of g. The following will be our main tool to ensure that all iterates do, in fact, remain in L_0.

Lemma 3.1.2. Suppose $u^p \in L_0$, $g'(u^p)e^p > 0$, and, for $0 < C \leqq \infty$, $I = \{u^p - te^p : 0 \leqq t \leqq C\}$. Then $I \cap L_0$ contains an open interval $\{u^p - te^p : 0 < t < t_0\}$. Moreover, if for every positive t such that $[u^p, u^p - te^p] \subset I \cap L_0$ we have

$$g(u^p - te^p) < g(u^p), \tag{3.1.5}$$

then $I \subset L_0$.

Proof: Since D is open, the continuity of g' and the fact that $g'(u^p)e^p > 0$ ensure the existence of an interval $[0, t_0)$, $t_0 > 0$, such that $g'(u^p - te^p)e^p > 0$ for $t \in [0, t_0)$. Hence, by the mean value theorem, (3.1.5) holds for $t \in [0, t_0)$ and therefore $I \cap L_0$ contains the open interval $(u^p, u^p - t_0 e^p)$. Now $I \cap L_0$ is closed, since L_0 is closed, and if $I \cap L_0$ were a proper subset of I we could write the component of $I \cap L_0$ containing u^p as $[u^p, z]$. Since $z \neq u^p$, then by (3.1.5) $g(z) < g(u^p)$. But z is a boundary point of L_0 and therefore $g(z) = g(u^\circ) \geqq g(u^p)$. This is a contradiction and we have $I \subset L_0$.

3.2 Minimization, Curry, and Altman Step-size Algorithms.

The two classical step-size algorithms,

(3.2.1) $\quad t_p$ = the smallest non-negative solution of

$$g'(u^p - te^p)e^p = 0,$$

due to Curry [7], and minimization on L_0, where $u^p - t_p e^p$ is a point of L_0 such that

(3.2.2) $\quad g(u^p - t_p e^p) = \min\{g(u^p - te^p) : t \geqq 0;\ u^p - te^p \in L_0\},$

are in general different for non-quadratic functionals. (They will coincide if g is strictly convex or even strictly pseudo-convex.) Each has been investigated by various authors for particular choices of direction. We will obtain results for these and other step-size algorithms as corollaries of an analysis of the following general algorithm. For fixed $0 \leqq q < 1$, define $a_p = 0$ if $g'(u^p)e^p = 0$ and otherwise

(3.2.3) $\quad a_p = \sup\{t : 0 \leqq r < t \text{ implies } g'(u^p - re^p)e^p > qg'(u^p)e^p\}.$

Then set

(3.2.4) $$t_p = w_p a_p$$

where $1 \geqq w_p \geqq d_1(g'(u^p)e^p)$ for a fixed function d_1 which forces its argument to zero. (In particular, $d_1(t)$ may be constant.) When $q \equiv 0$ and $w_p \equiv 1$ this is the Curry algorithm; for $q > 0$, $u^p - a_p e^p$ is the first point along the line $\{u^p - te^p : t \geqq 0\}$ at which the slope is the fraction

q of the value at $t = 0$. This algorithm is a variation of the following algorithm proposed by Altman [2] for gradient directions. We choose a fixed $C > 0$ and $0 \leqq q < 1$ and set

$$(3.2.5) \qquad a_p = \sup\{t : 0 \leqq r < t \leqq Cg'(u^p)e^p \text{ implies } g'(u^p - re^p)e^p \geqq qg'(u^p)e^p\}.$$

Then we may (although Altman didn't) use a relaxation factor w_p and define t_p by (3.2.4) It is possible to define a version of (3.2.3) with a bound $t \leqq Cg'(u^p)e^p$ or a version of (3.2.5) without the bound and to analyse these modifications in a way analogous to the following results for (3.2.3) and (3.2.5).

<u>Theorem 3.2.1</u>. Assume that L_0 is bounded and for fixed $0 \leqq q < 1$ let $\{t_p\}$ be defined by (3.2.4) and (3.2.3). Then the iterates $u^{p+1} = u^p - t_p e^p$ are well-defined, remain in L_0, and $g'(u^p)e^p \to 0$.

<u>Proof</u>: Suppose that $u^p \in L_0$. If $g'(u^p)e^p = 0$, then $u^{p+1} = u^p$. If $g'(u^p)e^p > 0$ then a_p is positive (and possibly $= \infty$) and we set $I = \{u^p - te^p : 0 \leqq t \leqq a_p\}$. Let $u \neq u^p$ be any point in $I \cap L_0$ such that $[u^p, u] \subset I \cap L_0$. By Lemma 3.1.2 such points u exist. From the mean value theorem and (3.2.3) we have

$$(3.2.6) \qquad g(u^p) - g(u) = g'(v)(u^p - u) > qg'(u^p)e^p\|u^p - u\| \geqq 0$$

for some $v \in (u^p, u) \subset I \cap L_0$ and thus $g(u) < g(u^p)$. By Lemma 3.1.2 we then have $I \subset L_0$ and since L_0 is bounded, a_p is finite. Hence, since $w_p \leqq 1$, $u^{p+1} \in L_0$ and it follows by induction that $\{u^p\} \subset L_0$.

To establish (3.1.4) we initially assume $q > 0$. Letting $u = u^p - w_p a_p e^p$ in (3.2.6) and using (3.2.4) we have

$$\begin{aligned} g(u^p) - g(u^{p+1}) &> qg'(u^p)e^p w_p a_p \\ &\geqq qg'(u^p)e^p\; d_1(g'(u^p)e^p)a_p \;. \end{aligned}$$

Setting $\overline{u}^p = u^p - a_p e^p$, it follows from (3.2.3) that $g'(\overline{u}^p)e^p = qg'(u^p)e^p$, and then for the function δ defined by (1.1.7), i.e., $\delta(t) = \sup\{\|u-v\| : \|g'(u) - g'(v)\| \geqq t; u,v \in D\}$, we have

$$a_p = \|u^p - \overline{u}^p\| \geqq \delta(\|g'(u^p) - g'(\overline{u}^p)\|).$$

Since $\|e^p\| = 1$, and δ is monotonic

$$\begin{aligned} \delta(\|g'(u^p) - g'(\overline{u}^p)\|) &\geqq \delta(|[g'(u^p) - g'(\overline{u}^p)]e^p|) \\ &= \delta((1-q)\; g'(u^p)e^p). \end{aligned}$$

and thus

(3.2.8) $$a_p \geqq \delta((1-q)\; g'(u^p)e^p).$$

As usual, the uniform continuity of g' implies that δ is a forcing function. For $q > 0$ we have therefore established (3.1.4) with

(3.2.9) $$d(t) = qtd_1(t)\delta((1-q)t),$$

and d, as the product of forcing functions qt, $d_1(t)$, and $\delta((1-q)t)$, is again a forcing function.

If $q = 0$ we use the comparison principle to establish (3.1.4). Let a_p be defined by (3.2.3) with $q = 0$ and a'_p with $q = \frac{1}{2}$. Set $u^{p+1} = u^p - w_p a_p e^p$ and $\overline{u}^p = u^p - w_p a'_p e^p$. Using the mean value theorem we have

$$g(\overline{u}^p) - g(u^{p+1}) = g'(u)(\overline{u}^p - u^{p+1}) = g'(u)e^p \|\overline{u}^p - u^{p+1}\|$$

for some $u \in (\overline{u}^p, u^{p+1})$ which therefore satisfies $g'(u)e^p > 0$. Thus $g(u^{p+1}) < g(\overline{u}^p)$. This and (3.2.9) with $q = \frac{1}{2}$ give

$$g(u^p) - g(u^{p+1}) > \tfrac{1}{2}g'(u^p)e^p\delta(\tfrac{1}{2}g'(u^p)e^p)d_1(g'(u^p)e^p)$$

which is of the form of (3.1.4). Applying Lemma 3.1.1 we may conclude for $0 \leqq q < 1$ that $g'(u^p)e^p \to 0$. This completes the proof.

We note that we may remove the assumption that L_0 is bounded if $q > 0$ since then our standard assumption that g is bounded below on L_0 suffices to guarantee that a_p is well-defined. For the algorithm (3.2.5), no assumption of boundedness of L_0 is needed even if $q = 0$ as we next show.

<u>Theorem 3.2.2</u>. Let fixed constants $0 \leqq q < 1$ and $0 < C$ be given, and assume $\{t_p\}$ is defined by (3.2.5) and (3.2.4). Then the iterates $u^{p+1} = u^p - t_p e^p$ are well-defined and remain in L_0, $g'(u^p)e^p \to 0$, $p \to \infty$, and $\|u^p - u^{p+1}\| \to 0$, $p \to \infty$.

Proof: The proof is similar to that of Theorem 3.2.1 and we only point out the differences. Firstly, it is the bound $t \leqq Cg'(u^p)e^p$ rather than the boundedness of L_0 that guarantees the existence of the supremum in (3.2.5). Therefore u^{p+1} is well-defined although we do not yet know it is in L_0. However, assume for $t_0 > 0$ that the interval $[u^p, u^p - t_0 e^p]$ is a subset of L_0. Then for any $0 < t < t_0$ there is by Lemma 3.1.2 a $0 < t' < t$ such that $g(u^p - t'e^p) < g(u^p)$. By the mean value theorem and (3.2.5)

$$\begin{aligned} g(u^p - t'e^p) - g(u^p - te^p) &= g'(u^p - t''e^p)e^p(t - t') \\ &\geqq q\, g'(u^p)e^p(t - t') \qquad (3.2.10) \\ &\geqq 0. \end{aligned}$$

Thus $g(u^p) > g(u^p - te^p)$, and we can apply Lemma 3.1.2 to conclude $u^{p+1} \in L_0$.

To derive an inequality of the form of (3.1.4) we again assume initially that $q > 0$. As in the proof of Theorem 3.2.1 we have (3.2.7). Only the lower bound on the norm of $u^p - \overline{u}^p$ differs. For the iteration given by (3.2.5) either $\|u^p - \overline{u}^p\| = Cg'(u^p)e^p$ or $g'(\overline{u}^p)e^p = qg'(u^p)e^p$ and (3.2.8) holds. Thus

$$\|u^p - \overline{u}^p\| \geqq \min\{Cg'(u^p)e^p,\ \delta((1-q)g'(u^p)e^p)\}, \qquad (3.2.11)$$

and for $q > 0$ we have established (3.1.4) with

$$d(t) = qtd_1(t) \cdot \min\{Ct, \delta((1-q)t)\}. \qquad (3.2.12)$$

d forces its argument to zero and enables us to conclude that $g'(u^p)e^p \to 0$. The case where $q = 0$ is handled by the comparison principle just as it was in the proof of Theorem 3.2.1. This concludes the proof.

Recall that in Theorems 2.3.1 and 2.4.1, a key assumption was that $\|u^p - u^{p+1}\| \to 0$. Thus the combination of Theorems 2.4.1 and 3.3.2 allows us, for this algorithm, to conclude that $g'(u^p) \to 0$ for the Gauss-Seidel directions without any convexity assumption. On the other hand, the combination of 3.3.2 with 2.3.1 shows that $g'(u^p) \to 0$ if only the solutions of $g'(u) = 0$ are isolated.

If a_p is defined using the following variation of (3.2.3)

(3.2.13) $a_p = \sup\{t : 0 \leqq r < t \text{ implies } g'(u^p - re^p)e^p \geqq qg'(u^p)e^p\}$

for fixed $0 \leqq q < 1$ then the resulting iteration is well-defined and $g'(u^p)e^p \to 0$. We state this as a corollary and omit the proof.

<u>Corollary 1</u>. Assume L_0 is bounded, and for fixed $0 \leqq q < 1$ let t_p be given by (3.2.4) and (3.2.13). Then the iterates $u^{p+1} = u^p - t_p e^p$ are well-defined, remain in L_0, and $g'(u^p)e^p \to 0$.

With the aid of the comparison principle we obtain a theorem on the minimization algorithm (3.2.2) essentially as a corollary of Theorem 3.2.1.

<u>Theorem 3.2.3</u>. Assume L_0 is bounded and let t_p be defined by (3.2.2). Then the iterates $u^{p+1} = u^p - t_p e^p$ are defined, remain in L_0 and satisfy $g'(u^p)e^p \to 0$.

<u>Proof</u>: Assume $u^p \in L_0$. Since L_0 is closed and bounded the set $\{u^p - te^p : t \geqq 0\} \cap L_0$ is compact, and t_p is well-defined. Clearly, $u^{p+1} \in L_0$. Now let a_p be defined by (3.2.3) with $0 < q < 1$ and set $\overline{u}^p = u^p - a_p e^p$. Then by definition $g(u^{p+1}) \leqq g(\overline{u}^p)$. With (3.2.7) and (3.2.8) this implies

$$g(u^p) - g(u^{p+1}) \geqq g(u^p) - g(\overline{u}^p) \geqq qg'(u^p)e^p\delta((1-q)g'(u^p)e^p)$$

so that by Lemma 3.1.1 we have $g'(u^p)e^p \to 0$.

We need not restrict ourselves to minimization on a component of the level set, L_0. If we assume the entire level set is closed and bounded, then by continuity g is bounded below on the entire set and we obtain the same result for a step-size algorithm which chooses the minimum from the entire half-line $\{u^p - te^p : t \geqq 0\}$.

Another minimization algorithm is to choose some fixed $C > 0$ and let $u^{p+1} = u^p - t_p e^p \in L_0$ satisfy

$$g(u^{p+1}) = \min\{g(u^p - te^p) : 0 \leqq t \leqq Cg'(u^p)e^p ; u^p - te^p \in L_0\}.$$

An argument that the iterates are well-defined and satisfy $g'(u^p)e^p \to 0$ would parallel the proof of Theorem 3.2.3 but use the comparison principle with (3.2.12) instead of (3.2.9).

Moreover the result would follow without the assumption that L_0 is bounded, and we have in addition that $\|u^p - u^{p+1}\| \to 0$.

