A COMPARATIVE STUDY OF SEVERAL GENERAL CONVERGENCE CONDITIONS FOR ALGORITHMS MODELED BY POINT-TO-SET MAPS

by

S. Tishyadhigama, E. Polak and R. Klessig

Memorandum No. UCB/ERL M77/78

16 November 1977

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

A COMPARATIVE STUDY OF SEVERAL GENERAL CONVERGENCE CONDITIONS FOR ALGORITHMS MODELED BY POINT-TO-SET MAPS

S. Tishyadhigama, E. Polak and R. Klessig†

Department of Electrical Engineering and Computer Sciences
and the Electronics Research Laboratory
University of California, Berkeley, California 94720

ABSTRACT

A general structure is established that allows the comparison of various conditions that are sufficient for convergence of algorithms that can be modeled as the recursive application of a point-to-set map. This structure is used to compare several earlier sufficient conditions as well as three new sets of sufficient conditions. One of the new sets of conditions is shown to be the most general in that all other sets of conditions imply this new set. This new set of conditions is also extended to the case where the point-to-set map can change from iteration to iteration.

---

†On leave from Bell Labs., Holmdel, New Jersey 07733.

Research sponsored by the National Science Foundation Grant ENG73-08214A01, and the National Science Foundation (RANN) Grant ENV76-05264.

1. Introduction

In recent years, the study of optimization algorithms has included a substantial effort to identify properties of algorithms that will guarantee their convergence (in some sense) e.g. [1]-[29]. A number of these results have used an abstract algorithm model that consists of the recursive application of a point-to-set map. It is this type of result with which we are concerned in this paper and, in particular with the results presented in [13], [16], [21], [24] and [29].

We have two purposes. First, we wish to introduce three new general convergence results. Second, we wish to identify the relationships among the general convergence results including both our new results and previously published results.

In order to compare results, it is necessary to have a common framework. Unfortunately, different authors have used slightly different abstract algorithm models and have arrived at slightly different conclusions partly because they have used somewhat different concepts of convergence. Thus, before a comparison can be made, it is necessary to establish a common framework and then to translate the various theories into this framework. Our approach to this task is as follows. In Section 2, we define an abstract algorithm model and formally define a concept of convergence for this model. Our new convergence results establish that certain conditions are sufficient for the algorithm model to be convergent in the sense of our concept of convergence. The earlier results use a similar approach, but occasionally differ from each other by the algorithm model and concept of convergence used. We take the essential features of these earlier sufficient conditions and use these to create analogous conditions that are sufficient in our present framework. We then establish relationships between the various sufficient conditions by showing which conditions imply other conditions.

In view of our approach to the interpretation of earlier work, we make no claim that, and the reader should not infer that, the contents of this paper fully describe the various earlier results. When we associate an author's name with a set of sufficient conditions, we mean that the original conditions from which we derived the conditions in question, were first proposed by that author. The interested reader can find all of the new results stated in this paper in [26]. [26] also shows how the sufficient conditions used in this paper are derived from the original sufficient conditions.

Section 3 contains the main results of this paper. These results are summarized by Fig. 1. Each box represents a set of conditions and an arrow indicates that the conditions at the tail of the arrow imply the comditions at the head. We have included in Section 3 results that show that under special conditions, some sets of sufficient conditions are equivalent. The most important of these special cases is when the cost (or surrogate cost) function, c, is continuous. The special cases are noted in Fig. 1.

In Section 4 we illustrate how the sufficient conditions presented in Section 3 can be modified to apply to an algorithm model which may use a different point-to-set map at each iteration. We do this by extending the most general sufficient conditions of Section 3.

Finally, in the Appendix we present some counterexamples to show that there are meaningful differences between the sets of sufficient conditions.

2. Framework for Comparison and Preliminaries

In this section, we present an abstract algorithm model and define a concept of convergence. In addition, we present some results and notation that will be extensively used in the sequel.

(2.1) Definition: Ω is a Hausdorff topological space that satisfies the first axiom of constability. $\Delta \subset \Omega$ is called the set of desirable points. ¤

(2.2) Remark: The set Δ consists of points that we will accept as "solutions" to the problem being solved by the algorithm. For example, it may consist of all points satisfying a necessary condition of optimality. Ω is usually taken as the set of feasible points for a problem. Thus, Ω may be a subset of a larger topological space. If such is the case, the relative topology on Ω is used. ¤

(2.3) Algorithm Model: Let $A:\Omega \to 2^{\Omega} \sim \phi$ where 2^{Ω} denotes all subsets of Ω.

Step 0: Set $i=0$. Choose any $z_o \in \Omega$.

Step 1: Choose any $z_{i+1} \in A(z_i)$.

Step 2: Set $i=i+1$ and go to Step 1. ¤

(2.4) Remark: Algorithm Model (2.3) has no provision for stopping and thus always generates an infinite sequence. However, many algorithms have stopping tests and stop only when $z_i \in \Delta$. This can be accounted for in (2.3) by defining $A(z_i) = \{z_i\}$ whenever z_i satisfies the stopping condition. Thus our analyses are shortened because we do not have to consider the trivial finite sequence case. ¤

We now state our concept of convergence.

(2.5) <u>Definition</u>: We say the Algorithm Model (2.3) is <u>convergent</u>[†] if the accumulation points of any sequence $\{z_i\}$ constructed by (2.3) are in Δ. ⌑

(2.6) <u>Remark</u>: We make no assumption that $\{z_i\}$ will have accumulation points. Thus, it is possible for (2.3) to be convergent and for $\{z_i\}$ to have no accumulation points. For example, for an optimization problem with no solution, defining Δ as the set of solutions means that $\Delta = \phi$ and the applications of a convergent algorithm results in a sequence $\{z_i\}$ that cannot have accumulation points. ⌑

The definitions (2.1), (2.3) and (2.5) constitute the common structure within which we shall carry out our analysis.

To conclude this section, we establish some notation and state some results that will be useful later. All of the sufficient conditions in Section 3 assume the existence of a function $c:\Omega \to \mathbb{R}^1$ and imply that $c(z') \leq c(z)\ \forall z' \in A(z),\ z \in \Omega$. In most applications, c is the cost function in the optimization problem. Since c is used frequently, we establish the following notational convention and state a lemma whose proof is straightforward and therefore omitted.

(2.7) <u>Notation</u>: The symbol c always represents a function $c:\Omega \to \mathbb{R}^1$. ⌑

(2.8) <u>Lemma</u>: Suppose $\{z_i\} \subset \Omega$ is such that $c(z_{i+1}) \leq c(z_i)\ i = 0,1,\ldots$. Then $\{c(z_i)\}$ converges if and only if some subsequence of $\{c(z_i)\}$ converges. ⌑

The following properties of first countable Hausdorff topological spaces are well known (see [31]) and are stated here for reference.

[†] When we say that a <u>sequence</u> $\{y_i\}$ converges, we still mean it in the usual sense, i.e., for some $\hat{y}$, $y_i \to \hat{y}$ as $i \to \infty$.

(2.9) Facts:

(i) If z is an accumulation point of $\{z_i\} \subset \Omega$, then there exists a subsequence of $\{z_i\}$ converging to z.

(ii) For each $z \in \Omega$, there exists a sequence of neighborhoods of z, $\{U_i\}$, such that $U_{i+1} \subset U_i$ $i = 0,1,\ldots$ and $z_i \in U_i$ $i = 0,1,\ldots$ implies that $z_i \to z$.

(iii) For any sequence $\{z_i\} \subset S \subset \Omega$ with S compact, there exists a subsequence converging to a point in S.

3. Comparison of Sufficient Conditions

In this section we present a number of sets of sufficient conditions for Algorithm Model (2.3) to be convergent in the sense of (2.5). Three of these sets, (3.3), (3.18) and (3.38) are new while the remaining sets have been extracted from previous results. We start by proving (3.3) is sufficient. Then we establish the relationships among the various sets of conditions as indicated in Fig. 1. As can be seen from this Figure, all conditions ultimately imply (3.3). Thus, all conditions presented are indeed sufficient.

