A FINITE ALGORITHM FOR SOLVING INFINITE DIMENSIONAL OPTIMIZATION PROBLEMS

IRWIN E. SCHOCHETMAN AND ROBERT L. SMITH

Department of Mathematical Sciences
Oakland University
Rochester, Michigan 48309
and
Department of Industrial and Operations Engineering
The University of Michigan
Ann Arbor, Michigan 48109

Technical Report 97-12
September 29, 1997

A FINITE ALGORITHM FOR SOLVING INFINITE DIMENSIONAL OPTIMIZATION PROBLEMS

IRWIN E. SCHOCHETMAN AND ROBERT L. SMITH

Department of Mathematical Sciences
Oakland University
Rochester, Michigan 48309
and
Department of Industrial and Operations Engineering
The University of Michigan
Ann Arbor, Michigan 48109

September 29, 1997

ABSTRACT. We consider the general optimization problem (P) of selecting a continuous function x over a σ-compact Hausdorff space T to a metric space A, from a feasible region X of such functions, so as to minimize a functional c on X. We require that X consist of a closed equicontinuous family of functions lying in the product (over T) of compact subsets Y_t of A. (An important special case is the optimal control problem of finding a continuous time control function x that minimizes its associated discounted cost $c(x)$ over the infinite horizon.) Relative to the uniform-on-compacta topology on the function space $C(T, A)$ of continuous functions from T to A, the feasible region X is compact. Thus optimal solutions x^* to (P) exist under the assumption that c is continuous. We wish to approximate such an x^* by optimal solutions to a net $\{P_i\}_i \in I$ of approximating problems of the form $\min_{x \in X_i} c_i(x)$ for each $i \in I$, where (1) the net of sets $\{X_i\}_I$ converges to X in the sense of Kuratowski and (2) the net $\{c_i\}_I$ of functions converges to c uniformly on X. We show that for large i, *any* optimal solution x_i^* to the approximating problem (P_i) arbitrarily well approximates *some* optimal solution x^* to (P). It follows that if (P) is well-posed, i.e., $\limsup X_i^*$ is a singleton $\{x^*\}$, then *any* net $\{x_i^*\}_I$ of (P_i)-optimal solutions converges in $C(T, A)$ to x^*. For this case, we construct a *finite* algorithm with the following property: given any prespecified error δ and any compact subset Q of T, our algorithm computes an i in I and an associated x_i^* in X_i^* which is within δ of x^* on Q. We illustrate the theory and algorithm with a problem in continuous time production planning over an infinite horizon.

§1 Introduction

Consider the abstract optimization problem

$$\min_{x \in X} c(x) \tag{P}$$

1991 *Mathematics Subject Classification.* Primary 90C20, Secondary 49A99.

Key words and phrases. Continuous time optimization, optimal control, infinite horizon optimization, production planning.

This work was supported in part by the National Science Foundation under Grants DDM–9214894 and DMI–9713723. The first author was partially supported by an Oakland University Research Fellowship.

Typeset by $\mathcal{A}_{\mathcal{M}}\mathcal{S}$-TEX

where the feasible region X is a non-empty compact subset of a function space Y and the objective function c is a real-valued continuous function over Y. The Weierstrass Theorem assures us of the existence of an optimal solution $x^* \in X$ that attains the minimum value c^* of c over X. However, optimization problems in this class are difficult to numerically solve in general, since there is seldom a concrete representation for solutions in Y. In this paper, we explore general methods for approximating a solution to (P) via solutions to a net $(P_i), i \in I$, of simpler approximating problems, where (P_i) is given by

$$\min_{x \in X_i} c_i(x) \tag{P_i}$$

for $i \in I$, c_i is continuous on $Y, \lim X_i = X$ (Kuratowski) and $c_i \to c$ uniformly on Y. These approximating problems are typically finite dimensional (or special in other ways that render them more easily solvable than the original problem (P)). For example, we may choose to approximate an optimal solution to a continuous time infinite horizon optimization problem (P) by optimal solutions to discrete time finite horizon versions $(P_i), i \in \mathbb{N}$, where the time periods decrease to zero and the horizon increases to infinity as $i \to \infty$ (see e.g. Schochetman and Smith [1989, 1992a, 1997], Bes and Sethi [1988], Bean and Smith [1984, 1993]).