3.3 Using One Newton Step. To carry out the Curry algorithm (3.2.1) it is necessary to solve the one dimensional non-linear equation

$$h(t) = g'(u^p + te^p)e^p = 0. \tag{3.3.1}$$

A standard one dimensional method for solving (3.3.1) would be the Newton algorithm. We now analyse the tactic of taking only one Newton step from u^p towards solving (3.3.1). Since

$$h'(t) = g''(u^p + te^p)e^p e^p$$

this yields the iteration

$$u^{p+1} = u^p - \frac{g'(u^p)e^p}{g''(u^p)e^p e^p} e^p \tag{3.3.2}$$

which was first suggested by Cauchy [5] for gradient directions if u° is sufficiently close to a minimum. More recently, a local convergence theorem for the more general algorithm

$$u^{p+1} = u^p - w_p \frac{g'(u^p)e^p}{g''(u^p)e^p e^p} e^p \tag{3.3.3}$$

using Gauss-Seidel directions and $0 < w_p < 2$ was given by Ortega and Rockoff [30]. For the same directions Schechter [33] gives a global convergence theorem for $0 < w_p < 2\gamma$, where γ may be small if u° is far from a minimum. We will now prove a theorem which contains a local convergence result for the full range of $w_p (0 < w_p < 2)$, a global convergence result for a suitably small range of w_p, and for the quadratic minimization problems yields global convergence for the entire interval $0 < w_p < 2$.

For a given set of directions $\{e^p\}$ and a given initial vector u°, let

$$(3.3.4) \qquad \alpha_p = \left\{\frac{g''(u)e^pe^p}{g''(u-te^p)e^pe^p} : [u,u-te^p] \subset L_0\right\}$$

and set

$$(3.3.5) \qquad \gamma = \inf_p\{\inf_{t,u} \alpha_p\}.$$

Theorem 3.3.1. Suppose that in L_0 g'' exists and satisfies for $u \in L_0$,

$$(3.3.6) \qquad c\|h\|^2 \leqq g''(u)h,h \leqq C\|h\|^2, \qquad 0 < c \leqq C.$$

Let γ be defined by (3.3.5). Then γ is positive and if w_p satisfies

$$(3.3.7) \qquad d_1(g'(u^p)e^p) \leqq w_p \leqq 2\gamma - d_1(g'(u^p)e^p)$$

for a forcing function $d_1(t) \leqq \gamma$, (in particular d_1 may be a positive constant) then the iteration (3.3.3) is well-defined, the sequence $\{u^p\}$ remains in L_0, $g'(u^p)e^p$ tends to zero, and $\|u^p - u^{p+1}\| \to 0$.

Proof: Suppose $u^p \in L_0$. If $g'(u^p)e^p = 0$ then $u^{p+1} = u^p$. Otherwise $u^p \neq u^{p+1}$ and to show $[u^p,u^{p+1}] \subset L_0$ we must establish (3.1.5). Let $0 < w < 2\gamma$ be any number for which $[u^p,u_w] \subset L_0$ where

$$u_w = u^p - \frac{w\, g'(u^p)e^p}{g''(u^p)e^pe^p}\, e^p$$

(The existence of such w is guaranteed by Lemma 3.1.1.)

We have, by the mean value theorem

$$(3.3.8)\qquad g(u_w)-g(u^p) = g'(u^p)(u_w-u^p) + \tfrac{1}{2}g''(v)(u_w-u^p,u_w-u^p)$$

for some v in the interval $(u^p,u_w) \subset L_0$. Substituting the definition of u_w into (3.3.8), we obtain

$$g(u_w)-g(u^p) = \frac{-(g'(u^p)e^p)^2 w}{g''(u^p)e^pe^p} + \frac{w^2 g''(v)e^pe^p(g'(u^p)e^p)^2}{2(g''(u^p)e^pe^p)^2}$$

$$= \frac{(g'(u^p)e^p)^2 w}{g''(u^p)e^pe^p}\left(\tfrac{1}{2}w\left(\frac{g''(v)e^pe^p}{g''(u^p)e^pe^p}\right)-1\right).$$

But

$$\frac{g''(u^p)e^pe^p}{g''(v)e^pe^p} \geqq \gamma$$

and thus

$$g(u^p) - g(u_w) \geqq \frac{(g'(u^p)e^p)^2 w}{g''(u^p)e^pe^p}\quad(1-\tfrac{1}{2}w/\gamma).$$

Therefore, for any $0 < w < 2\gamma$ for which $[u^p,u_w] \subset L_0$ we have $g(u_w) < g(u^p)$ and Lemma 3.1.2 implies that the entire interval $[u^p,u^{p+1}]$ belongs to L_0. Setting $w = w_p$ for w_p satisfying (3.3.7) we have

$$(3.3.9)\qquad g(u^p)- g(u^{p+1}) \geqq \frac{(g'(u^p)e^p)^2\ \ d_1(g'(u^p)e^p)^2}{2\gamma\ \ g''(u^p)e^pe^p}$$

$$\geqq \frac{(g'(u^p)e^p)^2 d_1(g'(u^p)e^p)^2}{2C\gamma}$$

which establishes (3.1.4) for $d(t) = t^2d_1^2(t)/2C\gamma$, and by Lemma 3.1.1, $g'(u^p)e^p \to 0$. Moreover, (3.3.6) and (3.3.3) imply that $\|u^p-u^{p+1}\| \leqq \frac{2\gamma}{C}\, g'(u^p)e^p$ so that $\|u^p-u^{p+1}\| \to 0$ whenever $g'(u^p)e^p \to 0$. This completes the proof.

<u>Corollary 1</u>. Let g satisfy the conditions of Theorem 3.3.1 and for $0 < \epsilon \leqq \frac{c}{C}$ let $\epsilon \leqq w_p \leqq \frac{2c}{C} - \epsilon$. Then the conclusions of the theorem hold.

<u>Proof</u>: For g'' satisfying (3.3.6) it is immediate that $\gamma \geqq \frac{c}{C}$.

<u>Corollary 2</u>. Suppose x^* is an interior local minimum of g in D and, in a neighborhood N of x^*, g'' exists, is continuous and satisfies (3.3.6). Then for any $0 < \epsilon \leqq 1$, if u° is sufficiently close to x^*, w_p may be taken to satisfy

$$\epsilon \leqq w_p \leqq 2-\epsilon$$

and the conclusions of the theorem follow.

<u>Proof</u>: It suffices to show that if u° is sufficiently close to x^* then $\gamma \geqq 1-\frac{1}{2}\epsilon$. From the mean value theorem and (3.3.6) we have

$$\tfrac{1}{2}c\|u-x^*\|^2 \leqq g(u) - g(x^*) \leqq \tfrac{1}{2}C\|u-x^*\|^2$$

and therefore for any $u \in L_o$,

$$\tfrac{1}{2}C\|u^\circ-x^*\|^2 \geqq g(u^\circ)-g(x^*) \geqq g(u)-g(x^*) \geqq \tfrac{1}{2}c\|u-x^*\|^2.$$

Hence L_o lies in a sphere about x^* of radius $(\frac{C}{c})^{\frac{1}{2}}\|u^\circ-x^*\|$. Since g'' is uniformly continuous in a neighborhood N of x^* there is an $r > 0$ such that the sphere $S(x^*,r)$ of radius r about x^* lies in N and $u,v \in S(x^*,r)$ implies that $\|g''(u)-g''(v)\| \leqq \frac{1}{2}\epsilon c$. If $\|u^\circ-x^*\| \leqq (\frac{c}{C})^{\frac{1}{2}}r$ then L_o lies in the $S(x^*,r)$. Moreover, if $u,v \in S(x^*,r)$ we have

$$|g''(v)e^pe^p - g''(u)e^pe^p| \leqq \|g''(v) - g''(u)\| \leqq \tfrac{1}{2}\epsilon c$$

and since $c \leqq g''(u)e^pe^p$ we have that

$$\frac{g''(u)e^pe^p}{g''(v)e^pe^p} \geqq \frac{g''(u)e^pe^p}{g''(u)e^pe^p + \frac{1}{2}\epsilon c}$$

$$\geqq 1 - \frac{\epsilon c}{2g''(u)e^pe^p}$$

$$\geqq 1 - \tfrac{1}{2}\epsilon$$

Therefore, if u^o is sufficiently close to x^*, $\gamma \geqq 1-\frac{1}{2}\epsilon$.

Theorem 3.3.1 is optimum in the sense that it contains the best known result for the quadratic minimization problem. In fact we have a generalization to arbitrary directions of the relaxation results for free-steering methods discussed by Schechter [37].

<u>Corollary 3</u>. If g'' is constant and strictly positive definite and w_p satisfies

$$d_1(g'(u^p)e^p) \leqq w_p \leqq 2-d_1(g'(u^p)e^p)$$

for some forcing function $d_1(t) \leqq 1$ then the iteration (3.3.3) is well-defined, remains in L_0, $g'(u^p)e^p \to 0$, and $\|u^p-u^{p+1}\| \to 0$.

<u>Proof</u>: The proof follows from Theorem 3.3.1 and the observation that γ is unity for g'' constant and strictly positive definite.

<u>Corollary 4</u>. (Schechter [38]): Suppose that $\{e^p\}$ consists only of coordinate directions and the definition (3.3.4), (3.3.5) of γ is replaced by

(3.3.10)
$$\gamma = \min_{1\le i\le n} \left\{ \inf_{u\in L_o} a_{ii}(u) \;/\; \sup_{u\in L_o} a_{ii}(u) \right\},$$

where $(a_{ij}(u))$ is the matrix representation of $g''(u)$. Then the statement of Theorem 3.3.1 remains valid.

The proof is immediate from the observation that, when $\{e^p\}$ consists of coordinate directions, the γ defined by (3.3.4),(3.3.5) is at least as great as that defined by (3.3.10). Note that the obvious extension of (3.3.10) to arbitrary directions,

$$(3.3.11)\qquad \gamma = \inf_{p} \left\{ \inf_{u\in L_o} g''(u)e^pe^p \;/\; \sup_{u\in L_o} g''(u)e^pe^p \right\},$$

is, in general, smaller and hence less satisfactory than the γ of (3.3.5)

To conclude this section we consider a variation of

$$(3.3.3)\qquad u^{p+1} = u^p - w_p \frac{g'(u^p)e^p}{g''(u^p)e^pe^p} e^p$$

in which $g''(u^p)e^pe^p$ is replaced by a constant C, thus giving

$$(3.3.12)\qquad u^{p+1} = u^p - \left(\frac{w_p}{C}\right)(g'(u^p)e^p)\; e^p.$$

To carry out this iteration it is necessary to know a Lipschitz constant or a bound on the second derivative, but it does not require knowing the functional g, or solving a one-dimensional nonlinear equation.

This iteration was first proposed by Goldstein [14] with C as an upper bound on the norms of $g''(u^p)$ and is also discussed by Ostrowski [32]. Altman [2] and Armijo [3] consider this iteration when C is a Lipschitz constant for g' and g need only have one derivative. The following is a minor extension of their results.

Theorem 3.3.2. Suppose that

$$\|g'(u)-g'(v)\| \leqq C\|u-v\|, \qquad u,v \in L_0$$

and that for some forcing function $d_1(t) \leqq 1$,

$$d_1(g'(u^p)e^p) \leqq w_p \leqq 2-d_1(g'(u^p)e^p) \tag{3.3.13}$$

(in particular d_1 may be a positive constant). Then the iteration given by (3.3.12) is well-defined, the iterates $\{u^p\}$ remain in L_0, $g'(u^p)e^p \to 0$, and $\|u^p - u^{p+1}\| \to 0$.

Proof: The proof parallels that of Theorem 3.3.1. Suppose $u^p \in L_0$. If $g'(u^p)e^p = 0$, then $u^{p+1} = u^p$. If $g'(u^p)e^p > 0$ then $u^p \neq u^{p+1}$ and we show first that (3.1.5) holds. Let $u_w = u^p - (wg'(u^p)e^p/C)e^p$ for any $w_p > w > 0$ such that $[u^p, u_w] \subset L_0$. By the mean value theorem we obtain

$$\begin{aligned} g(u^p) - g(u^p - te^p) &\\ &= tg'(u^p)e^p - \int_0^1 [g'(u^p) - g'(u^p - ste^p)]e^p ds \\ &\geqq tg'(u^p)e^p - \tfrac{1}{2}Ct^2 \end{aligned} \tag{3.3.14}$$

which is positive provided $t < 2g'(u^p)e^p/C$. Thus

$g(u_w) < g(u^p)$ and by Lemma 3.1.2, $[u^p, u^{p+1}] \subset L_0$. Setting $t = w_p g'(u^p)e^p / C$ in (3.3.14) and combining terms we have

$$g(u^p) - g(u^{p+1}) \geqq w_p(2-w_p)(g'(u^p)e^p)^2/2C$$

which with (3.3.13) yields (3.1.4) with

$$d(t) = \tfrac{1}{2}t^2 d_1(t)^2 / C.$$

Thus Lemma 3.1.1 implies $g'(u^p)e^p \to 0$ and it follows directly from the definition of u^p that $\|u^p - u^{p+1}\| \to 0$.

3.4 Over-relaxed Curry Iteration. We have seen, in section 3.2, that the Curry iteration,

$$(3.4.1) \qquad u^{p+1} = u^p - w_p a_p e^p,$$

for a_p equal to the smallest non-negative solution of

$$(3.4.2) \qquad g'(u^p - te^p)e^p = 0$$

may be under-relaxed, i.e., $w_p < 1$, with no difficulty. For quadratic g, the Curry iteration coincides with taking one Newton step towards solving (3.4.2) and our result in section 3.3 indicates that in this case $0 < w_p < 2$ may be used. We will now consider when, for non-quadratic g, we may take $w_p > 1$ in (3.4.1). The result for quadratic g is a consequence of the symmetry of the one-dimensional functional $h(t) = g(u^p - te^p)$ about its minimum and to extend it to non-quadratic g we must measure the deviation from this symmetry. Therefore, suppose $u^p \in L_0$, a_p given by (3.4.2) is well-defined and set $\bar{u}^p = u^p - a_p e^p$. Define

$$(3.4.3) \qquad \mathcal{A}_p = \{ \left| \frac{g'(\bar{u}^p - te^p)e^p}{t} \right| : t \neq 0 \text{ and } g(\bar{u}^p - te^p) \leqq g(u^p) \},$$

and set

$$(3.4.4) \qquad \gamma_p = \inf \mathcal{A}_p / \sup \mathcal{A}_p .$$

Note that if g is twice continuously differentiable and $t \neq 0$, then

$$(3.4.5) \qquad \left| \frac{g'(\bar{u}^p - te^p)e^p}{t} \right| = |g''(\bar{u}^p - se^p)e^p e^p|$$

for some $0 < s < t$; if g is quadratic, γ_p must therefore be unity. We shall now give the following general result and then consider various corollaries.