(3.1) Definition: $c(\cdot)$ is said to be locally bounded from below at z if there exist a neighborhood U of z and $b \in \mathbb{R}^1$ (possibly depending on z) such that

(3.2) $c(z') \geq b \quad \forall z' \in U.$ ¤

(3.3) Conditions:

(i) $c(\cdot)$ is locally bounded from below on $\Omega \sim \Delta$.

(ii) $c(z') \leq c(z) \quad \forall z' \in A(z),\ z \in \Delta$.

(iii) For each $z \in \Omega \sim \Delta$, if $\{x_i\} \subset \Omega$ is such that $x_i \to z$ and $c(x_i) \to c^*$, then there exists an integer N such that

(3.4) $c(y) < c^* \quad \forall y \in A(x_N)$. ◻

(3.5) <u>Theorem</u>: If Conditions (3.3) hold, then Algorithm Model (2.3) is convergent.

<u>Proof</u>: First we note that (3.3) (iii) implies that $c(z') < c(z)$ $\forall z' \in A(z)$, $z \in \Omega \sim \Delta$ and hence, together with (3.3) (ii), that

(3.6) $c(z') \leq c(z) \quad \forall z' \in A(z),\ z \in \Omega$.

Let $\{z_i\}$ be any infinite sequence constructed by the Algorithm Model (2.3) and suppose that $z^* \in \Omega \sim \Delta$ is an accumulation point of $\{z_i\}$. We shall establish a contradiction. Let $K \subset \{0,1,\ldots\}$ index a subsequence such that $z_i \underset{K}{\to} z^*$. Because of (3.3) (i) and (3.6), there exists a c^* such that $c(z_i) \underset{K}{\to} c^*$. Lemma (2.8) then implies that $c(z_i) \to c^*$ and it follows that

(3.7) $c(z_i) \geq c^*$ for $i = 0,1,\ldots$.

On the other hand, using $\{z_i\}_{i \in K}$ for $\{x_i\}$ in (3.3) (iii) yields the existence of N such that $c(z_{N+1}) < c^*$ which contradicts (3.7). Thus, we have shown that any accumulation point of $\{z_i\}$ must be in Δ. The proof is now complete. ◻

(3.8) <u>Conditions</u> (R. Meyer [16]):

(i) $c(\cdot)$ is locally bounded from below on $\Omega \sim \Delta$.

(ii) $c(z') \leq c(z) \quad \forall z' \in A(z),\ z \in \Delta$.

(iii) For each $z \in \Omega \sim \Delta$, if $\{x_i\}, \{y_i\} \subset \Omega$ are such that $x_i \to z$, $y_i \in A(x_i)$, $c(x_i) \to c^*$ and $c(y_i) \to \bar{c}$, then $\bar{c} < c^*$. ◻

(3.9) Theorem: Conditions (3.8) imply Conditions (3.3).

Proof: Suppose that conditions (3.8) hold. Then conditions (3.3)(i) and (3.3)(ii) hold since these are identical to (3.8)(i) and (3.8)(ii). Conditions (3.8)(ii) and (3.8)(iii) imply that

$$c(z') \leq c(z) \quad \forall z' \in A(z),\ z \in \Omega.$$

Now let $z \in \Omega \sim \Delta$ and let $\{x_i\} \subset \Omega$ be such that $x_i \to z$ and $c(x_i) \to c^*$. We now assume that (3.3)(iii) does not hold and establish a contradiction. If (3.3)(iii) does not hold, there exist $y_i \in A(x_i)$ $i = 0,1,\ldots$ such that $c(y_i) \geq c^*$ $i = 0,1,\ldots$. Thus, we have

$$c(x_i) \geq c(y_i) \geq c^* \quad i = 0,1,\ldots$$

which implies that $\lim c(y_i) = c^*$ and this contradicts (3.8)(iii). The proof is now complete. ¤

(3.10) Definition: The pair (c,A) is locally uniformly monotonic at z if there exists $\bar{\delta}(z) > 0$ (possibly depending on z) and a neighborhood $U(z)$ of z such that

(3.11) $$c(z'') - c(z') \leq -\bar{\delta}(z) \quad \forall z'' \in A(z'),\ z' \in U(z).$$ ¤

(3.12) Remark: Polak [21] was the first to use local uniform monotonicity. G. Meyer [13] later generalized the Polak conditions by using local boundedness from below of $c(\cdot)$ instead of boundedness from below. ¤

(3.13) Conditions (G. Meyer [13]):

(i) $c(\cdot)$ is locally bounded from below on $\Omega \sim \Delta$.

(ii) $c(z') \leq c(z) \quad \forall z' \in A(z),\ z \in \Delta$.

(iii) The pair (c,A) is locally uniformly monotonic on $\Omega \sim \Delta$.

¤

(3.14) Theorem:

a) Conditions (3.13) imply Conditions (3.8)

b) Conditions (3.8) with the additional assumption that $c(\cdot)$ is locally bounded imply Conditions (3.13).

Proof:

a) Suppose that Conditions (3.13) hold. Then Conditions (3.8)(i) and (3.8)(ii) hold since these are identical to Conditions (3.13)(i) and (3.13)(ii). Consider $z \in \Omega \sim \Delta$ and let $\bar{\delta}(z)$ and $U(z)$ be as in Definition (3.10). Let $\{x_i\}$, $\{y_i\}$, c^* and $\bar{c}$ be as given by (3.8)(iii). Then (3.13)(iii) implies that there exists an integer N such that $c(y_i) \leq c(x_i) - \bar{\delta}(z)$ for all $i \geq N$. Hence, $\bar{c} = \lim c(y_i) \leq \lim c(x_i) - \bar{\delta}(z) < c^*$ and (3.8)(iii) is established.

b) As pointed out above, (3.13)(i) and (3.13)(ii) are identical to (3.8)(i) and (3.8)(ii). To complete the proof, we assume (3.8) and local boundedness of $c(\cdot)$ hold, but that (3.13)(iii) does not hold and establish a contradiction.

Suppose $z \in \Omega \sim \Delta$. By (2.9)(ii), there exists a sequence $\{U_i\}$ of neighborhoods of z such that $U_{i+1} \subset U_i$ and $z_i \in U_i$ implies that $z_i \to z$. If (3.13)(iii) does not hold, there exist $\{\delta_i\}$, $\{x_i\}$ and $\{y_i\}$ such that $\delta_i > 0$, $\delta_i \to 0$, $x_i \in U_i$, $y_i \in A(x_i)$ and

(3.15) $c(x_i) \geq c(y_i) > c(x_i) - \delta_i \quad i = 0,1,\ldots$.

Since c is locally bounded, there exists $b \geq 0$ such that

(3.16) $|c(x_i)| \leq b \quad i = 0,1,\ldots$.

Thus, there exist a subsequence $\{c(x_i)\}_{i \in K}$ and c^* such that $c(x_i) \underset{K}{\to} c^*$. (3.15) then implies that

(3.17) $\lim_K c(y_i) = \lim_K c(x_i) = c^*.$

Since $x_i \in U_i$, $x_i \to z$ and $y_i \in A(x_i)$, (3.8)(iii) requires that $\lim_K c(y_i) < c^*$ which contradicts (3.17). Thus (3.13)(iii) must hold. The proof is now complete.

¤

(3.18) Conditions: There exists $\delta:\Omega \to \mathbb{R}^+$ with the following properties.

(i) $c(\cdot)$ is locally bounded from below on $\Omega \sim \Delta$.