The emphasis in this paper is on the case where Y is an infinite dimensional function space. For example, problems in infinite horizon optimal control seek a control function x from a space Y of continuous functions on $\mathbb{R}^+$ that minimizes its associated infinite horizon discounted cost. In this case, the feasible region X is specified through a differential or integral equation that relates system state evolution to the control policy employed (see e.g. Carlson, Haurie and Leizarowitz [1987], Luenberger [1969]). In fact, the special case of a problem in infinite horizon production control is considered here in detail in section 5. More generally, best approximation problems also fall within the framework of (P), where $c(x)$ is a measure of the error of x from a fixed target function x_0 (see e.g. Schochetman and Smith [1991, 1992b], Dontchev and Zolezzi [1993]).

Since the objective functions and feasible regions of the approximating problems $(P_i)_{i \in I}$ are also defined over the common space Y, it is important to endow Y with a topology that is fine enough to allow for the error in solution approximation to be suitably small, but at the same time, coarse enough for desirable properties like compactness of X to hold. Toward this end, we begin by embedding Y within the set of all functions from a σ-compact Hausdorff space T to a metric space A. By requiring X to be a non-empty closed subset of equicontinuous functions from the product (over T) of compact subsets in A, the pointwise convergence and uniform convergence on compacta topologies agree on X. If follows by Tychonoff that X remains compact in the stronger topology of uniform convergence on compacta. It is this stronger topology that appropriately measures solution error; roughly speaking, two solutions over T are "close" when their difference is uniformly small over a compact subset Q of T. For example, if $T = \mathbb{N}$, and is discrete, then this reduces to actual agreement on finite subsets of $\mathbb{N}$ and in particular on $\{1, \ldots, n\}$. Moreover since T is σ-compact, it is the countable union of such subsets, and the near agreement can thus be required over nearly all of T. In fact, there exist compact $Q_i \to T$ (Kuratowski), as $i \to \infty$, where $Q_i = \cup_{k=1}^{i} Q'_k$, Q'_k compact for all k.

We require only two properties of the approximating problems (P_i) in their relation to (P): 1) their feasible regions $X_i \to X$ (Kuratowski) and 2) their objective functions $c_i \to c$ (uniformly). Requirement 1) is a significant relaxation over assumptions commonly made in the literature. For example, in Schochetman and Smith [1989], it is required that $X_i = X$, for all i, while in Schochetman and Smith [1992] and Semple [1996], it is assumed that $X_{i+1} \subseteq X_i$, all i, and $X = \cap_{i=1}^{\infty} X_i$, so that $X_i \downarrow X$ in all all cases. Also, the action functions in these papers are defined within a discrete time framework where $T = \{1, 2, \ldots\}$.

In section 2, the class of optimization problems considered is formally defined, as well as their associated nets of approximating problems. We prove optimal value convergence, i.e., that the net of optimal values $\{c_i^*\}$ to the approximating problems converges to the optimal value c^* of the original optimization problem.

We also establish the fundamental result that solutions to the approximating problems (P_i) become (for large i) arbitrarily close to optimal solutions to the original problem (P). Section 3 turns to establishing conditions under which optimal policy convergence takes place, i.e., conditions under which optimal solutions of approximating problems converge to an optimal solution of the original problem. Existence of a unique accumulation point of approximating optima is established as a sufficient condition for this to take place. Under this well-posedness condition, in section 4, an algorithm is provided, together with stopping rule, that is guaranteed to finitely compute an approximating solution within any prespecified error from the optimal solution. Finally, in section 5, we apply the theory and algorithms developed in the preceding section to a problem in continuous-time infinite horizon production control.

§2 Problem Formulation and Value Convergence

Let T be a σ-compact Hausdorff space (Dugundji [1966, p. 240]) and A a metric space with metric d. The space T is the space of the *action index* (e.g., time) and A is the space of all possible *actions*. Let A^T denote the set of all functions from T into A. An element y of A^T will be called an *action strategy* in A over T.

Next let $C(T, A)$ denote the set of continuous functions from T into A. Although there are several natural Hausdorff topologies for $C(T, A)$, there is one which is of particular interest to us, namely the topology of uniform convergence on compact sets or more briefly, the uniform-on-compacta topology (Kelly [1955]. Moreover, since T is σ-compact, it follows (Dugundji [1966 , p.272]) that this topological space $C(T, A)$ is metrizable and hence, first countable. We next observe that the uniform-on-compacta topology is a *jointly continuous* topology on $C(T, A)$; this is not true in general for the topology of pointwise convergence (Kelley [1955, p. 223]). It is primarily for this reason that we adopt the topology of uniform convergence on compacta on $C(T, A)$.