<u>Theorem 3.4.1</u>. Suppose that L_0 is bounded, γ_p is defined by (3.4.4) and

$$(3.4.6) \qquad 1 \leqq w_p \leqq 1 + \sqrt{\gamma_p}(1-\epsilon_p)$$

where for some forcing function d_1,

$$1 \geqq \epsilon_p \geqq d_1(g'(u^p)e^p).$$

Then the iterates (3.4.1) are well-defined, lie in L_0 and $g'(u^p)e^p \to 0$.

<u>Proof</u>: Assume $u^p \in L_0$, and consider the case when $g'(u^p)e^p > 0$. Let $\hat{u}^p = u^p - \epsilon_p a_p e^p$. To apply Lemmas 3.1.1 and 3.1.2 we must verify (3.1.5) and then (3.1.4). Therefore set $u_w = u^p - wa_p e^p$ for any $0 < w \leqq w_p$ such that $[u^p, u_w] \subset L_0$. By the argument of Theorem 3.2.1 $\bar{u}^p \in L_0$,

and

$$g(u^p) - g(\hat{u}^p) \geqq d(g'(u^p)e^p) \tag{3.4.7}$$

for some forcing function d. But

$$\begin{aligned} g(u^p) &- g(u_w) \\ &= (g(u^p) - g(\hat{u}^p)) + (g(\hat{u}^p) - g(\overline{u}^p)) + (g(\overline{u}^p) - g(u_w)), \end{aligned} \tag{3.4.8}$$

and by the mean value theorem,

$$\begin{aligned} g(\hat{u}^p) - g(\overline{u}^p) &= \int_0^1 g'(\overline{u}^p + t(\hat{u}^p - \overline{u}^p))(\overline{u}^p - \hat{u}^p)\,dt \\ &= \int_0^1 \frac{g'(\overline{u}^p + t(1-\epsilon_p)a_p e^p)((1-\epsilon_p)a_p)^2 te^p dt}{t(1-\epsilon_p)a_p} \\ &\geqq \inf \mathcal{a}_p \tfrac{1}{2}(1-\epsilon_p)^2 a_p. \end{aligned} \tag{3.4.9}$$

The mean value theorem also gives

$$\begin{aligned} g(u_w) - g(\overline{u}^p) &= \int_0^1 g'(\overline{u}^p + t(u_w - \overline{u}^p))(u_w - \overline{u}^p)\,dt \\ &= \int_0^1 \frac{g'(\overline{u}^p + t(w-1)a_p e^p)((w-1)a_p)^2 te^p dt}{t(w-1)a_p} \\ &\leqq \sup \mathcal{a}_p \tfrac{1}{2}(w-1)^2\ a_p^{\,2}. \end{aligned} \tag{3.4.10}$$

Combining (3.4.8) and (3.4.9) we have

$$g(\hat{u}^p) - g(u_w) \geqq \inf \mathcal{a}_p \tfrac{1}{2}(1-\epsilon_p)^2 a_p - \sup \mathcal{a}_p \tfrac{1}{2}(w-1)^2 a_p$$

which will be non-negative if

$$(\inf \mathcal{a}_p \ /\ \sup \mathcal{a}_p)(1-\epsilon_p)^2 \geqq (1-w)^2.$$

Since $1 \leqq w_p \leqq 1 - \sqrt{\gamma}(1-\epsilon_p)$ we have from (3.4.7) and (3.4.8)

(3.4.11) $$g(u^p) - g(u_w) \geqq d(g'(u^p)e^p) > 0.$$

Thus Lemma 3.1.2 shows $u^{p+1} \in L_0$, and by Lemma 3.1.1 and (3.4.11) with $w = w_p$ we have $g'(u^p)e^p \to 0$.

Without further assumptions on g it is possible that γ_p may be zero; perhaps because $g'(u^p - te^p)e^p = 0$ has solutions other than a_p, perhaps because $g''(\overline{u}^p)e^pe^p = 0$. If $\gamma_p = 0$ then $w_p = 1$ and we have not extended Theorem 3.2.2. However, there are important cases when relaxation factors greater than unity may be used.

The simplest case is when g'' is a constant, strictly positive definite matrix. Then $\gamma_p \equiv 1$, $0 < w_p < 2$ may be used, and we have reproved Corollary 3 of Theorem 3.3.1. We next consider the case when g' is Lipschitz continuous, i.e.,

(3.4.12) $$\|g'(u) - g'(v)\| \leqq C\|u-v\|$$

and uniformly pseudo-convex with a linear forcing function d, i.e.,

(3.4.13) $$g(u) \leqq g(v) \text{ implies } g'(v)(v-u) \geqq c\|u-v\|^2, \; c > 0.$$

<u>Corollary 1</u>. Assume that the conditions of Theorem 3.4.1 as well as (3.4.12) and (3.4.13) hold. Then the conclusions of the theorem are valid for w_p satisfying

$$1 \leqq w_p \leqq 1 + (c/C)^{\frac{1}{2}}(1-\epsilon_p).$$

<u>Proof</u>: For a_p given by (3.4.3) an easy calculation

gives $c \leqq \inf \alpha_p$ and $C \geqq \sup \alpha_p$ and thus $\gamma_p \geqq c/C$ for all p.

If g'' exists and satisfies

$$c\|h\|^2 \leqq g''(u)hh \leqq C\|h\|^2, \qquad u \in D \tag{3.4.14}$$

then (3.4.12) and (3.4.13) follow. With (3.4.14) and the assumption that g'' is uniformly continuous we can get a much better result locally, just as we did with Theorem 3.3.1.

Corollary 2. Suppose x^* is an interior local minimum of g in D and, in a neighborhood N of x^*, g'' exists, is continuous and satisfies (3.4.13), for u belonging to N. Then for any $0 < \epsilon \leqq 1$, if u^0 is sufficiently close to x^*, w_p may be taken to satisfy

$$\epsilon \leqq w_p \leqq 2 - \epsilon$$

and the conclusions of the theorem hold.

Proof: From (3.4.5) it follows that

$$\begin{aligned} \inf \alpha_p &\leqq \inf\{g''(u)e^pe^p: u \in L_0\} \\ &\leqq \sup\{g''(u)e^pe^p: u \in L_0\} \\ &\leqq \sup \alpha_p \end{aligned}$$

and thus

$$\gamma_p = \inf \alpha_p / \sup \alpha_p \leqq \frac{\inf\{g''(u)e^pe^p : u \in L_0\}}{\sup\{g''(u)e^pe^p : u \in L_0\}}$$

In the proof of corollary 2 of Theorem 3.3.1 we showed that if u^p was sufficiently close to x^*, L_0 lay within a sphere $S(x^*,r)$ and that for any two points u,v of $S(x^*,r)$

$$|g''(u)e^pe^p - g''(v)e^pe^p| \leq \tfrac{1}{2}\epsilon c.$$

Hence, if $L_0 \subset S(x^*,r)$

$$|\inf\{g''(u)e^pe^p : u \in L_0\} - \sup\{g''(u)e^pe^p : u \in L_0\}| \leq \tfrac{1}{2}\epsilon c$$

and we have $\inf \mathcal{a}_p / \sup \mathcal{a}_p \leq 1-\tfrac{1}{2}\epsilon$. For $\gamma = \inf \gamma_p$ we have $\gamma \geq 1-\tfrac{1}{2}\epsilon$ and the result follows.

If g'' exists and is continuous then (3.4.5) implies that for every p, γ_p is greater than the γ defined by (3.3.11). Moreover, if g'' is expressed as a matrix (a_{ij}) and the sequence $\{e^p\}$ consists only of the n distinct coordinate vectors, then γ_p is greater than γ defined by (3.3.10), for all p.

3.5 Goldstein and Armijo Algorithms. Given $u^p \in D$ we might consider choosing u^{p+1} by verifying directly that

$$(3.5.1) \qquad g(u^p) - g(u^{p+1}) \geq d(g'(u^p)e^p), \qquad u^{p+1} \in D,$$

for some simple forcing functions such as $d(t) \equiv qt$ or qt^2. However, it is not possible to guarantee, in general, that for any forcing function d there exists a u^{p+1} satisfying (3.5.1). But a related approach does succeed.

Suppose $u^{p+1} = u^p$ when $g'(u^p)e^p = 0$. If $g'(u^p)e^p > 0$ we seek to verify instead

$$(3.5.2) \qquad g(u^p) - g(u^{p+1}) \geq d_1(g'(u^p)e^p)\|u^p - u^{p+1}\|, \qquad u^{p+1} \in D$$

for a forcing function d_1 satisfying $d_1(t)/t \leq q \leq \tfrac{1}{2}$. However, (3.5.2) will imply (3.5.1) only if we also know that

$$\|u^p - u^{p+1}\| \geqq d_2(g'(u^p)e^p) \tag{3.5.3}$$

for a forcing function d_2. In this section we consider algorithms which verify (3.5.2) directly and show how they also satisfy (3.5.3).

We consider first an algorithm proposed by Goldstein [15], [16], [17], and also discussed by Altman [2]. The algorithm is a procedure for taking another iterative method and modifying it to obtain a decreasing sequence which satisfies $g'(u^p)e^p \to 0$.

Assume that the original algorithm produces at $u^p \in D$ an iterate $\overline{u}^p$ which equals u^p if $g'(u^p)e^p = 0$ and otherwise

$$\|u^p - \overline{u}^p\| \geqq d_3(g'(u^p)e^p) \tag{3.5.4}$$

for a forcing function d_3 such that $d_3(0) = 0$. If setting $u^{p+1} = \overline{u}^p$ fails to satisfy

$$g(u^p) - g(u^{p+1}) \geqq q \cdot g'(u^p)e^p \cdot \|u^p - u^{p+1}\|, \tag{3.5.5}$$

$$0 < q \leqq \tfrac{1}{2}, \qquad u^{p+1} \in D$$

then $g'(u^p)e^p > 0$ and a number $0 < w_p < 1$ is found such that

$$u^{p+1} = w_p\overline{u}^p + (1-w_p)u^p \tag{3.5.6}$$

satisfies both (3.5.5) and

$$g(u^p) - g(u^{p+1}) \leqq (1-q)g'(u^p)e^p \cdot \|u^p - u^{p+1}\|. \tag{3.5.7}$$

We now show that if $g'(u^p)e^p > 0$ and (3.5.5) is false for $u^{p+1} = u^p$ there is some $w_p \in (0,1)$ such that u^{p+1} given by (3.5.4) satisfies (3.5.5) and (3.5.7). Define

$$h(u) = \frac{g(u^p) - g(u)}{\|u^p - u\| g'(u^p) e^p} .$$

It follows from the differentiability of g that h is continuous for $u \in \{u^p - t e^p \in D: t \geqq 0\}$ and tends to unity as u tends to u^p. If $\overline{u}^p \notin D$, then by Lemma 3.1.2 there is interval $[u^p, v] \subset D$ where $v \neq u^p$ is a boundary point of L_0 and $g(v) = g(u^\circ) \geqq g(u^p)$, thus $h(v) \leqq 0$. On the other hand if $\overline{u}^p \in D$ but $g(u^p) - g(\overline{u}^p) < qg'(u^p)e^p \|u^p - \overline{u}^p\|$ then $h(\overline{u}^p) < \frac{1}{2}$. In either event $h(u)$ must assume all values between $\frac{1}{2}$ and 1 in the interval $[u^p, \overline{u}^p]$ and therefore for w_p sufficiently small, both (3.5.5) and (3.5.7) can be satisfied, for some $u^{p+1} \in D$.

This algorithm can be generalized by replacing (3.5.5) with (3.5.2) and more significantly replacing (3.5.7) by the weaker condition

$$(3.5.8) \qquad \left| \frac{g(u^p) - g(u^{p+1})}{\|u^p - u^{p+1}\|} - g'(u^p) e^p \right| \geqq d_1(g'(u^p) e^p)$$

for the d_1 of (3.5.2). Note that even if $d_1(t)$ is taken as qt for $q \leqq \frac{1}{2}$, (3.5.8) is weaker than (3.5.7). In the following lemma we show that (3.5.8) implies (3.5.3).

<u>Lemma 3.5.1</u>. Assume D is convex, u^p, $u^{p+1} \in D$. If

(3.5.8) holds with a forcing function d_1 then (3.5.3) also holds with $d_2(t) = \delta(d_1(t))$, where δ is defined by (1.1.7).

<u>Proof</u>: If $u^p, u^{p+1} \in D$ and (3.5.8) holds, then by the convexity of D, $(u^p, u^{p+1}) \subset D$ and by the mean value theorem there is some $v \in (u^p, u^{p+1})$ such that

$$\left[\frac{g(u^p) - g(u^{p+1})}{\|u^p - u^{p+1}\|}\right] = g'(v)e^p.$$

(3.5.8) then implies

$$\|g'(v) - g'(u^p)\| \geqq |[g'(v) - g'(u^p)]e^p| \geqq d_1(g'(u^p)e^p).$$

From (1.1.7) we have

$$\|v - u^p\| \geqq \delta(\|g'(v) - g'(u^p)\|) \geqq \delta(d_1(g'(u^p)e^p))$$

and since $v \in (u^p, u^{p+1})$, we may conclude

$$\|u^p - u^{p+1}\| \geqq \delta(d_1(g'(u^p)e^p)). \tag{3.5.3'}$$

Although there is a number w_p which satisfies both (3.5.2) and (3.5.8), in practice such a number may be obtained only by trial and error. Presumably one would let w_p be successively $1, \frac{1}{2}, \frac{1}{4}, \ldots 2^{-n}, \ldots$ until, for some value of w_p, u^{p+1} satisfies (3.5.2). Then one would test if (3.5.8) also held, and if not, increase w_p until both (3.5.2) and (3.5.8) hold.

We shall show in the next two lemmas that if w_p is chosen as the first value in the sequence $1, \frac{1}{2}, \frac{1}{4}, \ldots$ which satisfies (3.5.2) then it is unnecessary to verify (3.5.8).