(ii) $c(z') - c(z) \leq -\delta(z) \leq 0 \quad \forall z' \in A(z),\ z \in \Omega.$

(iii) For each $z \in \Omega \sim \Delta$, if $\{x_i\} \subset \Omega$ is such that $x_i \to z$, then $\sum_{i=0}^{\infty} \delta(x_i) = \infty.$

¤

(3.19) Lemma: Suppose $z \in \Omega$ and $\delta:\Omega \to \mathbb{R}^+$ is such that $\sum_{i=0}^{\infty} \delta(x_i) = \infty$ for all sequences $\{x_i\}$ that converge to z. Then there exist $V(z)$ a neighborhood of z, and $\underline{\delta} > 0$ such that $\delta(z') \geq \underline{\delta}$ for all $z' \in V(z)$.

Proof: Suppose the lemma is false. Then we can find $\{z_i\}$ with $z_i \to z$ and $\delta(z_i) \to 0$. Define the map n as

(3.20) $n(i) = \min\{j \mid j \geq i+1,\ \delta(z_j) \leq (\tfrac{1}{2})^i\}.$

n is well defined since $\delta(z_i) \to 0$. Let the sequence $\{x_i\}$ be defined as $x_o = z_o$, $x_1 = z_{n(o)}$, $x_2 = z_{n(n(o))}$, $x_3 = z_{n(n(n(o)))}$ and so forth. Then $\sum_{i=0}^{\infty} \delta(x_i) \leq \sum_{i=0}^{\infty} (\tfrac{1}{2})^i = 2$. But $x_i \to z$ and thus by the hypothesis, $\sum_{i=0}^{\infty} \delta(x_i) = \infty$. So we have a contradiction and the lemma must be true.

¤

(3.21) Theorem: Conditions (3.18) imply Conditions (3.13) and vice versa (i.e. (3.18) $\Leftrightarrow$ (3.13)).

Proof: ($\Rightarrow$) Clearly, (3.18)(i) implies (3.13)(i) and (3.18)(ii) implies (3.13)(ii). Let $z \in \Omega \sim \Delta$. Conditions (3.18)(ii), (3.18)(iii) and Lemma (3.19) imply that there eixst $V(z)$ a neighborhood of z and $\underline{\delta} > 0$ such that

$$c(z'') - c(z') \leq -\delta(z') \leq -\underline{\delta} \quad \forall z'' \in A(z'),\ z' \in V(z).$$

Consequently, (c,A) is locally uniformly monotonic at z. Condition (3.13)(iii) is therefore established.

($\Leftarrow$) Condition (3.13)(i) is identical to (3.18)(i). For each $z \in \Omega \sim \Delta$, let $\bar{\delta}(z) > 0$ and $U(z)$ be as given in Definition (3.10). Define $\delta:\Omega \to \mathbb{R}^1$ by

(3.22) $\delta(z) \triangleq \inf\{c(z)-c(z') \mid z' \in A(z)\}$.

Conditions (3.13)(ii) and (3.13)(iii) imply that $\delta(z) \geq 0$ for all $z \in \Omega$ and thus (3.18)(ii) holds. Now, for $z \in \Omega \sim \Delta$, $\delta(z') \geq \bar{\delta}(z) > 0$ for all $z' \in U(z)$. Therefore, whenever $x_i \to z$, $\sum_{i=0}^{\infty} \delta(x_i) = \infty$ and (3.18)(iii) holds. The proof is now complete. ¤

(3.23) <u>Conditions</u> (Polak [21]):

(i) $c(\cdot)$ is either lower semicontinuous on $\Omega \sim \Delta$ or bounded from below on Ω.

(ii) $c(z') \leq c(z) \quad \forall z' \in A(z),\ z \in \Delta$

(iii) (c,A) is locally uniformly monotonic on $\Omega \sim \Delta$. ¤

(3.24) <u>Theorem</u>:

a) Conditions (3.23) imply Conditions (3.13).

b) Conditions (3.13) with the additional assumption that $c(\cdot)$ is lower semicontinuous imply Conditions (3.23).

<u>Proof</u>:

a) Clearly, (3.23)(i) implies that $c(\cdot)$ is locally bounded from below, Thus a) is true because (3.23)(ii) and (3.23)(iii) are identical to (3.13)(ii) and (3.13)(iii).

(b) When $c(\cdot)$ is lower semicontinuous, (3.23)(i) is satisfied and therefore Conditions (3.13) imply Conditions (3.23). Thus, b) is true.

¤

(3.25) <u>Conditions</u> (R. Meyer [16]): $\tilde{\delta}:\Omega \to \mathbb{R}^+$ is such that

(i) $c(\cdot)$ is either lower semicontinuous on $\Omega \sim \Delta$ or bounded from below on Ω.

(ii) $c(z') \leq c(z) - \tilde{\delta}(z) \quad \forall z' \in A(z),\ z \in \Omega$

(iii) For each $z \in \Omega$, if $z_i \to z$ and $\tilde{\delta}(z_i) \to 0$, then $\tilde{\delta}(z) = 0$.

(iv) $\{z' \in \Omega | \tilde{\delta}(z') = 0\} \subset \Delta$.

¤

(3.26) <u>Theorem</u>:

a) Conditions (3.25) imply Conditions (3.23).

b) Conditions (3.23) imply Conditions (3.25) when Δ is closed.

<u>Proof</u>:

a) Conditions (3.25)(i) and (3.23)(i) are identical. Since $\tilde{\delta}$ has only nonnegative values, (3.25)(ii) implies (3.23)(ii). Now suppose $z \in \Omega \sim \Delta$. By (3.25)(iv),

(3.27) $\tilde{\delta}(z) > 0$.

Let $\{U_i\}$ be a sequence of neighborhoods of z satisfying (2.9)(ii). Assume that (3.25) holds, but (3.23)(iii) does not hold at z. Then there exist $\{z_i\}$, $\{y_i\}$ and $\{\delta_i\}$ such that $z_i \in U_i$ (and hence $z_i \to z$), $y_i \in A(z_i)$, $\delta_i \downarrow 0$ and

$$c(y_i) > c(z_i) - \delta_i \quad i = 0,1,\ldots .$$

Combining this inequality with (3.25)(ii) yields

$$c(z_i) - \tilde{\delta}(z_i) \geq c(y_i) > c(z_i) - \delta_i \quad i = 0,1,\ldots$$

Rearranging, we have that $0 \leq \tilde{\delta}(z_i) < \delta_i \quad i = 0,1,\ldots$ which means that $\tilde{\delta}(z_i) \to 0$. But since $z_i \to z$, (3.25)(iii) requires $\tilde{\delta}(z) = 0$ which contradicts (3.27) and the proof of a) is complete. b) As before, (3.23)(i) and (3.25)(i) are identical. Now define $\tilde{\delta}: \Omega \to \mathbb{R}^+$ by

$$(3.28) \quad \tilde{\delta}(z) = \begin{cases} 0 & \text{if } z \in \Delta \\ \inf\{c(z)-c(z') \mid z' \in A(z)\} & \text{if } z \in \Omega \sim \Delta. \end{cases}$$

Then (3.23)(iii) implies that $\tilde{\delta}(z) \geq 0$ for all $z \in \Omega$ and (3.23)(ii) along with (3.28) imply (3.25)(ii). Now consider $z \in \Omega \sim \Delta$. Because Δ is closed and because of (3.23)(iii), there exists a neighborhood U of z and $\delta > 0$ such that $U \subset \Omega \sim \Delta$ and $\tilde{\delta}(z') \geq \delta > 0$ for all $z' \in U$. Thus if we have $z_i \to z^* \in \Omega$ and $\tilde{\delta}(z_i) \to 0$, we must have that $z^* \in \Delta$ which implies that $\tilde{\delta}(z^*) = 0$. Consequently, (3.25)(iii) holds. Also, $\Delta = \{z' \in \Omega \mid \tilde{\delta}(z') = 0\}$ and therefore (3.25)(iv) holds. The proof is now complete.