Lemma 2.1. *The canonical mapping $(y, t) \to y(t)$ of $C(T, A) \times T$ into A is continuous. If T is discrete, then it is also continuous relative to the topology of pointwise convergence on $C(T, A)$.*

Proof. This follows from the definitions of the relevant topologies, together with the fact that T is locally compact. □

Let E be a non-empty, *equicontinuous* (Kelley [1955, p.232]) family of functions in $C(T, A)$, i.e., for each $t_0 \in T$ and $\epsilon > 0$, there exists a neighborhood U of t_0 in T such that

$$d(y(t), y(t_0)) < \epsilon, \quad \forall t \in U, \quad \forall y \in E.$$

For such a subspace E it follows (Kelley [1955, p.232]) that the relative topologies of uniform convergence on compacta and pointwise convergence are equal. Note that a sufficient condition for E to be equicontinuous is for E to be compact in $C(T, A)$ (Kelley [1955, p.233]); if T is discrete, then it is sufficient for E to be merely pointwise-compact (Lemma 2.1).

Since E is equicontinuous, its pointwise-closure in A^T is also equicontinuous (Kelley [1955, p.232]), as is its relative pointwise-closure in $C(T, A)$. Thus, there is no loss of generality in also assuming that E is pointwise-closed in $C(T, A)$. Consequently, E is closed in $C(T, A)$ as well.

In general, given $t \in T$, it is unlikely that all actions in A will be feasible for t. Thus, we will let Y_t denote the space of actions in A which are feasible at action index t, $\forall t \in T$. We assume that Y_t is a non-empty, compact subset of A, $\forall t \in T$. Let

$$\begin{aligned} Y &= E \cap \prod_{t \in T} Y_t \\ &= \{y \in E : y(t) \in Y_t, \forall t \in T\}, \end{aligned}$$

so that $Y \subseteq E \subseteq C(T,A) \subseteq A^T$ and $Y \subseteq \prod_{t\in T} Y_t \subseteq A^T$. Note that $\prod_{t\in T} Y_t$ is not contained in $C(T,A)$ in general, unless T is discrete. (In fact, if T is discrete, we could choose $E = \prod_T Y_t = Y$.) Thus, our choice of Y is necessitated by the fact that the decision index t need not be discrete. We assume $Y \neq \emptyset$. The space Y will play the role of the space of all possible strategies (feasible or not) over T with values in the appropriate decision spaces Y_t, $\forall t \in T$. Note that Y is an equicontinuous family also, so that the restrictions to Y of the uniform-on-compacta and pointwise convergence topologies agree. We will simply refer to these restrictions as *the* topology of Y. Finally, note that Y being a subspace of $C(T,A)$ is also first countable.

The next result explains why we require E to be equicontinuous.

Lemma 2.2. *The space Y is a non-empty compact Hausdorff space.*

Proof. The space $\prod_{t\in T} Y_t$ is pointwise-compact by the Tychonoff Theorem (Kelley [1955]), and contained in A^T which is Hausdorff relative to pointwise convergence. Hence, $\prod_{t\in T} Y_t$ is pointwise-closed in A^T (Kelley [1955]). Thus, Y is pointwise-closed in A^T and pointwise-closed in $C(T,A)$. Consequently, Y is closed in $C(T,A)$. Thus, the result follows from Ascoli's Theorem (Kelley [1955, p.233]). □

Example 2.3. (Discrete-Time, Discrete-Action Space) Let T be the positive integers $\mathbb{N}$, $Y_t = \{0,1,\ldots,m_t\}, \forall t = 1,2,\ldots$, and $A = \{0,1,\ldots,m\}$, where we assume

$$0 \leq m_t \leq m, \quad \forall t = 1,2,\ldots,$$

and

$$d(x,y) = |x-y|, \quad \forall x,y = 0,1,\ldots,m.$$

Then $E = \prod_{t=1}^{\infty} Y_t$ is pointwise-compact and hence, equicontinuous (Kelley [1955, p.233]) and pointwise-closed. In this case, $Y = \prod_{t=1}^{\infty} Y_t$.

Example 2.4. (Discrete-Time, Continuous-Action Space) Let $T = \mathbb{N}$ and let Y_t be a compact subset of the Euclidean space $\mathbb{R}^{n_t}$ of dimension n_t, $\forall t = 1,2,\ldots$. Define A to be the set of infinite sequences $x = (x_t)$ of real vectors in the $\mathbb{R}^{n_t}$ having the property that $x_t = 0$, for all but finitely many $t = 1,2,\ldots$. Then A is a metric space with metric given by

$$d(x,y) = (\sum_{t=1}^{\infty} \|x_t - y_t\|^2)^{1/2}, \quad \forall x,y \in A.$$

Note that the canonical restriction of d to $\mathbb{R}^{n_t}$ yields the usual Euclidean metric, $\forall t = 1,2,\ldots$. Once again, $E = \prod_{t=1}^{\infty} Y_t$ is pointwise-compact, equicontinuous (Kelley [1955, p.233]) and pointwise-closed; $Y = \prod_{t=1}^{\infty} Y_t$ also.