We refer to this procedure as the Armijo algorithm because it contains as a special case an algorithm of Armijo [3]. There are two ways in which the preceding trial value of u^{p+1} can have failed to satisfy (3.5.2) Either it did not belong to D or it belonged to D but $g(u^{p+1})$ was too large. Each of these cases is handled in a separate lemma.

Lemma 3.5.2. Let D be convex, $0 < a < 1$, $u^p, u^p - a^{-1}(u^p - u^{p+1}) \in D$, and suppose

(3.5.9) $g(u^p) - g(u^p - a^{-1}(u^p - u^{p+1})) < a d_1(g'(u^p)e^p)\|u^p - u^{p+1}\|$

where $d_1(t) \leqq qt$, $0 < q \leqq \frac{1}{2}$. Then, for δ defined by (1.1.7) we have

(3.5.3") $$\|u^p - u^{p+1}\| \geqq a\delta((1-q)g'(u^p)e^p)$$

Proof: By the convexity of D, the interval $[u^p, \hat{u}^p]$ lies in D, where $\hat{u}^p = u^p - a^{-1}(u^p - u^{p+1})$, and by the mean value theorem and (3.5.9) we have

$$g'(v)e^p = \frac{g(u^p) - g(\hat{u}^p)}{\|u^p - \hat{u}^p\|} < d_1(g'(u^p)e^p) \leqq qg'(u^p)e^p$$

where $v \in (u^p, \hat{u}^p)$ and $e^p = (u^p - u^{p+1})/\|u^p - u^{p+1}\|$. Thus

$$\|g'(u^p) - g'(\hat{u}^p)\| \geqq [g'(u^p) - g'(\hat{u}^p)]e^p \geqq (1-q)g'(u^p)e^p.$$

By the monotonicity of δ we have

$$\|u^p - u^{p+1}\| = a\|u^p - \hat{u}^p\| \geqq a\delta(\|g'(u^p) - g'(\hat{u}^p)\|)$$
$$\geqq a\delta((1-q)g'(u^p)e^p)$$

and (3.5.3") is therefore satisfied.

Lemma 3.5.3. Let D be convex, $0 < a < 1$, $u^p \in D$ and suppose

$$\hat{u}^p = u^p - a^{-1}(u^p - u^{p+1}) \notin D. \tag{3.5.10}$$

Then

$$\|u^p - u^{p+1}\| \geqq a\delta(g'(u^p)e^p).$$

Proof: In the proof of Theorem 3.2.1 we showed ((3.2.8) with $q = 0$) that the set $\{u^p - te^p \in L_0 : t \geqq 0\}$ contains at least the interval $\{u^p - te^p : 0 \leqq t \leqq \delta(g'(u^p)e^p)\}$, and therefore if $\hat{u}^p \notin L_0$, $\|\hat{u}^p - u^p\| \geqq \delta(g'(u^p)e^p)$. Hence $\|u^p - u^{p+1}\| \geqq a\delta(g'(u^p)e^p)$.

In the next theorem we show that the Goldstein and Armijo algorithms produce well-defined sequences for which $g'(u^p)e^p \to 0$ $p \to \infty$. In general, u^p and u^{p+1} need not belong to the same component of the level set $L = \{u \in D : g(u) \leqq g(u^0)\}$, and therefore in the next theorem we strengthen the underlying assumptions from conditions on L_0 to assumptions about L.

Theorem 3.5.1. Let D be convex, L closed, g bounded below on L, $0 < a < 1$ given and suppose for every $u^p \in D$ we have a $\overline{u}^p$ satisfying

$$\|u^p - \overline{u}^p\| \geqq d_3(g'(u^p)e^p) \tag{3.5.4}$$

where $\overline{u}^p = u^p$ if $g'(u^p)e^p = 0$, and $d_3(0) = 0$. Then either the Goldstein or Armijo procedure may be used to obtain a well-defined sequence of iterates such that $g'(u^p)e^p \to 0$.

Proof: Let $d_1(t)$ be a forcing function such that $d_1(t)/t \leqq q \leqq \frac{1}{2}$. If, setting $u^{p+1} = \bar{u}^p$, (3.5.2) is satisfied then (3.5.4) implies (3.5.3). Otherwise $d_3(0) = 0$ implies $g'(u^p)e^p$ must be positive and then we know that there is some $w_p \in (0,1)$ such that for u^{p+1} given by (3.5.6) both (3.5.2) and (3.5.8) are satisfied, and by Lemma 3.5.1, (3.5.3) holds. Alternately, letting w_p be the first value in the sequence $a^0, a^1, \ldots, a^i, \ldots$ for which (3.5.2) holds (it follows from $g'(u^p)e^p > 0$ that (3.5.2) ultimately does hold) we have either (3.5.9) or (3.5.10). Applying Lemma 3.5.2 or 3.5.3, respectively, we may conclude (3.5.3). But (3.5.2) implies $g(u^{p+1}) \leqq g(u^p)$; thus the sequence remains in L and by Lemma 3.1.1 and (3.5.3) $g'(u^p)e^p \to 0$.

The following result, due to Armijo, is a corollary of Theorem 3.5.1.

Corollary 1. Suppose $g: E^n \to R$, and

$$u^{p+1} = u^p - a_p[g'(u^p)]^T$$

where a_p is the first number in the sequence $1, \frac{1}{2}, \frac{1}{4}, \ldots$ to satisfy

$$g(u^p) - g(u^p - a_p[g'(u^p)]^T) \geqq \tfrac{1}{2}\|g'(u^p)\|^2 .$$

Then $g'(u^p) \to 0$.

Proof: For this iteration (3.5.4) holds with $d_3(t) = t$ and (3.5.2) with $d_1(t) = \frac{1}{2}t$. Moreover $g'(u^p)e^p \to 0$ is equivalent to $g'(u^p) \to 0$.

3.6. Searching for the Minimum of a Strictly Unimodal Functional. In this section we consider algorithms which evaluate a strictly unimodal functional at a finite number of points lying in the direction e^p and bracket the minimum in that direction. We will show that these algorithms produce step-sizes that are under-relaxed with respect to one of the variations of Altman's algorithm and it will follow from our results in Section 3.2 that they produce well-defined iterates such that $g'(u^p)e^p \to 0$.

Before discussing direct search algorithms we first give some results about a variation of the step-size algorithm given by (3.2.5). Let the direction e^p be given, set $a_p = 0$ if $g'(u^p)e^p = 0$, and otherwise, let

(3.6.1) $\quad a_p = \sup\{t: 0 \leqq r < t \text{ implies } g'(u^p - te^p)e^p \geqq 0\}.$

Let w_p satisfy

(3.6.2) $$c \leqq w_p \leqq 1, \qquad 0 < c \leqq 1,$$

and define $u^{p+1} = u^p - w_p a_p e^p$. It follows from Corollary 1 of Theorem 3.2.1 that if L_0 is bounded then the iterates are well-defined, remain in L_0 and $g'(u^p)e^p \to 0$.

Note that $\overline{u}^p = u^p - a_p e^p$ is a local minimum of g on $I_p = \{u^p - te^p \in L_0 : t \geqq 0\}$, since $h(t) = g(u^p - te^p)$ is non-increasing for $0 \leqq t \leqq a_p$ and increasing at least for some small interval beyond a_p. If we assume that g is strictly quasi-convex then g has precisely one local minimum on a bounded line segment and $\bar{u}^p$ must be a (global) minimum of g on I_p. Therefore, when g is strictly quasi-convex, this algorithm coincides with the minimization algorithm, but our result now permits the use of an under-relaxation factor. (In general this is not permissible with the minimization algorithm.) Under the assumption of strict quasi-convexity we can also show that the minimization algorithm produces iterates satisfying $\|u^p-u^{p+1}\| \to 0$ as $p \to \infty$.

Theorem 3.6.1. If L_0 is bounded, g is strictly quasi-convex, and $t_p = w_p a_p$ for a_p satisfying (3.6.1) and w_p satisfying (3.6.2) then $t_p \to 0$ as $p \to \infty$.

Proof: Since $g(u^p-te^p)$ is non-increasing for $t \leqq a_p$ it follows that

(3.6.3) $$g(u^p) \gneqq g(wu^p + (1-w)u^{p+1}) \geqq g(u^{p+1})$$

for $w \in (0,1)$. By Theorem 1.4.1 strict quasi-convexity implies property S and by Theorem 2.4.2, (3.6.3) then implies $t_p \to 0$.

We now consider direct search algorithms for strictly unimodal functionals. Let I denote a bounded interval, ordered in the usual fashion, and x^* the minimim of g on I. Recall, (section 1.3) that a functional g is strictly unimodal if for any bounded interval I and points u,v,x^*

of I, if $u < v < x^*$ or if $u > v > x^*$ then $g(u) > g(v) > g(x^*)$. Recall also that in Theorem 1.3.5 we showed that strict unimodality was equivalent to strict quasi-convexity, and we shall therefore use the terms interchangeably.

The property of strictly unimodal functionals exploited by direct search algorithms is this: if $u < w < v$ are three points of a bounded interval I, $g(w) \leqq g(u)$, and $g(w) \leqq g(v)$ then the minimum x^* of g on I is in the segment $[u,v]$. For if x^* is not in $[u,v]$, say $x^* > v$ then $w < v < x^*$ implies $g(w) > g(v) > g(x^*)$ which contradicts the assumption.

One very simple direct search technique for the minimum of g in the interval I_p is to place m equally spaced points $u^p = u_1 < u_2 < \ldots < u_m$ in the interval. If $g'(u^p)e^p > 0$ and m is sufficiently large then some $k > 2$ will satisfy

$$g(u_k) = \min\{g(u_i) : i = 1,\ldots,m\}$$

and x^* will lie in the interval $[u_{k-1}, u_{k+1}]$. Setting $u^{p+1} = u_{k-1}$ implies that $u^{p+1} = u^p - w_p a_p e^p$ for $\frac{1}{3} \leqq w_p \leqq 1$.

Therefore, if

$$u^{p+1} = u^p \quad \text{when} \quad g'(u^p)e^p = 0, \tag{3.6.4a}$$

otherwise,

(3.6.4b) $$g(u_k) = \min\{g(u_i) : i = 1,\ldots,m\}, \quad k > 2,$$

and

(3.6.4c) $$u^{p+1} = u_{k-1},$$

then this direct search algorithm will, by Corollary 1 of Theorem 3.2.1, produce well-defined iterates that remain in L_0 and satisfy $g'(u^p)e^p \to 0$, $p \to \infty$. Moreover, $\|u^p - u^{p+1}\| \to 0$, as $p \to \infty$.

In practice, one would use a more sophisticated strategy for placing the points u_k (see especially Wilde [44]), but it is clear that our analysis extends to more complicated direct search algorithms. The constant c will differ from method to method, but if (3.6.4) is satisfied then u^{p+1} will equal $u^p - w_p a_p e^p$ for some $0 < c \leqq w_p \leqq 1$, and we will again have well-defined iterates satisfying $g'(u^p)e^p \to 0$, and $\|u^p - u^{p+1}\| \to 0$, as $p \to \infty$.

CHAPTER IV

REPRESENTATIVE CONVERGENCE THEOREMS

The last chapter was devoted entirely to the analysis of suitable choices of step-size for minimization algorithms and the one preceding it to the discussion of the choice of directions and to questions of convergence. In this chapter we seek to tie these results together by applying them to a series of specific combinations of step-size and direction. There are dozens of possible combinations but many are uninteresting. We are motivated in our selections by several factors. Well known algorithms are discussed so that our results may be compared to those in the literature. Some new combinations that are computationally more convenient or apply to a wider class of problems than the well known algorithms are also considered. Finally we wish to illustrate as many methods as possible and some selections are made simply for completeness.

Proofs in this chapter will be brief, restricted essentially to quoting the relevant results in the previous chapters and proving their hypotheses are satisfied. We assume in this chapter that $g:D \subset E^n \to R$ has a continuous Frechet derivative on an open set D, that L_0 is a closed, bounded component of the level set $\{u \in D: g(u) \leqq g(u^\circ)\}$ which contains the initial iterate u°. Continuity of g and compactness of L_0 then imply that g' is uniformly continuous on L_0 and there exists at least one point $x^* \in L_0$ such that $g(u) \geqq g(x^*)$, for all $u \in L_0$ and $g'(x^*) = 0$.

4.1 Gauss-Seidel Directions. We begin with the Gauss-Seidel method, which has already been presented as an example in Chapter 2, writing the iteration as $u^{p+1} = u^p - t_p e^p$. As directions e^p, we use the n orthonormal coordinate vectors cyclically. Thus, if $e_0, e_1, \ldots, e_{n-1}$ are the coordinate vectors, e^p is given by

$$e^p = \operatorname{sgn}(g'(u^p)e_i)e_i, \qquad i = p \pmod{n}. \tag{4.1.1}$$

To define the step-size t_p let a_p be the smallest non-negative solution of

$$g'(u^p - te^p)e^p = 0 \tag{4.1.2}$$

and let w_p be a relaxation factor satisfying

$$d_1(g'(u^p)e^p) \leqq w_p \leqq 1, \tag{4.1.3}$$

for a forcing function d_1 (Definition 1.1.1). In particular, $d_1(t) \equiv c$ for $0 < c \leqq 1$ may be used. Setting $t_p = w_p a_p$ we have the Gauss-Seidel algorithm for minimizing a functional.

The best previous result for this method is due to Schechter [38] who used $w_p \equiv 1$ and required that g have a uniformly positive definite second derivative. We need only the assumption that g has one uniformly continuous derivative and property S (Definition 1.4.1). Recall that g has property S if whenever $g(u) = g(v)$ for $u \neq v$ there is some w in the open interval (u,v) such that $g(w) \neq g(v)$. Strict convexity, strict quasi-convexity and strict pseudo-convexity each imply property S, but are considerably stronger.

<u>Theorem 4.1.1</u>. Suppose that $g'(u) = 0$ has only isolated solutions in L_0, g has property S on L_0, the sequence $\{e^p\}$ is defined by (4.1.1), a_p is given by (4.1.2) and w_p satisfies (4.1.3). Then with

$$u^{p+1} = u^p - w_p a_p e^p \tag{4.1.4}$$

we may conclude that the iterates u^p are well-defined, remain in L_0, and converge to a solution of $g'(u) = 0$.