¤

(3.29) <u>Remark</u>: If Conditions (3.23) hold, they will also hold if Δ is replaced by any Δ' such that $\Delta \subset \Delta'$. Because of this latitude in selecting Δ in (3.23), Conditions (3.23) and (3.25) are not equivalent. However, if one chooses Δ to be as small as possible so that Conditions (3.23) hold for a given $c(\cdot)$ and $A(\cdot)$, then Corollary (3.30) shows that Conditions (3.23) are equivalent to Conditions (3.25).

¤

(3.30) <u>Corollary</u>: Suppose that Conditions (3.23) are satisfied. In addition, suppose that $\Delta = \Omega \sim \Lambda$ where $\Lambda \triangleq \{z \mid (c,A)$ is locally uniformly monotonic at $z\}$. (Conditions (3.23) only imply that $\Omega \sim \Delta \subset \Lambda$.) Then Conditions (3.25) are satisfied.

Proof: In view of Theorem (3.26), it will suffice to show that Δ is closed. We assume that Δ is not closed and establish a contradiction. If Δ is not closed, there exists z^*, an accumulation point of Δ, such that $z^* \notin \Delta$ (i.e. $z^* \in \Lambda$). Thus there exist a neighborhood N of z^* and a $\delta > 0$ such that

$$c(z'') - c(z') \leq -\delta \quad \forall z'' \in A(z'),\ z' \in N.$$

Since $N \cap \Delta \neq \phi$, we can choose a $\tilde{z} \in N \cap \Delta$. But N is also a neighborhood of $\tilde{z}$ and the above inequality implies that (c,A) is locally uniformly monotonic at $\tilde{z}$. That is, $\tilde{z} \in \Lambda$ and we have $\tilde{z} \in \Delta \cap \Lambda = \phi$ which is a contradiction. Therefore, Δ is closed and the proof is complete.

$\square$

(3.31) Conditions (G. Meyer [13]):

(i) $c(\cdot)$ is lower semicontinuous on $\Omega \sim \Delta$.

(ii) $c(z') \leq c(z) \quad \forall z' \in A(z),\ z \in \Delta$.

(iii) For each $z \in \Omega \sim \Delta$, there exists a neighborhood U of z such that

$$c(z'') < c(z) \quad \forall z' \in A(z'),\ z' \in U.$$

$\square$

(3.32) Theorem:

a) Conditions (3.31) imply Conditions (3.3)

b) Conditions (3.3) with the additional assumption that $c(\cdot)$ is continuous imply Conditions (3.31).

Proof:

a) Assume that Conditions (3.31) hold. Then (3.3)(i) holds since lower semicontinuity implies local boundedness from below. Next, (3.3)(ii) is identical to (3.31)(ii). Let $z \in \Omega \sim \Delta$ and let $\{x_i\} \subset \Omega$ be such that

$x_i \to z$ and $c(x_i) \to c^*$. The lower semicontinuity of $c(\cdot)$ implies that

(3.33) $\quad c(z) \leq c^* = \lim c(x_i)$

There must exist an integer N such that $x_N \in U$ where U is given by (3.31)(iii). Hence

(3.34) $\quad c(y) < c(z) \quad \forall y \in A(x_N)$

Combining (3.33) and (3.34) shows that (3.3)(iii) holds.

b) Assume that Conditions (3.3) hold and that c is continuous. Since we are assuming that $c(\cdot)$ is continuous, Condition (3.31)(i) holds. Condition (3.31)(ii) holds since it is identical to (3.3)(ii). Let $z \in \Omega \sim \Delta$ and let $\{U_i\}$ be a sequence of neighborhoods of z that satisfies (2.9)(ii). If (3.31)(iii) does not hold. we can find $x_i \in U_i$ and $y_i \in A(x_i)$ such that

(3.35) $\quad c(y_i) \geq c(z) \quad i = 0,1,\ldots$.

By the construction of $\{U_i\}$, $x_i \to z$, and therefore the continuity of $c(\cdot)$ implies that $\lim c(x_i) = c(z)$. But then (3.35) contradicts (3.3)(iii) so we must have that (3.31)(iii) holds. The proof is now complete. ¤

(3.36) Definition: The composit map $c(A(\cdot))$ is super upper semicontinuous at z if for each $\varepsilon > 0$, there exist $\bar{z} \in A(z)$ and a neighborhood U of z such that

(3.37) $\quad c(z'') \leq c(\bar{z}) + \varepsilon \quad \forall z'' \in A(z'),\ z' \in U,$ ¤

(3.38) Conditions:

(i) $c(\cdot)$ is lower semicontinuous on $\Omega \sim \Delta$.

(ii) $c(z') \leq c(z) \quad \forall z' \in A(z),\ z \in \Delta$.

(iii) For each $z \in \Omega \sim \Delta$, there exists $\gamma(z) > 0$ (possibly depending on z) such that

$$c(z') \leq c(z) - \gamma(z) \quad \forall z' \in A(z).$$

(iv) $c(A(\cdot))$ is super uppersemicontinuous on $\Omega \sim \Delta$.

¤

(3.39) Theorem:

a) Conditions (3.38) imply Conditions (3.31).

b) Conditions (3.38) imply Conditions (3.23).

Proof: Suppose Conditions (3.38) hold.

a) Then Conditions (3.31)(i) and (3.31)(ii) hold because they are identical to (3.31)(i) and (3.38)(ii). Let $z \in \Omega \sim \Delta$ and let $\gamma(z) > 0$ be as given in (3.38)(iii). Set $\varepsilon = \frac{\gamma(z)}{2}$ in Definition (3.36) and denote the required neighborhood by U. Then, with $\bar{z} \in A(z)$, as in (3.36),

$$(3.40) \quad c(z'') \leq c(\bar{z}) + \frac{\gamma(z)}{2} \leq c(z) - \gamma(z) + \frac{\gamma(z)}{2} < c(z)$$

$$\forall z'' \in A(z'),\ z' \in U$$

and consequently, (3.31)(iii) holds.

b) Condition (3.38)(i) clearly implies (3.23)(i) and (3.38)(ii) is identical to (3.23)(ii). Let $z \in \Omega \sim \Delta$ and let $\gamma(z) > 0$ be as given by (3.38)(iii). Let $\bar{z} \in A(z)$ and U be as given by by Definition (3.36) with $\varepsilon = \frac{\gamma(z)}{3}$. Then we have

$$(3.41)\quad c(z'') \le c(\bar{z}) + \frac{\gamma(z)}{3} \le c(z) - \gamma(z) + \frac{\gamma(z)}{3} = c(z) - \frac{2\gamma(z)}{3}$$

$$\forall z'' \in A(z'),\ z' \in U$$

Since $c(\cdot)$ is lower semicontinuous, there exists a neighborhood U_1 of z with $U_1 \subset U$ such that

$$(3.42)\quad c(z) - \frac{\gamma(z)}{3} \le c(z') \quad \forall z' \in U_1$$

Combining (3.41) and (3.42) yields

$$(3.43)\quad c(z'') \le c(z') - \frac{\gamma(z)}{3} \quad \forall z'' \in A(z'),\ z' \in U_1.$$

Thus, (c,A) is locally uniformly monotonic on $\Omega \sim \Delta$, i.e., (3.23)(iii) holds. The proof is now complete.

¤

(3.44) <u>Conditions</u> (Polyak [24]):

(i) $c(\cdot)$ is lower semicontinuous on $\Omega \sim \Delta$.

(ii) $A(\cdot)$ is single valued (denoted by $a(\cdot)$) $\forall z \in \Omega$.

(iii) $c(a(z)) \le c(z) \quad \forall z \in \Delta$.

(iv) $c(a(z)) < c(z) \quad \forall z \in \Omega \sim \Delta$.