Example 2.5. (Continuous-Time, Continuous-Action Space) Now let $T = \mathbb{R}^+$ (non-negative reals) and $A = \mathbb{R}$, so that A^T is the set $\mathbb{R}^{\mathbb{R}^+}$ of real-valued functions on $\mathbb{R}^+$. Fix $M > 0$ and let

$$E = \{y \in \mathbb{R}^{\mathbb{R}^+} : 0 \leq y(t) - y(s) \leq M(t-s), \quad \forall 0 \leq s \leq t\}.$$

Note that the constant functions are contained in each such E. Thus, E is a non-empty equicontinuous subset of $C(\mathbb{R}^+,\mathbb{R})$. Moreover, it is pointwise-closed in $C(\mathbb{R}^+,\mathbb{R})$, and hence, uniform-on-compacta closed as well. Also let Y_t denote the closed interval $[0, Mt]$, $\forall t \geq 0$. Note that the zero function is in both E and $\prod_{t\geq 0} Y_t$, the non-zero constant functions are in E, but not in $\prod_{t\geq 0} Y_t$, and there exist non-continuous functions in $\prod_{t\geq 0} Y_t$, which are not in E. The set $\prod_{t\geq 0} Y_t$ is pointwise-closed in $\mathbb{R}^{\mathbb{R}^+}$ and thus, also closed. Therefore, $Y = E \cap \prod_{t\geq 0} Y_t$ is also closed in $C(\mathbb{R}^+, \mathbb{R})$, and hence, compact by Ascoli's Theorem (Kelley [1955, p.233]). Also, note that $y(0) = 0$, for each $y \in Y$; in fact, $y \in Y$ if and only if $y \in E$ and $y(0) = 0$. More generally, let

$$F = \{y \in \mathbb{R}^{\mathbb{R}^+} : |y(t) - y(s)| \leq M|t - s|, \quad \forall s, t \geq 0\},$$

i.e., F is the set of *Lipschitz continuous functions* with uniform Lipschitz constant M. Accordingly, also let $Y_t = [-Mt, Mt]$, $\forall t \geq 0$. Then the previous claims are true for F and $Y = F \cap \prod_{t\geq 0} Y_t$ as well.

Note that the remaining case where T is continuous-time while A is a discrete-action space is not considered, since the only continuous functions from a connected space to a discrete space are the constant functions, a trivial family.

Returning to the general discussion preceding Example 2.3, recall that Y is a compact Hausdorff space. We will let $\mathcal{K}(Y)$ denote the set of all closed (hence, compact), non-empty subsets of Y. If we assume that $\mathcal{K}(Y)$ is equipped with the relativized *Vietoris* topology (Klein and Thompson [1984, p.8]), then $\mathcal{K}(Y)$ is also a compact Hausdorff space (Klein and Thompson [1984, pp.15-17]). Moreover, the convergence in $\mathcal{K}(Y)$ underlying this topology is precisely *Kuratowski convergence* (Klein and Thompson [1984, p. 34]) which we describe next. Let $\{S_i\}_I$ be a net of subsets of Y (with I directed by $\preceq$) and $y \in Y$. Define:

(i) y is a *limit* point of the net $\{S_i\}_I$ if, for every neighborhood U of y in Y, there exists $i_U \in I$ such that $S_i \cap U \neq \emptyset$, for all $i \in I$ for which $i_U \preceq i$.

(ii) y is a *cluster* point of the net $\{S_i\}_I$ if, for every neighborhood U of y in Y, and every $i \in I$, there exists $j_i \in I$ such that $i \preceq j_i$ and $S_{j_i} \cap U \neq \emptyset$.

Then let $\liminf_I S_i$ (resp. $\limsup_I S_i$) denote the set of limit (resp. cluster) points of the $\{S_i\}_I$. If $S \subseteq Y$ and $S = \liminf_I S_i = \limsup_I S_i$, we write $\lim_I S_i = S$. In general, $\liminf_I S_i$ and $\limsup_I S_i$ are closed subsets of Y which may be empty, and which satisfy $\liminf_I S_i \subseteq \limsup_I S_i$. Thus, $\lim_I S_i = S$ if and only if $\limsup_I S_i \subseteq S$ and $S \subseteq \liminf_I S_i$.