<u>Proof</u>: Clearly (4.1.1) implies that $\|e^p\| = 1$ and $g'(u^p)e^p \geqq 0$. By taking $q = 0$ in Theorem 3.2.1 we have that all the iterates u^p are well-defined, remain in L_0 and $g'(u^p)e^p$ tends to zero. The mean value theorem implies that for this choice of step-size

$$g(u^p) \geqq g(tu^p+(1-t)u^{p+1}) \geqq g(u^{p+1}) \qquad t \in (0,1) \tag{4.1.5}$$

and then since L_0 is compact, and g has property S, Theorem 2.4.2 yields that $\|u^p - u^{p+1}\| \to 0$. With this, and the uniform linear independence of the directions e^p, Theorem 2.4.1 implies that $g'(u^p) \to 0$. But $g'(u) = 0$ has only isolated solutions and thus Theorem 2.2.1 yields the convergence of the iteration to a solution. This completes the proof.

If g is strictly pseudo-convex then by Theorem 1.3.1, 1.3.3, and 1.4.1, g has property S, and $g'(u) = 0$ has a unique solution. Therefore a strictly pseudo-convex functional satisfies the hypotheses of the theorem. We give, however, another example which satisfies the hypotheses but which is not even quasi-convex.

Let $g: E^2 \to R$ be the Rosenbrock [36] functional defined by

$$g(x_1, x_2) = 100(x_1^2 - x_2)^2 + (1-x_1)^2, \tag{4.1.6}$$

whose level curves are shown in Figure 4.1.1.

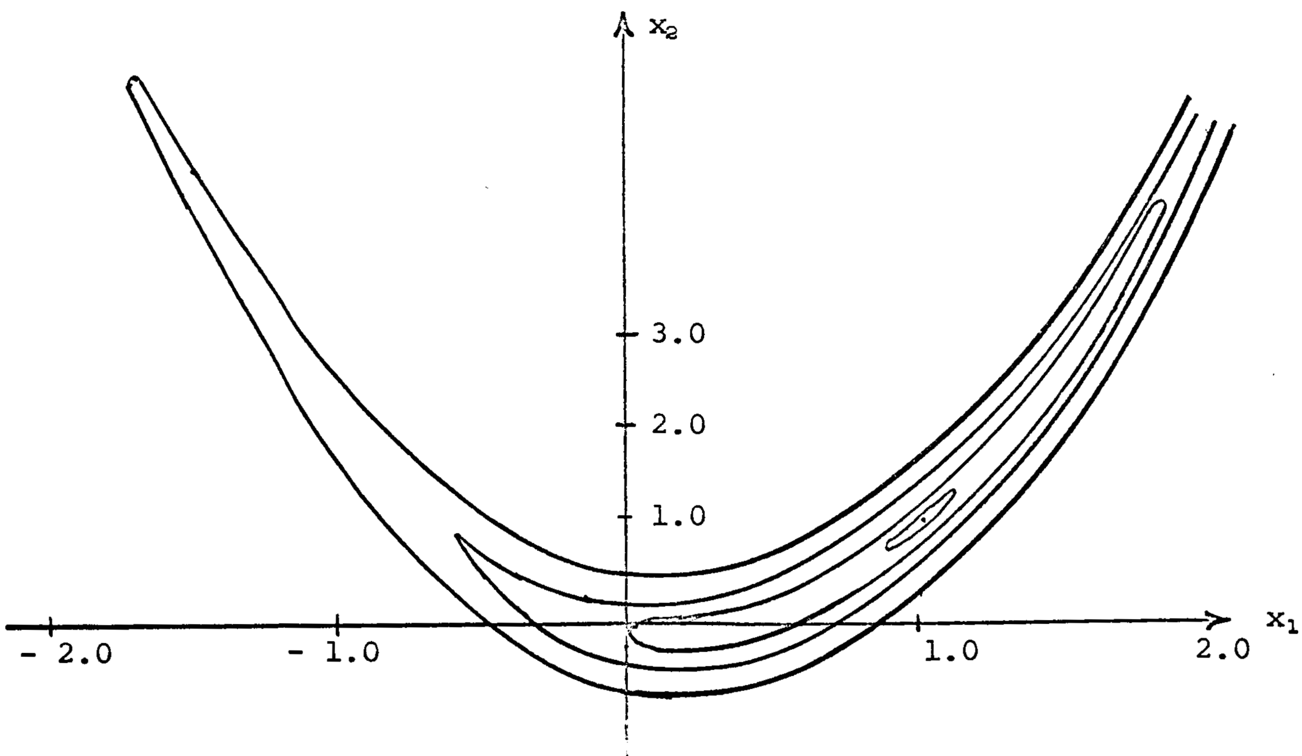

Figure 4.1.1 Level Curves of the Rosenbrock Functional

We will verify that the conditions of Theorem 4.1.1 are satisfied by this functional. It is easily seen that g has bounded level sets since $g(x) \to \infty$ as $\|x\| \to \infty$. Further,

$$g'(x) = \begin{bmatrix} 400x_1(x_1^2-x_2) - 2(1-x_1) \\ -200(x_1^2-x_2) \end{bmatrix}$$

and the only solution of $g'(x) = 0$ is $(1,1)$. To prove g has property S it is sufficient to show g is not constant on any line segment. Suppose for some constant c,

$$h(t) = g(u+te) = c, \qquad 0 < t \leqq t_0, \ \|e\| = 1.$$

Then all derivatives of h on $[0,t_0]$ must be zero but an easy calculation shows that $h^4(t) = 2400e_1^4$. Hence $e_1 = 0$ and then $h'(t) = 0$ implies that $u_1^2 = u_2+te_2$ for $t \in [0,t_0]$. Therefore g is not constant on an interval, has property S and thus satisfies the conditions of Theorem 4.1.1. However, g is not convex or even quasi-convex so that Schechter's result does not apply to it.

In the proof of Theorem 4.1.1 property S was used only to establish that $\|u^p - u^{p+1}\|$ tended to zero. We now suggest a modification of the Gauss-Seidel algorithm which eliminates the need to assume property S by implying directly that $\|u^p - u^{p+1}\|$ tends to zero. Suppose for some fixed $C > 0$ we let t_p be the minimum of $w_p a_p$ and

$Cg'(u^p)e^p$. Then we would still know, using the arguments of Theorem 3.2.2, that u^p is well-defined and all the u^p remain in L_0. Moreover, $g'(u^p)e^p$ still tends to zero, but now this implies directly that $\|u^p - u^{p+1}\| \to 0$. Thus a convergence theorem for this modified Gauss-Seidel algorithm can be given without assuming property S.

Let us digress for the moment from our examination of algorithms and consider whether the other hypotheses of Theorem 4.1.1 are essential. Elementary examples, even in one dimension, show that neither the boundedness nor the closure of L_0 may be dispensed with in general. If g is continuous but not continuously differentiable everywhere, then (even if it is uniformly convex) the Gauss-Seidel algorithm may not converge to a minimum of g. For example, if $g(x) = x_1^2 + x_2^2 + |x_1 - x_2|$ then g is bounded below on E^n, uniformly convex, has closed bounded level sets, and a unique minimum. Nonetheless, there are u^o for which the sequence $\{u^p\}$ will not converge to the minimum of g.

The hypothesis that g' has isolated zeros is probably unnecessary. Kahan [20] has shown that the Gauss-Seidel method will converge for quadratic functionals with a continuum of minima. The extension of this result to non-quadratic functionals is still an open question.

A major objection to using the Curry algorithm, (4.1.2), to determine the step-size is that at each iteration it requires the solution of a nonlinear equation. If a Lipschitz constant C for g' is known, we may use the next iteration which is simpler to apply.

$$u^{p+1} = u^p - w_p (g'(u^p)e^p)/C \quad e^p \tag{4.1.7}$$

where w_p satisfies

$$d_1(g'(u^p)e^p) \leqq w_p \leqq 2 - d_1(g'(u^p)e^p) \tag{4.1.8}$$

for a forcing function d_1, and C is a Lipschitz constant for g', i.e.,

$$\|g'(u) - g'(v)\| \leqq C\|u-v\| \qquad u,v \in L_0. \tag{4.1.9}$$

It is clear that $\|u^p - u^{p+1}\| \to 0$ whenever $g'(u^p)e^p \to 0$, and therefore we need not assume property S to prove convergence for this algorithm.

<u>Theorem 4.1.2</u>. If $g'(u) = 0$ has isolated solutions in L_0 and (4.1.9) holds then the iterates $\{u^p\}$ given by (4.1.1), (4.1.7), (4.1.8) and (4.1.9) are well-defined, remain in L_0 and converge to a solution of $g'(u) = 0$.

<u>Proof</u>: By Theorem 3.3.2 the iterates u^p are well-defined, remain in L_0 and $g'(u^p)e^p \to 0$. Hence, $\|u^p - u^{p+1}\| \to 0$ and Theorems 2.4.1 and 2.2.1 show that $g'(u^p) \to 0$ and the sequence $\{u^p\}$ converges to a solution of $g'(u) = 0$.

<u>4.2 Rosenbrock Directions</u>. Although the Gauss-Seidel

iterates converge, the rate may be slow. This is particularly true if the functional g has a long narrow valley which is not parallel to any coordinate axis. The next direction algorithm (based on a proposal by Rosenbrock [36]) attempts to speed the convergence by using more information about the past behavior of the iteration sequence.

For the Rosenbrock algorithm the directions are computed in blocks of n, and depend on the results of the last n iterations. The first n iterations of this algorithm are precisely the same as for the Gauss-Seidel directions. Once u^n has been computed, the Gram-Schmidt orthonormalization process is applied to the sequence $u^n - u^0, u^{n-1} - u^0, \dots, u^1 - u^0$, or if these are linearly dependent, to the set $B_k = \{u^n - u^0, \dots, u^1 - u^0, e^0, e^1, \dots, e^k\}$ for the smallest value of k such that B_k spans E^n. This will produce n new mutually orthogonal vectors, $e^n, \dots, e^{2n-1}$, of norm unity, and when using these direction vectors, $u^{n+1}, u^{n+2}, \dots, u^{2n}$ have been computed, the Gram-Schmidt process is then applied to the sequence $u^{2n-1} - u^n, \dots, u^{n+1} - u^n$. Again if these vectors are linearly dependent we continue the Gram-Schmidt process with one or more of the vectors $e^n, e^{n+1}, \dots, e^{2n}$ to produce $e^{2n}, \dots, e^{3n-1}$. The Gram-Schmidt process can always be carried out because the vectors $e^n, \dots,$

e^{2n-1} are linearly independent. Continuing in this manner we produce a succession of blocks of n mutually orthogonal vectors. Such a sequence of directions satisfies our definition of uniform linear independence.

In the next algorithm we combine the Rosenbrock directions with a modification of the usual minimization algorithm similar to the modification we suggested for the Curry algorithm in the last section. In this modified algorithm we fix $C > 0$ and choose $u^{p+1} = u^p - t_p e^p \in L_0$ such that

(4.2.1) $$g(u^p - t_p e^p)$$

$$= \min\{g(u^p - te^p) : u^p - te^p \in L_0;\ 0 \leqq t \leqq Cg'(u^p)e^p\}.$$

Theorem 4.2.1. Let the sequence of directions $\{e^p\}$ be chosen as described above, with the sign taken so that $g'(u^p)e^p \geqq 0$. Define the sequence of iterates $u^{p+1} = u^p - t_p e^p$ by (4.2.1). If $g'(u) = 0$ has isolated solutions then the iterates u^p are well-defined and converge to a solution of $g'(u) = 0$.

Proof: Since L_0 is closed the intersection of $[u^p, u^p - Cg'(u^p)e^p)e^p]$ with L_0 is compact and therefore u^{p+1} is always defined and in L_0. As noted in our remarks following Theorem 3.2.3, $g'(u^p)e^p \to 0$ and $\|u^p - u^{p+1}\| \to 0$. Now recall that a sequence of vectors is uniformly linearly independent (Definition 2.4.1) if there is some $m \geqq n$ and $c > 0$

such that for any p' and any $x \in E^n$

$$\max_{p'+1 \leqq p \leqq p'+m} \{|x^T e^p|\} \geqq c\|x\|. \tag{4.2.2}$$

For the Rosenbrock directions, in every $2n-1$ successive elements of $\{e^p\}$ there are n that are mutually orthogonal and therefore we can satisfy (4.2.2) with $m = 2n-1$ and $c = n^{-\frac{1}{2}}$. By Theorem 2.4.1, $g'(u^p) \to 0$, and then by Theorem 2.2.1, the sequence $\{u^p\}$ converges to a zero of g'.

If we had used the usual minimization algorithm, $u^{p+1} = u^p - t_p e^p$ such that

$$g(u^{p+1}) = \min\{g(u^p - te^p) : t \geqq 0;\ u^p - te^p \in L_0\}, \tag{4.2.3}$$

we would not have been able to conclude that $\|u^p - u^{p+1}\| \to 0$, directly from $g'(u^p)e^p \to 0$. Recall that in Theorem 4.1.1 we assumed property S, to prove that $\|u^p - u^{p+1}\| \to 0$ verified that the algorithm always satisfies

$$g(u^p) \geqq g(tu^p + (1-t)u^{p+1}) \geqq g(u^{p+1}), \quad 0 \leqq t \leqq c \tag{4.2.4}$$

for some fixed $c > 0$, and applied Theorem 2.4.2. But for (4.2.3) only the second inequality in (4.2.4) follows directly. If, however, g is quasi-convex (i.e., for any $u,v \in D$, $g(u) \leqq g(v)$ implies $g(w) \leqq g(v)$ for all w in the interval (u,v)) then (4.2.4) follows. Now, in Theorem 1.4.1 we showed that quasi-convexity plus property S is equivalent to strict quasi-convexity. Therefore we can prove convergence for (4.2.3) with uniformly linearly independent directions, by assuming that g is strictly quasi-convex.

Since Theorems 4.2.1, 4.1.1 and 4.1.2 all rest on the uniform linear independence of $\{e^p\}$, it is clear that 4.2.1 is also valid if the sequence $\{e^p\}$ consists of the Gauss-Seidel directions, and by the same token 4.1.1 and 4.1.2 apply to the Rosenbrock direction algorithm.