(v) $c(a(\cdot))$ is upper semicontinuous on $\Omega \sim \Delta$.

¤

(3.45) <u>Theorem</u>:

a) Conditions (3.44) imply Conditions (3.38)

b) Conditions (3.38) imply Conditions (3.44) when $A(\cdot)$ is a single valued map.

<u>Proof</u>: First we note that when $A(\cdot)$ $(=a(\cdot))$ is single valued, Definition (3.36) implies that $c(a(\cdot))$ is upper semicontinuous. Thus for $A(\cdot) = a(\cdot)$, (3.44)(i) and (3.38)(i) are equivalent, (3.44)(iii)

and (3.38)(ii) are equivalent, and (3.44)(v) and (3.38)(iv) are equivalent. Furthermore, when $A(\cdot) = a(\cdot)$, it is easy to see that (3.44)(iv) is equivalent to (3.38)(iii). Thus both a) and b) follow immediately. The proof is now complete. ¤

(3.46) Definition: $A(\cdot)$ is said to be closed at z if $z_i \to z$, $y_i \in A(z_i)$ and $y_i \to y$ imply that $y \in A(z)$. ¤

(3.47) Conditions (Zangwill [29]):

(i) $c(\cdot)$ is continuous on Ω.

(ii) $c(z') \leq c(z) \quad \forall z' \in A(z),\ z \in \Delta$.

(iii) $c(z') < c(z) \quad \forall z' \in A(z),\ z \in \Omega \sim \Delta$.

(iv) A is closed on $\Omega \sim \Delta$.

(v) For each $z \in \Omega \sim \Delta$, if $x_i \to z$ and $y_i \in A(x_i)$, then $\{y_i\}$ is compact. ¤

(3.48) Theorem: Conditions (3.47) imply Conditions (3.38).

Proof: Suppose Conditions (3.47) hold. Condition (3.47)(i) clearly implies (3.38)(i). Condition (3.47)(ii) is identical to (3.38)(ii) and hence (3.38)(i), (ii) hold. Let $z \in \Omega \sim \Delta$ and let $\{\xi_i\} \subset A(z)$ be such that $c(\xi_i) \to \sup\{c(\xi') | \xi' \in A(z)\}$. Condition (3.47)(v) implies that $\{\xi_i\}$ is compact and hence there exist a subsequence $\{\xi_i\}_{i \in K_1}$ and ξ^* such that $\xi_i \xrightarrow[K_1]{} \xi^*$. Condition (3.47)(iv) implies that $\xi^* \in A(z)$ and thus, because $c(\cdot)$ is continuous ((3.47)(i)),

$$(3.49) \quad c(\xi^*) = \max\{c(\xi') | \xi' \in A(z)\}.$$

Now (3.49) and (3.47)(iii) yield

$$(3.50) \quad c(z') \leq c(\xi^*) < c(z) \quad \forall z' \in A(z)$$

and so we have

$$(3.51)\quad c(z') \leq c(z) - (c(z)-c(\xi^*)) \triangleq c(z) - \gamma(z) \quad \forall z' \in A(z)$$

with $\gamma(z) > 0$ which establishes (3.38)(iii). To show that (3.38)(iv) holds, we assume the contrary and establish a contradiction. Let $z \in \Omega \sim \Delta$ and let $\xi^* \in A(z)$ be as above. Let $\{U_i\}$ be a sequence of neighborhoods of z satisfying (2.9)(ii). If (3.38)(iv) does not hold, we can find $\varepsilon > 0$, $z_i' \in U_i$ and $z_i'' \in A(z_i')$ such that

$$(3.52)\quad c(z_i'') \geq c(\xi^*) + \varepsilon \quad i = 0,1,\ldots .$$

By construction, $z_i' \to z$. Condition (3.47)(v) then implies that $\{z_i''\}$ is compact. Hence, there exist a subsequence $\{z_i''\}_{i\in K_2}$ and z_*'' such that $z_i'' \underset{K_2}{\to} z_*''$ and of course $z_i' \underset{K_2}{\to} z$. From (3.52) we conclude that

$$(3.53)\quad c(z_*'') \geq c(\xi^*) + \varepsilon$$

Since $A(\cdot)$ is closed ((3.47)(iv)), $z_*'' \in A(z)$ and (3.50) combines with (3.53) to yield

$$(3.54)\quad c(z_*'') \geq c(\xi^*) + \varepsilon \geq c(z_*'') + \varepsilon > c(z_*'')$$

and we have a contradiction. Consequently, (3.38)(iv) must hold and the proof is complete.

¤

4. Extension to the Time Varying Case

In this section, we modify Conditions (3.3) to apply to the case where the point-to-set map depends upon the iteration number, i. The other sufficient conditions can be extended in a similar fashion. These extensions are relatively straightforward. Therefore, we extend only Conditions (3.3) (the most general conditions) as an example of what can be done.

(4.1) Time Varying Algorithm Model: Let $A_i : \Omega \to 2^{\Omega}$ for $i = 0,1,\ldots$.

Step 0: Set $i = 0$. Choose any $z_o \in \Omega$.

Step 1: Choose any $z_{i+1} \in A_i(z_i)$.

Step 2: Set $i = i+1$ and go to step 1.

¤

(4.2) Conditions:

(i) $c(\cdot)$ is locally bounded from below on $\Omega \sim \Delta$.

(ii) There exists an integer $N_1 \geq 0$ such that $c(z') \leq c(z)$ $\forall z' \in A_i(z)$, $z \in \Omega$, $i \geq N_1$.

(iii) For each $z \in \Omega \sim \Delta$, if $\{x_i\} \subset \Omega$ is such that $x_i \to z$ and $c(x_i) \to c^*$, then there exists an integer $N_2 \geq N_1$ such that $c(y) < c^* \quad \forall y \in A_{N_2}(x_{N_2})$.

¤

(4.3) Theorem: If Conditions (4.2) hold, then Algorithm Model (4.1) is convergent (in the sense of Definition (2.5)).

Proof: Let z^* be an accumulation point of $\{z_i\}$, the sequence constructed by (4.1). We assume that $z^* \in \Omega \sim \Delta$ and establish a contradiction. There exists a subsequence $\{z_i\}_{i \in K}$ such that $z_i \underset{K}{\to} z^*$. Without loss of generality, we can also assume that $\{c(z_i)\}_{i \in K}$ is monotonically decreasing because of (4.2)(ii). If $z^* \in \Omega \sim \Delta$, (4.2)(i) implies that $\{c(z_i)\}_{i \in K}$ is bounded from below and hence $c(z_i) \underset{K}{\to} c^*$. Lemma (2.8) can be applied to obtain that $c(z_i) \to c^*$ and

(4.4) $c(z_i) \geq c^* \quad \forall i \geq N_1$.

But, if $z^* \in \Omega \sim \Delta$, (4.2)(iii) implies that

(4.5) $c(z_{N_2+1}) < c^*$

which contradicts (4.4). Thus, we must have $z^* \in \Delta$ and the proof is complete.

¤

(4.6) <u>Remark</u>: It is immediately obvious that Conditions (3.3) imply Conditions (4.2) when $A_i = A$ for $i = 0,1,2,\ldots$.

¤

Appendix A - Selected Counterexamples

The purpose of this Appendix is to show, by means of counterexamples, that certain of the implications not proved in Section 3 cannot, in fact, be proved. In the first set of counterexamples c is continuous, Δ is closed and A is single valued. Under these restrictions, the sets of sufficient conditions aggregate into four equivalence classes (see Figure 1). There are:

<u>Class I</u>: Conditions (3.3) and (3.31).

<u>Class II</u>: Conditions (3.8), (3.13), (3.18), (3.23) and (3.25).

<u>Class III</u>: Conditions (3.38) and (3.44).

<u>Class IV</u>: Condition (3.47).