4.3 The Seidel and Gauss-Southwell Directions. Some direction algorithms use only the coordinate directions e_1, e_2, ... , e_n, but instead of choosing them cyclically, as in the Gauss-Seidel algorithm, the coordinate is selected according to some particular criterion. For example, the Gauss-Southwell algorithm chooses a coordinate vector $\bar{e}$ satisfying

$$(4.3.1) \qquad |g'(u^p)\bar{e}| = \max_{1 \leq i \leq n} |g'(u^p)e_i|$$

and sets $e^p = \operatorname{sgn}(g'(u^p)\bar{e})\bar{e}$. This is not a free-steering method (Definition 2.5.1) since in general every coordinate direction need not appear infinitely often in the sequence $\{e^p : p = 0,1, \ldots\}$. However, since e^p is the direction corresponding to the largest component of the gradient, it is a gradient-related method (Definition 2.3.1). In fact, (4.3.1) easily implies

$$(4.3.2) \qquad g'(u^p)e^p \geq n^{-\frac{1}{2}}\|g'(u^p)\|.$$

Goldstein [14] has analysed this choice of direction in conjunction with several step-size algorithms.

In this section we will study a similar method, the Seidel algorithm, which selects a coordinate vector using a different criterion.

For the quadratic functional $g(u) = u^TAu + b^Tu + c$, where A is the matrix (a_{ij}), set

$$u^TA + b^T = (r_1, r_2, \ldots, r_n) = r^T.$$

Then the Seidel algorithm uses the coordinate direction e_i for which $(r_i)^2/a_{ii}$ is a maximum. For the non-quadratic minimization problem this corresponds to choosing $\bar{e}$ as a coordinate vector which satisfies

$$(4.4.3) \qquad \left|\frac{(g'(u^p)\bar{e})^2}{g''(u^p)\bar{e}\bar{e}}\right| = \max_{1 \leqq i \leqq n} \left|\frac{(g'(u^p)e_i)^2}{g''(u^p)e_ie_i}\right|$$

and setting

$$(4.3.4) \qquad e^p = \operatorname{sgn}(g'(u^p)\bar{e})\bar{e}.$$

In Theorem 4.3.1 we shall consider using these directions in conjunction with taking one relaxed Newton step towards the solution of $g'(u^p - te^p)e^p = 0$. That is, the iteration is

$$(4.3.5) \qquad u^{p+1} = u^p - \frac{w_p(g'(u^p)e^p)}{g''(u^p)e^pe^p}\, e^p,$$

for

$$(4.3.6) \qquad d_1(g'(u^p)e^p) \leqq w_p \leqq 2\gamma - d_1(g'(u^p)e^p)$$

where d_1 is a forcing function and γ is given by

$$(4.3.7)\quad \gamma = \min_{1\leq i\leq n}\left\{\inf_{u,t}\left\{\frac{g''(u)e_ie_i}{g''(u-te_i)e_ie_i} : t \geqq 0;\ [u,u-te_i]\in L_0\right\}\right\}.$$

Theorem 4.3.1. Suppose that g has a bounded, uniformly positive definite second derivative in L_0, i.e.,

$$(4.3.8)\quad c\|h\|^2 \leqq g''(u)hh \leqq C\|h\|^2, \qquad u \in L_0,\ 0 < c \leqq C,$$

and the iterates $\{u^p : p = 0,1, \ldots\}$ are given by (4.3.3)-(4.3.7). Then the iterates are well-defined, remain in L_0 and the sequence converges to the unique minimum of g in L_0.

Proof: From (4.3.4) the sequence of directions $\{e^p\}$ contains only n distinct directions (up to a choice of sign) and therefore

$$(3.3.5)\quad \gamma = \inf_p\left\{\inf_{t,u}\left\{\frac{g''(u)e^pe^p}{g''(u-te^p)e^pe^p} : t \geqq 0;\ u,u-te^p \in L_0\right\}\right\}$$

reduces to (4.3.7). Therefore, Theorem 3.3.1 shows that the sequence $\{u^p\}$ is well-defined, remains in L_0 and $g'(u^p)e^p \to 0$. By (4.3.8) and (4.3.3)

$$\frac{(g'(u^p)e^p)^2}{c} \geqq \frac{(g'(u^p)e^p)^2}{g''(u^p)\,e^pe^p}$$

$$\geqq \frac{(g'(u^p)e_i)^2}{g''(u^p)\,e_i\,e_i} \qquad (i = 1, \ldots ,n)$$

$$\geqq \frac{(g'(u^p)e_i)^2}{C} \qquad (i = 1, \ldots ,n)$$

and thus

$$(4.3.9) \qquad g'(u^p)e^p \geqq (c/nC)^{-\frac{1}{2}}\|g'(u^p)\|.$$

Hence the sequence of directions $\{e^p\}$ is gradient-related and it follows immediately from $g'(u^p)e^p \to 0$ and (4.3.9) that $g'(u^p) \to 0$. Moreover (4.3.8) implies that $g'(u) = 0$ has only one solution x^* in L_0 and $g(x^*)$ is the minimum of g in L_0. Therefore the sequence u^p has precisely one limit point and converges.

As a corollary we have the following result whose proof follows easily from the appropriate corollaries of Theorem 3.3.1.

Corollary 1. Suppose g satisfies the conditions of Theorem 4.3.1. Then:

(a) (4.3.6) may be replaced by

$$d_1(g'(u^p)e^p) \leqq w_p \leqq 2c/C - d_1(g'(u^p)e^p)$$

and the conclusions of the theorem remain valid;

(b) if g'' is constant (4.3.6) may be replaced by

$$d_1(g'(u^p)e^p) \leqq w_p \leqq 2 - d_1(g'(u^p)e^p)$$

and the conclusions of the theorem remain valid;

(c) if g'' satisfies (4.3.8) in all of D and g'' is uniformly continuous in a neighborhood of x^*, the minimum of g in D, then for any $\varepsilon > 0$ it is possible to choose u^o sufficiently close to x^* that the results of the theorem follow for w_p satisfying

$$\varepsilon \leqq w_p \leqq 2-\varepsilon.$$

4.4 Modified Jacobi and Newton Iterations. In this section we consider two algorithms, the Jacobi method and Newton's method and show how they may be modified to obtain iterates satisfying $g(u^p) \geqq g(u^{p+1})$ and $g'(u^p) \to 0$. We describe the Jacobi method first.

Let $e_1, \ldots, e_n$ be the orthonormal coordinate vectors of E^n and let t_i be the smallest non-negative solution of

$$g'(u^p - t\ \mathrm{sgn}(g'(u^p)e_i)e_i)e_i = 0, \qquad 1 \leqq i \leqq n. \tag{4.4.1}$$

We then set

$$\overline{u}^p = u^p - \sum_i [t_i\, \mathrm{sgn}(g'(u^p)e_i)e_i]. \tag{4.4.2}$$

In general $g(\overline{u}^p)$ need not be smaller than $g(u^p)$ and therefore we modify this algorithm using the Goldstein algorithm as generalized in section 3.5.

Set $u^{p+1} = \overline{u}^p$ if $u^{p+1} \in D$ and

$$g(u^p) - g(\overline{u}^p) \geqq \tfrac{1}{4} g'(u^p)(u^p - \overline{u}^p), \tag{4.5.3}$$

otherwise choose some $0 < w_p < 1$ such that $u^{p+1} = w_p \overline{u}^p + (1-w_p)u^p$ satisfies

$$g(u^p) - g(u^{p+1}) \geqq \tfrac{1}{4} g'(u^p)(u^p - u^{p+1}), \quad u^{p+1} \in D, \tag{4.4.4}$$

and

$$|g(u^p) - g(u^{p+1}) - g'(u^p)(u^p - u^{p+1})| \geqq \tfrac{1}{4} g'(u^p)(u^p - u^{p+1}). \tag{4.4.5}$$

The next theorem gives sufficient conditions for this iteration to converge.

<u>Theorem 4.4.1</u>. Suppose D is convex and $L = \{u \in D: g(u) \leqq g(u^\circ)\}$ is closed and bounded. Let g be bounded below on L, and let the iteration $u^{p+1} = u^p - t_p e^p$ be defined by (4.4.1) through (4.4.5). Then the iterates are well-defined, remain in L and $g'(u^p) \to 0$.

<u>Proof</u>: Since L is closed and bounded, the equations (4.4.1) have solutions such that $u^p - t_i \operatorname{sgn}(g'(u^p)e_i)e_i \in L$ and therefore $\bar{u}^p$ is well-defined (though it may not belong to D). Moreover, $u^p = \bar{u}^p$ if and only if $g'(u^p) = 0$. Suppose, therefore, that $g'(u^p) \neq 0$ and let k satisfy

$$(4.4.6) \qquad |g'(u^p)e_k| = \max_{1\leqq i\leqq n} |g'(u^p)e_i|;$$

then $|g'(u^p)e_k| \geqq n^{-\frac{1}{2}} \|g'(u^p)\|$. We shall show first that (3.5.4) is satisfied. Let $\delta(t)$ be defined by (1.1.7) and set $\hat{u}^p = u^p - t_k \operatorname{sgn}(g'(u^p)e_k)e_k$. It follows that $\hat{u}^p \in L$ and

$$\begin{aligned} \|u^p - \bar{u}^p\| \geqq \|u^p - \hat{u}^p\| = t_k &\geqq \delta(\|g'(u^p) - g'(\hat{u}^p)\|) \\ &\geqq \delta(|[g'(u^p) - g'(\hat{u}^p)]e_k|) \end{aligned}$$

and since $g'(\hat{u}^p)e_k = 0$, we then have

$$(4.4.7) \qquad \begin{aligned} \|u^p - \hat{u}^p\| \geqq t_k &\geqq \delta(|g'(u^p)e_k|) \\ &\geqq \delta(n^{-\frac{1}{2}}\|g'(u^p)\|) \end{aligned}$$

and (3.5.4) holds. We can also conclude that the direction $e^p = \sum_{i=1}^{n} t_i \operatorname{sgn}(g'(u^p)e_i)e_i / \|\sum_{i=1}^{n} t_i e_i\|$ is gradient-related. We have

$$g'(u^p)e^p = \sum_{i=1}^{n} t_i |g'(u^p)e_i| / \|\sum_{i=1}^{n} t_i e_i\|$$

$$\geq \sum_{i=1}^{n} t_i |g'(u^p)e_i| / \sum_{i=1}^{n} t_i .$$

Since L is bounded, it has a diameter $C > 0$ and hence $t_i \leqq C$. By (4.4.7) we then have

$$g'(u^p)e^p \geqq t_k|g'(u^p)e_k| / nC \tag{4.4.9}$$

$$\geqq \delta(n^{-\frac{1}{2}}\|g'(u^p)\|)\|g'(u^p)\|/Cn^{3/2}$$

$$= d(\|g'(u^p)\|)$$

where $d(t) = \delta(n^{-\frac{1}{2}}t)t/Cn^{3/2}$ is the product of forcing functions and hence forcing. Therefore the directions e^p are gradient-related, and $\|g'(u^p)\| \neq 0$ implies $g'(u^p)e^p > 0$. Since (3.5.4) is satisfied we may apply Theorem 3.5.1 and conclude that u^{p+1} is well-defined, lies in L and $g'(u^p)e^p \to 0$ as $p \to \infty$. Then (4.4.9) implies that $g'(u^p) \to 0$.

Newton's method, $u^{p+1} = u^p - [g''(u^p)]^{-1}g'(u^p)^T$, like the Jacobi iteration has gradient-related directions, but may not decrease the value of g at each step. We will therefore modify the Newton iterates using the Armijo procedure as described in section 3.5.

Let

$$u^{p+1} = u^p - w_p[g''(u^p)]^{-1}g'(u^p)^T, \tag{4.4.10}$$

where w_p is the first number in the sequence $1, \frac{1}{2}, \frac{1}{4}, \ldots$ to satisfy

$$g(u^p) - g(u^p - w_p[g''(u^p)]^{-1}g'(u^p)^T) \tag{4.4.11}$$

$$\geqq \tfrac{1}{2}w_p g'(u^p)[g''(u^p)]^{-1}g'(u^p)^T$$

for $u^p - w_p[g''(u^p)]^{-1} g'(u^p)^T \in D$.

<u>Theorem 4.4.2</u>. Suppose g'' is continuous and strictly positive definite in D, and D is convex. Then the iteration defined by (4.4.10) and (4.4.11) converges to the minimum of g in D.

<u>Proof</u>: g has convex, and hence connected, level sets and thus L_0 is the only component of $\{u \in D: g(u) \leqq g(u^o)\}$. Since g'' is continuous and strictly positive definite and L_0 is compact we have

$$c\|h\|^2 \leqq g''(u)hh \leqq C\|h\|^2, \quad u \in L_0,\ 0 < c \leqq C. \tag{4.4.12}$$

This implies $[g''(u)]^{-1}$ exists, satisfies

$$\frac{1}{C}\|h\|^2 \leqq h^T(g''(u))^{-1}h \leqq \frac{1}{c}\|h\|^2, \quad u \in L_0,$$

and thus $\bar{u}^p = u^p - [g''(u^p)]^{-1}g'(u^p)^T$ is well-defined for $u^p \in L_0$. If $g'(u^p) = 0$, $\bar{u}^p = u^p$; otherwise

$$\begin{aligned} \|u^p - \bar{u}^p\| &\geqq \|[g''(u^p)]^{-1}g'(u^p)^T\| \\ &\geqq \frac{1}{C}\|g'(u^p)\| \end{aligned} \tag{4.4.13}$$

and the iteration satisfies (3.5.4). Further,

$$\begin{aligned} g'(u^p)e^p &\geqq \frac{g'(u^p)[g''(u^p)]^{-1}g'(u^p)^T}{\|[g''(u^p)^{-1}]g'(u^p)^T\|} \\ &\geqq (1/C)\|g'(u^p)\|^2 / \|g''(u^p)^{-1}\|\cdot\|g'(u^p)\| \\ &\geqq (1/C)\|g'(u^p)\| \;/\; (1/c) \\ &\geqq c/C\ \|g'(u^p)\|, \end{aligned} \tag{4.4.14}$$

and the directions $e^p = [g''(u^p)]^{-1}g'(u^p)^T / \|g''(u^p)^{-1}g'(u^p)^T\|$ are gradient-related. By Theorem 3.5.1, u^{p+1} is well-defined, lies in L_0 and $g'(u^p)e^p \to 0$. This and (4.4.14) imply $g'(u^p) \to 0$ and since the only solution of $g'(u) = 0$ is x^*, the unique minimum of g in D, the sequence u^p converges to x^*.

We note that Goldstein [17] has shown that for g'' satisfying the hypothesis of the theorem there exists a number N such that for $p \geqq N$, $w_p = 1$ satisfies (4.4.11). Therefore, after a finite number of iterations, this algorithm coincides with the Newton algorithm.

CHAPTER V

BLOCK METHODS

5.1 Introduction. The block Gauss-Seidel algorithm we consider in this chapter is a generalization of the usual or "point" Gauss-Seidel method we considered in the preceeding chapters, and we will motivate it by considering the problem of solving a system of simultaneous equations written as

$$
\begin{gathered}
F_1(x_1,\ldots,x_n) = 0 \\
\vdots \\
F_n(x_1,\ldots,x_n) = 0.
\end{gathered}
$$

In the Gauss-Seidel algorithm the equations are treated cyclically, one at a time, the first equation solved for the first unknown, etc. In the block Gauss-Seidel method the unknowns and equations are divided into groups or blocks. The first block of equations is then solved simultaneously for the first block of unknowns. Then, substituting these values, the second block of equations is solved for the second block of unknowns, etc. In general, the blocks need not have the same number of elements. However, when all the blocks consist of precisely one element, this reduces to the usual Gauss-Seidel method.

Consider, for example, the system of simultaneous equations that arises from the discretization of a boundary value problem for a partial differential equation. In this case a grid is established in the region of interest and the unknowns represent values of the solution of the partial differential equation at grid points. One natural way of choosing blocks for such a system is to let a block consist of all unknowns associated with a certain grid line. For example, in a planar region we may choose our blocks as the horizontal or the vertical grid lines. In another variation we may alternate between them on successive "sweeps" of the system.

We now consider the corresponding minimization algorithms. Let $e_0, e_1, \ldots, e_{n-1}$ be the orthonormal coordinate vectors of E^n. In the Gauss-Seidel algorithm we select the next iterate u^{p+1} such that

$$u^{p+1} \in \{u^p - te_i : i = p(\text{mod } n),\ t \in R\}. \tag{5.1.1}$$

Suppose, however, the integers $0, 1, \ldots, n-1$ are grouped into k blocks so that if $\varkappa = \{i : 0 \leq i < n\}$ then there are k sets I_j such that $\cup I_j = \varkappa$ and $I_i \cap I_j = \emptyset$, $i \neq j$. In the block Gauss-Seidel method we require that

$$u^{p+1} \in \{u^p + v : v \in H_j;\ j = p(\text{mod } k)\}, \tag{5.1.2}$$

where

$$H_j = \text{span}\{e_i : i \in I_j\}.$$

Note that (5.1.2) reduces to (5.1.1) when all the blocks have precisely one element. Thus we seek the next iterate in the affine subspace $u^p \oplus H_j$, $j = p(\text{mod } k)$, where

(5.1.3) $$H_1 \oplus \cdots \oplus H_k = E^n.$$

When u^{p+1} is chosen as the solution of

(5.1.4) $g'(u)h = 0,\quad h \in H_j,\quad u \in u^p \oplus H_j,\quad j = p(\text{mod } k),$

for $H_j = \text{span}\{e_i : i \in I_j\}$, this is then the cyclic block Gauss-Seidel method described above.

The algorithms we considered in Chapter IV were composed of two parts - choosing a direction and picking a step-size. In a block algorithm we have three steps because the direction is computed in two stages. First we select a subspace H_p, then a direction $e^p \in H_p$ and finally the step-size t_p. Our analysis is also in three parts. In Chapter III we have seen that suitable choice of t_p implies that $g'(u^p)e^p \to 0$. In Section 5.3 we will show we can choose $e^p \in H_p$ such that $g'(u^p)e^p \to 0$ implies $P_p[g'(u^p)^T] \to 0$ where P_p is the orthogonal projection of E^n onto H_p. In Section 5.2 we shall give sufficient conditions on the choice of blocks H_p so that $P_p[g'(u^p)^T] \to 0$ implies $g'(u^p) \to 0$, and in the final section we combine these results to obtain complete convergence theorems for block methods.

The underlying assumptions in this chapter are the same as in Chapter IV. We assume that $g:D \subset E^n \to R$ has a continuous Frechet derivative on an open set D, and that L_o is a closed bounded component of the level set $\{u:g(u) \leqq g(u^\circ)\}$ containing the initial iterate u°.

5.2 Choice of Subspaces. In Section 2.4 we generalized the use of coordinate vectors by the concept of uniform linear independence. We now introduce the analogous concept for subspaces.

Definition 5.2.1. Let $\{H_p:p = 0,1,\ldots\}$ be a sequence of subspaces of E^n and P_p be the orthogonal projection of E^n onto H_p. The sequence $\{H_p\}$ is uniformly linearly independent if there is a $c > 0$ and an integer m such that for any $x \in E^n$ and integer p

$$\max_{1 \leqq i \leqq m} \{\|P_{p+i}(x)\|\} \geqq c\|x\|. \tag{5.2.1}$$

If the sequence $\{H_p\}$ is generated by the cyclic use of a fixed number of subspaces satisfying (5.1.3) it is easily seen to be uniformly linearly independent. The next theorem, whose proof is analogous to that of Theorem 2.4.1 gives conditions under which $P_p[g'(u^p)^T] \to 0$ implies $g'(u^p) \to 0$.

Theorem 5.2.1. Suppose $\{u^p:p = 0,1,\ldots\} \subset D_0 \subset D$, where D_0 is compact, and $\{H_p:p = 0,1,\ldots\}$ is a sequence of uniformly linearly independent subspaces. If

$\|u^p - u^{p+1}\| \to 0$ and $P_p[g'(u^p)^T] \to 0$ as $p \to \infty$, then $g'(u^p) \to 0$, $p \to \infty$.

Proof: Let $\epsilon > 0$ be given. Since g' is uniformly continuous on D_0 the function δ defined by (1.1.7), i.e.,

$$\delta(t) = \inf\{\|u-v\| : u,v \in D ; \|g'(u)-g'(v)\| \geqq t\},$$

satisfies $\delta(t) \geqq 0$. Therefore, because $\|u^p-u^{p+1}\| \to 0$ and $P_p[g'(u^p)^T] \to 0$, we can find a K sufficiently large that

(5.2.2) $$\|u^p-u^{p+1}\| \leqq \delta(\tfrac{1}{2}\epsilon c)/m, \qquad p \geqq K,$$

and

(5.2.3) $$\|P_p[g'(u^p)^T]\| \leqq \tfrac{1}{2}\epsilon c, \qquad p \geqq K,$$

where m and c are the constants of (5.2.1) in the definition of uniform linear independence. From (5.2.2) and the triangle inequality it follows that

(5.2.4) $$\|u^p-u^{p+i}\| \leqq \delta(\tfrac{1}{2}\epsilon c) \qquad 1 \leqq i \leqq m,$$

and then the definition of δ implies that

$$\|g'(u^p)-g'(u^{p+i})\| \leqq \tfrac{1}{2}\epsilon c, \qquad 1 \leqq i \leqq m.$$

Since P_p is a projection, and thus $\|P_p\| \leqq 1$, we have

$$\begin{aligned} \tfrac{1}{2}\epsilon c &\geqq \|P_p[g'(u^p)^T - g'(u^{p+i})^T]\| \\ &\geqq \|P_p[g'(u^p)^T]\| - \|P_p[g'(u^p)^T]\|, \end{aligned}$$

and thus for $1 \leqq i \leqq m$,

$$\tfrac{1}{2}\epsilon c + \|P_p[g'(u^{p+i})^T]\| \geqq \|P_p[g'(u^p)^T]\|.$$

It then follows from (5.2.3) that

$$\epsilon c \geqq \|P_{p+i}[g'(u^p)^T]\|, \quad 1 \leqq i \leqq m, \quad p \geqq K,$$

which with the uniform linear independence of $\{H_p\}$, implies that

$$\epsilon c \geqq \max_{1 \leqq i \leqq m} \{\|P_{p+i}[g'(u^p)^T]\|\} \geqq c\|g'(u^p)\|,$$

and therefore for $p \geqq K$ we have $\|g'(u^p)\| \leqq \epsilon$. But ϵ was arbitrary and hence $g'(u^p) \to 0$.

5.3 Directions Within a Subspace. In this section we study methods of picking $e^p \in H_p$ such that $g'(u^p)e^p \to 0$ implies $P_p(g'(u^p)^T) \to 0$. One method is immediately clear. If $e^p \in H_p$ is arbitrary when $P_p(g'(u^p)^T) = 0$, and otherwise satisfies

$$e^p = P_p(g'(u^p)^T)/\|P_p(g'(u^p)^T)\| \tag{5.3.1}$$

then $\|P_p(g'(u^p)^T)\| = g'(u^p)e^p$ and we have $g'(u^p)e^p \to 0$ implies $P_p(g'(u^p)^T) \to 0$. Less immediately we have a result that applies to the Block Gauss-Seidel algorithm. Here e^p is chosen such that $\{u^p - te^p : t \geqq 0\}$ contains the minimum of g in the intersection of L_o with $u^p \oplus H_p$.

Theorem 5.3.1. Suppose $\{u^p : p = 0,1,\ldots\} \subset L_0$, g is uniformly pseudo-convex on L_0 and $e^p \in H_p$ satisfies $\|e^p\| = 1$ and

$$g'(u^p - t_p e^p)h = 0, \quad \text{for all } h \in H_p, \tag{5.3.2}$$

and some $t_p \geqq 0$. Then $g'(u^p)e^p \to 0$ implies $P_p(g'(u^p)^T) \to 0$.

Proof: Since L_0 is compact g has a minimum on

$[u^p \oplus H_p] \cap L_0$ at some x^* and by Theorem 1.3.1, uniform pseudo-convexity implies that $x^* = u^p - t_p e^p$. Thus $g(u^p - t_p e^p) \leqq g(u^p)$, and by the uniform pseudo-convexity of g,

$$g'(u^p)(u^p - (u^p - t_p e^p)) \geqq d(\|t_p e^p\|)\|t_p e^p\|,$$

where d is the forcing function in the definition of uniform pseudo-convexity, so that

$$[g'(u^p) - g'(u^p - t_p e^p)]e^p \geqq d(t_p).$$

But for δ defined by (1.1.7)

$$t_p = \|u^p - (u^p - t_p e^p)\| \geqq \delta(\|g'(u^p) - g'(u^p - t_p e^p)\|)$$

$$\geqq \delta(\|P_p[g'(u^p)^T - g'(u^p - t_p e^p)^T]\|)$$

and since $P_p(g'(u^p - t_p e^p)^T) = 0$,

$$g'(u^p)e^p \geqq d(\delta(\|P_p[g'(u^p)^T]\|)),$$

and the result follows.

The next theorem applies to the direction obtained by taking one Newton step from u^p towards solving

$$(5.3.3) \qquad g'(u)h = 0, \qquad u \in u^p \oplus H_p, \quad h \in H_p.$$

To compute this direction suppose H_p is an m dimensional subspace of E^n and $v_1, \ldots, v_m$ are m mutually orthogonal vectors that span H_p. If A_p is the $n \times m$ matrix $(v_1 \ \ldots \ v_m)$ then all elements of H_p are linear combinations of $v_1, \ldots, v_m$ and thus

$$u^p \oplus H_p = \{u^p + A_p v : v \in E^m\}.$$

In particular, if $h(v) = g(u^p + A_p v)$, $v \in E^m$, then solving $h'(v) = 0$ is equivalent to solving (5.3.3), and we consider the Newton iteration for $h'(v) = 0$. Since

$$h'(v) = g'(u^p + A_p v)A_p$$

and

$$h''(v) = A_p^T g''(u^p + A_p v)A_p$$

taking one Newton step with v initially zero yields

$$v_p = 0 - [A_p^T g''(u^p)A_p]^{-1}[g'(u^p)A_p]^T$$

and thus one Newton step from u^p is

$$\begin{aligned} u^{p+1} &= u^p - A_p v_p \\ &= u^p - A_p[A_p^T g''(u^p)A_p]^{-1}A_p^T g'(u^p)^T \\ &= u^p - B_p g'(u^p)^T, \end{aligned}$$

where

$$B_p = A_p[A_p^T g''(u^p)A_p]^{-1}A_p^T. \tag{5.3.4}$$

When $\|B_p g'(u^p)^T\| = 0$ let $e^p \in H_p$ be arbitrary. Otherwise set

$$e^p = B_p g'(u^p)^T / \|B_p g'(u^p)^T\|. \tag{5.3.5}$$

<u>Theorem 5.3.2.</u> Suppose g has a continuous, strictly positive definite second derivative on L_0, the sequence $\{u^p\} \subset L_0$ is given, and e^p is defined by (5.3.5). Then $g'(u^p)e^p \to 0$ implies $P_p(g'(u^p)^T) \to 0$.

<u>Proof:</u> Since L_0 is compact $g''(u^p)$ satisfies

$$c\|h\|^2 \leq h^T g''(u^p)h \leq C\|h\|^2 \qquad h \in E^n, \quad 0<c\leq C$$

and since $\|A_p h\| = \|h\|$, $h \in E^m$, we have

$$c\|h\|^2 \leqq h^T A_p^T g''(u^p) A_p h \leqq C\|h\|^2, \qquad h \in E^m.$$

Thus

$$(1/C)\|h\|^2 \leqq h^T (A_p^T g''(u^p) A_p)^{-1} h \leqq (1/c)\|h\|^2,$$

for $h \in E^m$. Further,

$$g'(u^p)e^p = \frac{(g'(u^p)A_p)[A_p^T g''(u^p)A_p]^{-1}(g'(u^p)A_p)^T}{\|A_p[A_p^T g''(u^p)A_p]^{-1}(g'(u^p)A_p)^T\|}$$

$$\geqq \frac{(1/C)\|g'(u^p)A_p\|^2}{\|A_p\| \; \|(A_p^T g''(u^p)A_p)^{-1}\| \; \|(g'(u^p)A_p)^T\|}$$

$$\geqq \quad (c/C)\|g'(u^p)A_p\|.$$

But $\|P_p(g'(u^p)^T)\| = \|g'(u^p)A_p\|$ and the result follows.