It is immediately evident that IV implies III, III implies II, II implies I. We shall present counterexamples to show that the converse is false. The first two counterexamples will be constructed from the following optimization problem and algorithm

(A.1) <u>Problem</u>: $\min\{c(z) \mid z \in \mathbb{R}^1\}$ where $c: \mathbb{R}^1 \to \mathbb{R}^1$ is defined by

$$\text{(A.2)} \quad c(z) = \begin{cases} -z - 1 & \text{for } z \leq -1/2 \\ z^2 - 3/4 & \text{for } -1/2 < z < 1/2 \\ z - 1 & \text{for } 1/2 \leq z . \end{cases}$$

¤

(A.3) <u>Remark</u>: c is continously differentiable and

$$\text{(A.4)} \quad c'(z) = \begin{cases} -1 & \text{for } z \leq -1/2 \\ 2z & \text{for } -1/2 < z < 1/2 \\ 1 & \text{for } 1/2 \leq z . \end{cases}$$

¤

(A.5) Algorithm:

Data: $z_o \in \mathbb{R}^1$ arbitrary.

Parameters: $\alpha \in [0,1)$, $\beta \in (0,1)$, $\sigma > 0$.

Step 0: Set $i = 0$.

Step 1: If $c'(z_i) = 0$, set $z_{i+1} = z_i$. Else, set $z_{i+1} = z_i - \alpha\beta^{j(z_i)}$ $c'(z_i)$, where $j(z_i)$ is the smallest nonnegative integer satisfying

$$\text{(A.6)} \quad c(z_i - \sigma\beta^{j(z_i)} c'(z_i)) - c(z_i) < -\alpha\sigma\beta^{j(z_i)} \left|c'(z_i)\right|^2 .$$

Step 3: Set $i = i+1$ and go to Step 1.

¤

(A.7) Remark: When $\alpha \neq 0$, this algorithm is a version of the Armijo gradient method [32]. Our counterexamples are constructed by taking various values of the parameters α, β and σ.

¤

(A.8) Definition:

$$\Omega \triangleq \mathbb{R}^1$$

$$\Delta \triangleq \{z | c'(z) = 0\} = \{0\}$$

$$A(z) \triangleq a(z) \triangleq \begin{cases} z & \text{if } z \in \Delta \\ z - \sigma\beta^{j(z)} c'(z) & \text{if } z \in \Omega \sim \Delta. \end{cases}$$

¤

(A.9) Counterexample - I $\not\Rightarrow$ II: We take $\alpha = 0$, $\beta = 1/2$ and $\sigma = 1$ and show that Conditions (3.31) hold but Condition (3.8)(iii) does not hold. Straightforward calculations show that in this case

$$\text{(A.10)} \quad a(z) = \begin{cases} z + 1 & \text{for } z < -1/2 \\ 0 & \text{for } -1/2 \leq z \leq 1/2 \\ z - 1 & \text{for } 1/2 < z . \end{cases}$$

If $z > 0$ and $|z'-z| < 1/2$, (A.10) implies that $-z < a(z') < z$ and hence $c(a(z')) < c(z)$. Similarly, if $z < 0$ and $|z'-z| < 1/2$, (A.10) implies that $z < a(z') < -z$ and again $c(a(z')) < c(z)$. Thus, Condition (3.31)(iii) is satisfied at any $z \neq 0$, for the neighborhood $(z-1/2, z+1/2)$. It is also clear that (3.31)(i) and (3.31)(ii) hold.

Now we show that Condition (3.8)(iii) is not satisfied at $z = \frac{1}{2}$. Let $\{z_i\}$ be defined as $\frac{1}{2} + \frac{1}{i}$ for $i = 1,2,\ldots$. Clearly $z_i \to z$ and $c(z_i) \to c(z) = -1/2$. Also, by (A.10), $y_i = a(z_i) = z_i - 1 \in [-1/2, 1/2]$ for $i = 1,2,\ldots$. Hence $c(y_i) = (z_i-1)^2 - 3/4 \to (z-1)^2 - 3/4 = -1/2 = c(z)$. However, (3.8)(iii) requires that $\lim c(y_i) < c(z) = \lim c(z_i)$. Consequently, we have shown that Conditions (3.31) do not imply Conditions (3.8).

¤

(A.11) <u>Counterexample - II $\not\Rightarrow$ III</u>: We take $\alpha = 1/2$, $\beta = 2/3$ and $\sigma = 3/2$ and show that Conditions (3.23) are satisfied but that Condition (3.44)(v) does not hold. It is immediately evident that (3.23)(i) and (3.23)(ii) hold. Suppose $z \notin \Delta$, i.e. $c'(z) \neq 0$. For a nonnegative integer j, we have, by use of the Taylor Series expansion,

$$\text{(A.12)}\quad c(z-(\tfrac{3}{2})(\tfrac{2}{3})^j c'(z)) - c(z) + (\tfrac{1}{2})(\tfrac{2}{3})^j |c'(z)|^2$$
$$= -(\tfrac{3}{2})(\tfrac{2}{3})^j c'(z)c'(z) + (\tfrac{1}{2})(\tfrac{2}{3})^j |c'(z)|^2 + o((\tfrac{3}{2})(\tfrac{2}{3})^j c'(z))$$
$$= (\tfrac{2}{3})^j [-|c'(z)|^2 + \frac{o((\tfrac{3}{2})(\tfrac{2}{3})^j c'(z))}{(\tfrac{2}{3})^j}]$$

where $\frac{o(x)}{|x|} \to 0$ as $x \to 0$. Thus since $c'(z) \neq 0$, there exists an integer $\hat{j}$ such that the right hand side of (A.12) is strictly less than zero for $j = \hat{j}$. Hence

(A.13) $c(z-(\frac{3}{2})(\frac{2}{3})^{\hat{j}}c'(z)) - c(z) < -(\frac{1}{2})(\frac{2}{3})^{\hat{j}}|c'(z)|^2.$

Since c and c' are continuous, there exists an $\varepsilon > 0$ such that

(A.14) $c(z'-(\frac{3}{2})(\frac{2}{3})^{\hat{j}}c'(z')) - c(z') < -(\frac{1}{2})(\frac{2}{3})^{\hat{j}}|c'(z')|^2 < -(\frac{1}{2})(\frac{2}{3})^{\hat{j}}\frac{|c'(z)|^2}{2}$

$\forall |z'-z| < \varepsilon.$

We therefore can conclude that $j(z') \leq \hat{j}$ for all $|z'-z| < \varepsilon$ and hence

(A.15) $c(a(z')) - c(z') < -(\frac{1}{2})(\frac{2}{3})^{j(z')}\frac{|c'(z)|^2}{2} \leq -(\frac{1}{2})(\frac{2}{3})^{\hat{j}}\frac{|c'(z)|^2}{2}$

$\triangleq -\delta < 0 \quad \forall |z'-z| < \varepsilon.$

Thus, we have established condition (3.23)(iii).

To show that Condition (3.44)(v) does not hold, we show that $c(a(\cdot))$ is not upper semicontinuous at $z = -1$. A straightforward calculation shows that $j(z) = 1$, $a(z) = 0$ and hence $c(a(z)) = -3/4$. Now let $z' = z-\varepsilon$ for some $\varepsilon > 0$. It is easy to show that for $\varepsilon > 0$ sufficiently small, $j(z') = 0$, $a(z') = \frac{1}{2} - \varepsilon$ and hence $c(a(z'))$ $= (\frac{1}{2} - \varepsilon)^2 - 3/4 = \varepsilon^2 - \varepsilon - 1/2$. Thus, there exists an $\bar{\varepsilon} > 0$ such that $c(a(z+\varepsilon)) \geq -5/8 > -3/4 = c(a(z))$ for all $\varepsilon \in (0,\bar{\varepsilon}]$ and $c(a(\cdot))$ is not upper semicontinuous at $z = -1$.