In the proof of the theorem the strict positive definiteness of $g''(u^p)$ was used only to conclude that $A_p^T g''(u^p)A_p$ was strictly positive definite. For the cyclic block Gauss-Seidel Newton method, therefore, it is sufficient to assume that the appropriate principle leading minors of g'' are strictly positive definite.

5.4 Convergence Theorems for Block Methods. We will now apply the results of Chapter III and the two previous sections to give complete convergence theorems for several block methods. Our first theorem applies to the cyclic block Gauss-Seidel method. Let $H_1, \ldots, H_k$ be given subspaces such that

$$\hat{H}_0 \oplus \cdots \oplus \hat{H}_{k-1} = E^n.$$

Set

(5.4.1) $$H_p = \hat{H}_j, \qquad j = p \pmod k$$

suppose that $a_p \geqq 0$, $e^p \in H_p$ satisfy

$$g'(u^p - a_p e^p)h = 0, \qquad \text{for all } h \in H_p,$$

and set

(5.4.2) $$u^{p+1} = u^p - w_p a_p e^p.$$

When the H_j are coordinate subspaces this is the cyclic block Gauss-Seidel algorithm. However, for any fixed set of subspaces we have the following result.

<u>Theorem 5.4.1.</u> Assume that g is uniformly pseudo-convex in L_0 and the sequence of iterates is given by (5.4.2). Then the iterates will be well-defined and converge to x^*, the unique minimum of g in L_0, if for some forcing function d,

(a) w_p satisfies

(5.4.3) $$d(g'(u^p)e^p) \leqq w_p \leqq 1,$$

(b) or if g' is Lipschitz continuous on L_0 with Lipschitz constant C, uniformly pseudo-convex with a forcing function ct for $c > 0$, ϵ_p satisfies $d(g'(u^p)e^p) \leqq \epsilon_p \leqq 1$, and w_p satisfies

(5.4.4) $$1 \leqq w_p \leqq 1 + (c/C)^{\frac{1}{2}}(1-\epsilon_p),$$

(c) or if g'' exists and is constant in L_0 and w_p satisfies

$$d(g'(u^p)e^p) \leqq w_p \leqq 2 - d(g'(u^p)e^p). \tag{5.4.5}$$

Proof: The existence of x^* follows from our assumption that L_0 is compact, and the uniqueness of x^* from the uniform pseudo-convexity of g on L_0. Since $e^p \in H_p$, (5.4.2) implies $g'(u^p - a_p e^p)e^p = 0$ and by Theorem 1.3.1 this implies $g(u^p - a_p e^p) \leqq g(u^p)$. Pseudo-convexity then implies $g'(u^p)e^p \geqq 0$ and thus (5.4.2) represents a relaxed Curry iteration. Applying respectively, Theorem 3.2.1, Corollary 1 of Theorem 3.4.1, or Corollary 3 of Theorem 3.3.1 in parts (a), (b), or (c) we may conclude the iterates u^p are well-defined and $g'(u^p)e^p \to 0$. From Theorem 5.3.1 $P_p\,[g'(u^p)^T] \to 0$, where P_p is the orthogonal projection of E^n to H_p. But the sequence of subspaces $\{H_p\}$ is uniformly linearly independent, by Theorem 2.4.3 $\|u^p - u^{p+1}\| \to 0$, and hence by Theorem 5.2.1, $g'(u^p) \to 0$. Since x^* is the only solution of $g'(u) = 0$, the iterates must converge to x^*.

Note that when all the subspaces H_p are one dimensional this theorem is a result for the usual or point Gauss-Seidel method. For the point method, part (a) has already been shown to hold in Theorem 4.1.1 assuming only property S instead of the much stronger condition of uniform pseudo-convexity. For $w_p \geqq 1$, parts (b), (c) and the theorem below represent our best result, even if the blocks are one dimensional.

We now give a local convergence result for this algorithm.

<u>Theorem 5.4.2.</u> Assume an interior local minimum x^* of g in D exists. If in a neighborhood N of x^*, g'' exists, is continuous, and satisfies

$$c\|h\|^2 \leqq g''(u)hh \leqq C\|h\|^2, \qquad h \in E^n,\ u \in N, \tag{5.4.6}$$

then for any $0 < \epsilon \leqq 1$ if u^0 is sufficiently close to x^* then w_p may be taken to satisfy

$$\epsilon \leqq w_p \leqq 2-\epsilon$$

and the iterates (5.4.2) will be well-defined and converge to x^*.

<u>Proof:</u> It follows from Corollary 2 of Theorem 3.4.1 that the iterates are well-defined, remain in N, and $g'(u^p)e^p \to 0$. As in the proof of Theorem 5.4.1 this implies $g'(u^p) \to 0$. By (5.4.6) x^* is the only local minimim of g in N and the iterates must converge to x^*.

Our next theorem is a result about the block Gauss-Seidel-Newton method when the subspaces $H_0, \ldots, H_{k-1}$ are coordinate subspaces. We only assume, however, that the subspaces $H_0, \ldots, H_{k-1}$ are fixed, satisfy

$$H_o \oplus \cdots \oplus H_{k-1} = E^n$$

and

(5.4.7) $$H_p = H_j, \; j = p (\mathrm{mod}\; k).$$

Instead of solving

(5.4.8) $$g'(u)h = 0, \qquad u \in u^p \oplus H_p, \quad h \in H_p,$$

exactly we take one Newton step from u^p towards the solution. As with Newton's method itself the algorithm must be modified and we use the Armijo proceedure. Thus

(5.4.9) $$u^{p+1} = u^p - w_p B_p g'(u^p)^T$$

where B_p is given by (5.3.4), and w_p is the first number in the sequence $1, \frac{1}{2}, \frac{1}{4}, \ldots$ to satisfy

(5.4.10) $$g(u^p) - g(u^p - w_p B_p g'(u^p)^T) \geqq \tfrac{1}{2} w_p g'(u^p) B_p g'(u^p)^T.$$

Theorem 5.4.3. Suppose g'' exists, is continuous and strictly positive definite in L_o, let $\{H_p : p=0,1,\ldots\}$ satisfy (5.4.7), and let u^{p+1} be given by (5.4.9) where w_p is the first element in the sequence $1, \frac{1}{2}, \frac{1}{4}, \ldots$ to satisfy (5.4.10). Then the iterates are well-defined and converge to the minimim of g in L_o.

Proof: We are applying the Armijo algorithm to the iteration $\bar{u}^p = u^p - B_p g'(u^p)^T$ which satisfies (3.5.4) and by Theorem 3.5.1 the iterates are well-defined and $g'(u^p)e^p \to 0$. By Theorem 5.3.2, $P_p(g'(u^p)^T) \to 0$. The sequence $\{H_p\}$ is uniformly linearly independent and thus by Theorem 5.2.1

$g'(u^p) \to 0$. Since the minimum x^* is the only solution of $g'(u) = 0$ the sequence $\{u^p\}$ converges to x^*.

The equation (5.4.8) is in general difficult to solve and even using one Newton step from u^p requires inverting a matrix at each iteration. We consider next an algorithm which takes one gradient step towards solving (5.4.8), and thus is simple to apply. When g is Lipschitz continuous we set

$$(5.4.11) \qquad u^{p+1} = u^p - (w_p/C)P_p(g'(u^p)^T)$$

where C is a Lipschitz constant for g' and w_p satisfies

$$(5.4.12) \qquad d(\|P_p[g'(u^p)^T]\|) \leq w_p \leq 2 - d(\|P_p[g'(u^p)^T]\|)$$

for a forcing function d. We then have

<u>Theorem 5.4.4.</u> Suppose g' has isolated zeros and is Lipschitz continuous in L_0, $\{H_p : p = 0,1,\ldots\}$ is defined by (5.4.7) and u^{p+1} is given by (5.4.11). Then the iterates are well-defined and converge to a solution of $g'(u) = 0$.

<u>Proof:</u> From Theorem 3.3.2 the iterates are well-defined and $g'(u^p)e^p \to 0$. Since $g'(u^p)e^p = \|P_p(g'(u^p)^T)\|$ and the sequence of directions is uniformly linearly independent, by Theorem 5.2.1 we have $g'(u^p) \to 0$. The convergence of the iterates follows from Theorem 2.2.1.

REFERENCES

1. M. Altman, Connection between gradient methods and Newton's method for functionals, Bull. Acad. Polon. Sci. Ser. Sci. Math. Astronom. Phys. 9 (1961), 877-880.

2. ________, Generalized gradient methods of minimizing a functional, Bull. Acad. Polon. Sci. Ser. Sci. Math. Astronom. Phys. 14 (1966), 313-318.

3. L. Armijo, Minimization of functions having Lipschitz continuous first partial derivatives, Pacific J. Math. 16 (1966), 1-3.

4. L. Bers, On mildly nonlinear partial difference equations of elliptic type, J. Res. Nat. Bur. Standards Sect. B. 51 (1953), 229-236.

5. A. Cauchy, Methods generale pour la resolution des systemes d'equations simultanees, C. R. Acad. Sci. Par. 25 (1847), 536-538.

6. J. B. Crockett and H. Chernoff, Gradient methods of maximization, Pacific J. Math. 5 (1955), 33-50.

7. H. B. Curry, The method of steepest descent for nonlinear minimization problems, Quart. Appl. Math. 2 (1944), 258-263.

8. J. Dieudonne, Foundations of modern analysis, Academic Press, New York, 1960.

9. N. Dunford and J. T. Schwartz, Linear operators part I, Interscience, New York, 1964.

10. W. Fenchel, Convex cones, sets and functions, Princeton Univ., 1953.

11. G. E. Forsythe and W. R. Wasow, Finite-difference methods for partial differential equations, John Wiley & Sons, New York, 1960.

12. I. M. Glazman, Gradient relaxation for non-quadratic functions, Soviet Math. Dokl. 5 (1964), 210-213.

13. ______________, *Relaxation on surfaces with saddle points*, Soviet Math. Dokl. 6 (1965), 487-490.

14. A. A. Goldstein, *Cauchy's method of minimization*, Numer. Math. 4 (1962), 146-150.

15. ______________, *Minimizing functionals on Hilbert space*, Computing Methods in Optimization Problems, A. V. Balakrishnan and L. W. Neustadt, eds., Academic Press, New York (1964), 159-166.

16. ______________, *On steepest descent*, SIAM J. Control 3 (1965), 147-151.

17. ______________, *Minimizing functionals on normed-linear spaces*, SIAM J. Control 4 (1966), 81-89.

18. ______________, *Constructive real analysis*, Harper & Row, New York, 1967.

19. M. N. Jakovlev, *The solution of nonlinear equations by an iteration method*, Soviet Math. Dokl. 5 (1964), 697-699.

20. W. Kahan, *Relaxation methods for semi-definite systems*, Technical report no. CS45, Computer Science Department, Stanford University, 1966.

21. L. V. Kantorovich, *The method of successive approximations for functional equations*, Acta. Math. 71 (1939), 63-77.

22. ______________ and G. P. Akilov, *Functional analysis in normed spaces*, transl. D. E. Brown, A. P. Robertson, MacMillan, New York, 1964.

23. J. Kiefer, *Sequential minimax search for a maximum*, Proc. Amer. Math. Soc. 4 (1953), 502-506.

24. J. Ljubic, *General theorems on quadratic relaxation*, Soviet Math. Dokl. 6 (1965), 588-591.

25. O. L. Mangasarian, *Pseudo-convex functions*, SIAM J. Control 3 (1965), 281-290.

26. G. J. Minty, On the monotonicity of the gradient of a convex function, Pacific J. Math. 14 (1964), 243-247.

27. M. J. Nashed, The convergence of the method of steepest descents for nonlinear equations with variational or quasi-variational operators, J. Math. Mech. 13 (1964), 765-794.

28. ____________, Supportably and weakly convex functionals with applications to approximation theory and nonlinear programming, J. Math. Anal. Appl. 18 (1967), 504-521.

29. D. J. Newman, Location of the maximum on unimodal surfaces, J. Assoc. Comput. Mach. 12 (1965), 395-398.

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

31. A. M. Ostrowski, On the linear iteration procedures for symmetric matrices, Rend. Mat. e Appl. (5) 14 (1954), 140-163.

32. ______________, Solution of equations and systems of equations, Second Edition, Academic Press, New York, 1966.

33. ______________, Contributions of the theory of steepest descent I, MRC Report No. 615, Mathematics Research Center, Madison, 1966.

34. B. T. Poljak, Existence theorems and convergence of minimizing sequences in extremum problems with restrictions, Soviet Math. Dokl. 7 (1966), 72-75.

35. J. Ponstein, Seven kinds of convexity, SIAM Rev. 9 (1967), 115-119.

36. H. H. Rosenbrock, An automatic method for finding the greatest or least value of a function, Comput. J. 3 (1960), 175-184.

37. S. Schechter, <u>Relaxation methods for linear equations</u>, Comm. Pure Appl. Math. 12 (1959), 313-335.

38. ____________, <u>Iteration methods for nonlinear problems</u>, Trans. Amer. Math. Soc. 104 (1962), 179-189.

39. R. V. Southwell, <u>Relaxation methods in theoretical physics</u>, Oxford University Press, Oxford, 1946.

40. M. M. Vainberg, <u>On the convergence of the method of steepest descents for nonlinear equations</u>, Soviet Math. Dokl. 1 (1960), 1-4.

41. ____________, <u>Variational methods for the study of nonlinear operators</u>, transl. A. Feinstein, Holden-Day, San Francisco, 1964.

42. R. S. Varga, <u>Matrix iterative analysis</u>, Prentice-Hall Englewood Cliffs, N. J., 1962.

43. A. Wilanski, <u>Functional analysis</u>, Blaisdell, New York, 1964, p. 23.

44. D. J. Wilde, <u>Optimum seeking methods</u>, Prentice-Hall, Englewood Cliffs, N. J., 1964.

45. E. H. Zarantonello, <u>Solving functional equations by contractive averaging</u>, MRC Report No. 160, Mathematics Research Center, Madison, 1960.