¤

(A.16) <u>Counterexample - III $\not\Rightarrow$ IV</u>: For this counterexample, we apply Algorithm (A.5) to a new function c where

(A.17) $c(z) = \begin{cases} z & \text{for } 1 < z \\ z^2 & \text{for } -1 \leq z \leq 1 \\ -z & \text{for } z < -1. \end{cases}$

This c is not differentiable at 1 and -1 so at these points we take $c'(1) = \lim_{z\uparrow 1} c'(z) = 2$ and $c'(-1) = \lim_{z\downarrow -1} c'(z) = -2$ in Algorithm (A.5).

We also select $\sigma = \frac{3}{4}$, $\alpha = \frac{1}{8}$ and β to any positive number. Under these conditions, it is easy to show that $j(z) = 0$ for all z and thus

$$\text{(A.18)} \quad a(z) = \begin{cases} x - 3/4 & 1 < z \\ -z/2 & -1 \leq z \leq 1 \\ z + 3/4 & z < -1 \end{cases}$$

We shall now show that Conditions (3.44) hold but that Condition (3.47)(iv) does not hold. It is clear that (3.44)(i) - (3.44)(iv) hold. Also, it is obvious that (3.44)(v) holds for all z, except, possibly at $z = +1$, and $z = -1$. We show that $c(a(\cdot))$ is uppersemicontinuous at $z = 1$ and thus by symmetry, (3.44)(v) holds for all z. First $c(a(1)) = c(-\frac{1}{2}) = (-\frac{1}{2})^2 = \frac{1}{4}$. Now consider $\varepsilon \in (0,3/4]$. Then $a(1+\varepsilon) \in (\frac{1}{4},1]$. Thus, $c(a(1+\varepsilon)) = [a(1+\varepsilon)]^2 = (1+\varepsilon-\frac{3}{4})^2 = (1+\varepsilon)^2 - \frac{3}{2}(1+\varepsilon) + \frac{9}{16}$. On the other hand, $a(1-\varepsilon) \in (-\frac{1}{2}, -\frac{1}{8}]$ and so $c(a(1-\varepsilon)) = (\frac{\varepsilon-1}{2})^2$. Consequently,

$$\text{(A.19)} \quad \lim_{\varepsilon \downarrow 0} c(a(1+\varepsilon)) = \frac{1}{16} < \frac{1}{4} = c(a(1)) = \lim_{\varepsilon \downarrow 0} c(a(1-\varepsilon)).$$

We conclude from (A.19) that $c(a(\cdot))$ is upper semicontinuous at 1 and thus everywhere. Therefore, we have shown that Conditions (3.44) hold. On the other hand, $\lim_{\varepsilon \downarrow 0} a(1+\varepsilon) = \frac{1}{4}$ but $a(1) = \frac{1}{2}$ so that $a(\cdot)$ is not closed (i.e. not continuous at 1 and (3.47)(iv) does not hold.[†] ⌑

[†] It is interesting to note that a counterexample can be constructed using a continuously differentiable function, c. In particular, the function and algorithm used in Counterexample (A.11) can be used with $\alpha = \frac{1}{3}$, $\beta = \frac{1}{6}$ and $\sigma = \frac{3}{2}$. After a substantial amount of calculation, it can be shown that the resulting map $a(\cdot)$ is discontinuous at $z = \pm 1$ because $j(z)$ is discontinuous at these points.

In our last counterexample, we show that Conditions (3.3) do not imply Conditions (3.31). As can be seen from Figure 1, we shall need a function $c(\cdot)$ that is not continuous. The following lower semicontinuous function will suffice.

$$\text{(A.20)} \quad c(z) = \begin{cases} z+1 & \text{for } 1 < z \\ z^2 & \text{for } z \leq 1. \end{cases}$$

(A.21) <u>Counterexample - (3.3) $\not\Rightarrow$ (3.31)</u>: We apply Algorithm (3.5) to c as defined by (A.20) where we take $c'(1) = \lim_{z \uparrow 1} c'(z) = 2$. We also select $\sigma = \frac{9}{4}$, $\alpha = \frac{1}{9}$ and $\beta = \frac{2}{9}$. After some computation, it can be shown that

$$\text{(A.22)} \quad j(z) = \begin{cases} 0 & z > 1 \\ 1 & z \leq 1 \end{cases}$$

and hence

$$\text{(A.23)} \quad a(z) = \begin{cases} z - \frac{9}{4} & z > 1 \\ 0 & z \leq 1. \end{cases}$$

We now show that Conditions (3.3) hold but Condition (3.31)(iii) does not hold. Clearly (3.3)(i) and (3.3)(ii) hold. For any $z \neq 1$ and $z \neq 0$, $c(a(\cdot))$ is continuous and $c(a(z)) < f(z)$. Thus there exists $\varepsilon > 0$ such that $|z'-z| < \varepsilon$ implies that $c(a(z')) < c(z)$. Since z is a point of continuity of c, we conclude that (3.3)(iii) holds for all $z \notin \Delta$ except possibly $z = 1$. Now consider $z_i = 1 + \frac{1}{i}$ and $\tilde{z}_i = 1 - \frac{1}{i}$ for $i = 1, \ldots$. Then

$$\text{(A.24)} \quad z_i \downarrow 1, \; \tilde{z}_i \uparrow 1$$

$$\text{(A.25)} \quad c(z_i) = 1 + \frac{1}{i} + 1 \to 2 \triangleq c^*$$

$$\text{(A.26)} \quad c(\tilde{z}_i) = (1 - \frac{1}{i})^2 \to 1 \triangleq \tilde{c}^*$$

$$\text{(A.27)} \quad c(a(z_i)) = (1 + \frac{1}{i} - \frac{9}{4})^2 \to \frac{25}{16}$$

$$\text{(A.28)} \quad c(a(z_i)) = c(0) = 0.$$

Thus, there exists an integer N such that

$$\text{(A.29)} \quad c(a(z_N)) < 2 = c^*$$

$$\text{(A.30)} \quad c(a(\tilde{z}_N)) = 0 < 1 = \tilde{c}^*$$

and (3.3)(iii) also holds at $z = 1$.

On the other hand, Condition (3.31)(iii) does not hold at $z = 1$. To see this, consider any $\varepsilon \in (0,1)$. Then, $a(1+\varepsilon) = 1 + \varepsilon - \frac{9}{4} = \varepsilon - \frac{5}{4} \in (-\frac{5}{4}, -\frac{1}{4})$. Therefore,

$$\text{(A.31)} \quad c(a(1+\varepsilon)) = (\varepsilon - \frac{5}{4})^2 = \varepsilon^2 - \frac{5}{2}\varepsilon + \frac{25}{16} \xrightarrow[\varepsilon\downarrow 0]{} \frac{25}{16}$$

Thus, there exists an $\bar{\varepsilon} > 0$ such that

$$\text{(A.32)} \quad c(a(1+\varepsilon)) > 1 = c(1) \quad \forall \varepsilon \in (0,\bar{\varepsilon}].$$

Consequently, Condition (3.31)(iii) does not hold at $z = 1$. ⌑

References

[1] J. W. Daniel, "Convergent step sizes for gradient-like feasible direction algorithms for constrained optimization," in Nonlinear Programming (J. B. Rosen, O. L. Mangasarian and K. Ritter, eds.), Academic Press, New York, 1970, 245-274.

[2] J. Dubois, "Theorems of convergence for improved nonlinear programming algorithms," Operations Research, 21 (1), 1973, 328-332.

[3] B. C. Eaves and W. I. Zangwill, "Generalized cutting plane algorithms, SIAM J. Control, 9 (4), 1971, 529-542.

[4] R. M. Elkin, "Convergence theorems for Gauss-Seidel and other minimization algorithms," Computer Science Center Tech. Report 68-59, University of Maryland, College Park, Maryland, 1968.

[5] W. W. Hogan, "Point-to-set maps in mathematical programming," SIAM Review, 15 (3), 1973, 591-603.

[6] W. W. Hogan, "Applications of general convergence theory for outer approximation algorithms," Math. Prog., 5, 1973, 151-168.

[7] P. Huard, "Optimization algorithms and point-to-set maps," Math. Prog., 8 (3), 1975, 308-331.

[8] R. Klessig, "A general theory of convergence for constrained optimization algorithms that use antizigzagging provisions," SIAM J. Control, 12 (4), 1974, 598-608.

[9] R. Klessig and E. Polak, "An adaptive precision gradient method for optimal control," SIAM J. Control, 11 (1), 1973, 80-93.

[10] D. G. Luenberger, Introduction to Linear and Nonlinear Programming, Addison-Wesley, Reading, Mass., 1973.

[11] M. L. Lenard, "Practical convergence conditions for unconstrained optimization," Math. Prog., 4, 1973, 309-323.

[12] G. G. L. Meyer and E. Polak, "Abstract models for the synthesis of optimization algorithms," SIAM J. Control, 9, 1971, 547-560.

[13] G. G. L. Meyer, "Convergence conditions for a type of algorithm model," Technical Report 75-14, The Johns Hopkins University, Baltimore, Maryland, 1975.

[14] G. G. L. Meyer, "Algorithm model for penalty-type iterative procedures," J. Comp. and Sys. Science, 9, 1974, 20-30.

[15] R. R. Meyer, "Sufficient conditions for the convergence of monotonic mathematical programming algorithms," J. Comp. and Sys. Science, 12, 1976, 108-121.

[16] R. R. Meyer, "A comparison of the forcing functions and point-to-set mapping approaches to convergence analysis," SIAM J. Control, 15 (4), 1977, 699-715.

[17] R. R. Meyer, "A convergence theory for a class of anti-jamming strategies," Technical Report, 1481, Math. Research Center, University of Wisconsin, Madison, Wisconsin, 1974.

[18] H. Mukai and E. Polak, "On the use of approximations in algorithms for optimization problems with equality and inequality constraints," Memorandum No. ERL-M489, Electronics Research Laboratory, University of California, Berkeley, California, 1974.

[19] J. M. Ortega and W. C. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, New York, 1970.

[20] J. M. Ortega and W. C. Rheinboldt, "A general convergence result for unconstrained minimization methods," SIAM J. Numer. Anal., 9 (1), 1972, 40-43.

[21] E. Polak, "On the convergence of optimization algorithms," Rev. Fr. Inform. Rech. Operation, (16-R1), 1969, 17-34.

[21a] E. Polak, Computational Methods in Optimization: A Unified Approach, Academic Press, New York, 1971.

[22] E. Polak, "On the implementation of conceptual algorithms," in Nonlinear Programming (J. B. Rosen, O. L. Mangasarian and K. Ritter, eds.), Academic Press, New York, 1970, 275-291.

[23] E. Polak, R. W. H. Sargent and D. J. Sebastian, "On the convergence of sequential minimization algorithms," JOTA, 14, 1974, 439-442.

[24] B. T. Polyak, "Gradient methods for the minimization of functionals," U.S.S.R. Comp. Math. and Math. Phys., 3, 1963, 864-878.

[25] S. W. Rauch, "A convergence theory for a class of nonlinear programming problems," SIAM J. Numer. Anal., 10 (1), 1973, 207-228.

[26] S. Tishyadhigama, "General convergence theorems: Their relationships and applications," Ph.D. Thesis, Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, California, 1977.

[27] D. M. Topkis and A. F. Veinotte, "On the convergence of some feasible direction algorithms for nonlinear programming," SIAM J. Control, 5 (2), 1976, 268-279.

[28] P. Wolfe, "Convergence conditions for ascent methods," SIAM Review, 11 (2), 1969, 226-235.

[29] W. I. Zangwill, Nonlinear Programming: A Unified Approach, Prentice-Hall, Englewood Cliffs, New Jersey, 1969.

[30] W. I. Zangwill, "Convergence conditions for nonlinear programming algorithms," Management Science, 16 (1), 1969, 1-13.

[31] J. L. Kelley, General Topology, Van Nostrand, Princeton, New Jersey, 1955.

[32] L. Armijo, "Minimization of functions having continuous partial derivatives," Pacific J. Math., 16, 1966, 1-3.

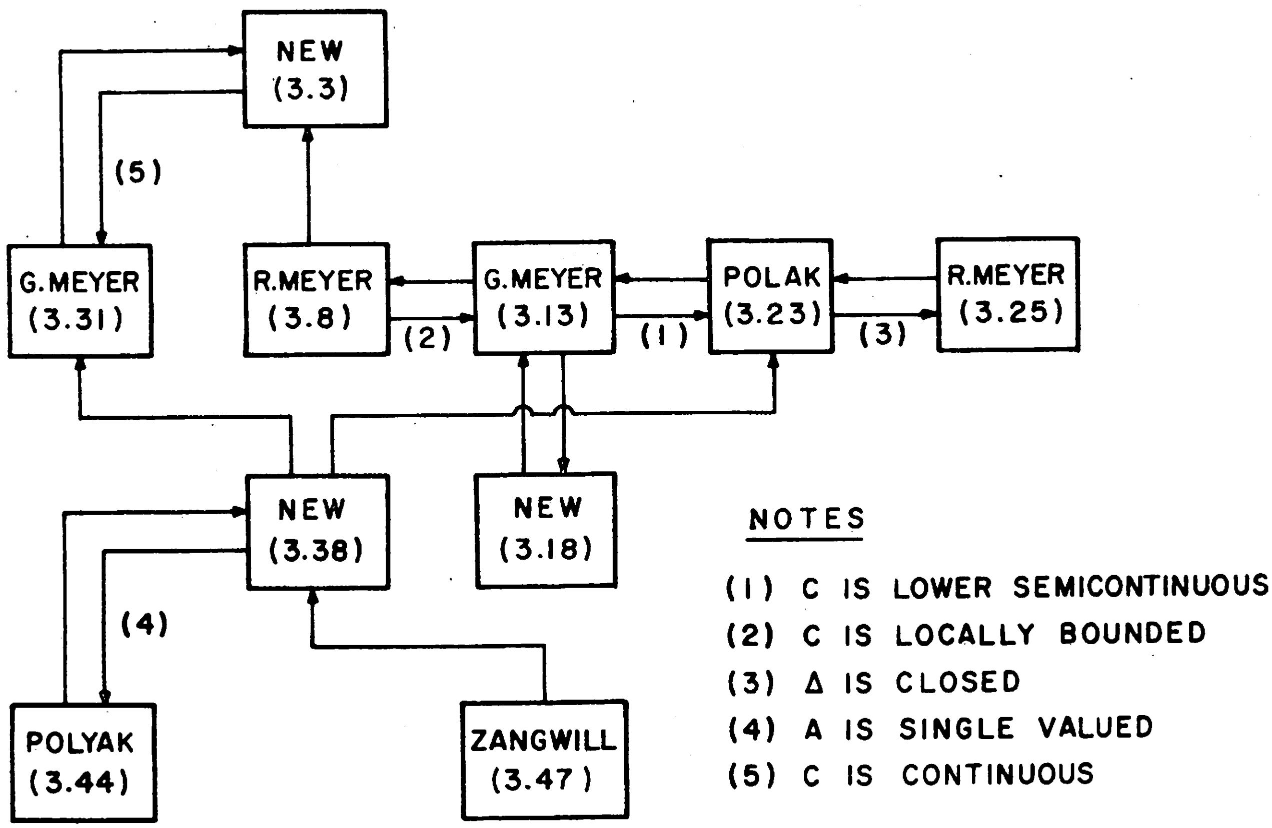

Figure 1 Results of Section 3