We assume that our feasible region X is a given closed, non-empty subset of Y, i.e., $X \in \mathcal{K}(Y)$. We also assume we are given a real-valued, continuous cost function c defined on Y. Our optimization problem $(\mathcal{P})$ is then defined as follows:

$$\min_{x\in X} c(x). \qquad (\mathcal{P})$$

Since X is compact and c is continuous, it follows that the minimum is attained. We will denote the set of optimal solutions to $(\mathcal{P})$ by X^* and the optimal cost by c^*. Of course, X^* is a non-empty, closed subset of X, i.e., $X^* \in \mathcal{K}(Y)$.

Our primary objective in this paper is to approximate the optimal solutions of $(\mathcal{P})$ by optimal solutions of problems which approximate $(\mathcal{P})$. To this end, let $\{X_i\}_I$ be any net in $\mathcal{K}(Y)$ such that $\lim_I X_i = X$. The directed set I (directed by $\preceq$) is the *approximation index* set.

Example 2.6. (Discrete, Partially-Ordered Approximation Index) Suppose that X can be expressed as a countable intersection $\cap_{n=1}^{\infty} K_n$, where $K_n \in \mathcal{K}(Y)$, $\forall n = 1, 2, \ldots$. For example, if the notation is as in Example 2.4, and X is the region satisfying countably infinitely many nonlinear constraints, then K_n could be the set of solutions in Y which satisfy the n-th constraint, $\forall n = 1, 2, \ldots$. Let I denote the collection of all finite subsets of $\mathbb{N}$, so that I is a countable set which is directed by $\subseteq$. For each $i \in I$, let $X_i = \cap_{n\in i} K_n$.

Then $\{X_i\}_I$ is a (countable) net in $\mathcal{K}(Y)$ where $i \subseteq j$ implies $X_j \subseteq X_i$. Moreover, by Klein and Thompson [1984, p.28],

$$\lim_I X_i = \cap_{i \in I} X_i = X.$$

Example 2.7. (Continuous-Time Approximation Index) Let the notation be as in Example 2.5, with $(I, \preceq) = (\mathbb{R}^+, \leq)$ also. Suppose we are given a function $D : \mathbb{R}^+ \to \mathbb{R}^+$. Define

$$X = \{y \in Y : y(t) \geq D(t),\ \forall t \geq 0\},$$

and suppose it is non-empty. Then X is pointwise-closed in Y, i.e., $X \in \mathcal{K}(Y)$. For each $t \geq 0$, define

$$X_t = \{y \in Y : y(s) \geq D(s),\ \forall 0 \leq s \leq t\},$$

so that $\{X_t\}_{t \geq 0}$ is a net in $\mathcal{K}(Y)$ for which $s \leq t$ implies $X_t \subseteq X_s$, and

$$\lim_{t \to \infty} X_t = \bigcap_{t \geq 0} X_t = X.$$

We now return to our general situation. For each $i \in I$, let c_i be a continuous, real-valued function on Y. We assume that the net $\{c_i\}_I$ converges *uniformly* to c on Y, i.e., given $\delta > 0$, there exists $i_\delta \in I$ such that

$$|c(y) - c_i(y)| < \delta, \quad \forall y \in Y,$$

for all $i \in I$ such that $i_\delta \preceq i$. For each $i \in I$, we define the i-th approximating problem $(\mathcal{P}_i)$ as follows:

$$\min_{x \in X_i} c_i(x). \qquad (\mathcal{P}_i)$$

(In the particular case where $I = \mathbb{N}$, we have that the sequence of problems $(\mathcal{P}_i)$ converges in the sense of Fiacco [1974] to the problem $(\mathcal{P})$.) The optimal solution set X_i^* to $(\mathcal{P}_i)$ is then a non-empty, closed subset of Y. Thus, $X_i^* \in \mathcal{K}(Y)$, for all $i \in I$. We will denote the optimal objective value to $(\mathcal{P}_i)$ by c_i^*, $\forall i \in I$.

Theorem 2.8. *(Value Convergence) The net $\{c_i^*\}_I$ of optimal values to the $(P_i), i \in I$, converges to the optimal value c^* of (P), i.e., $\lim_I c_i^* = c^*$.*

Proof. We first show that $\liminf_I c_i^* \geq c^*$. By definition of $\liminf_I c_i^*$, there exists a subnet $\{c_j^*\}_J$ of $\{c_i^*\}_I$ such that $\lim_J c_j^* = \liminf_I c_i^*$. But $c_j^* = c_j(x_j^*)$, for some $x_j^* \in X_j$, $\forall j \in J$. Hence, $\{x_j^*\}_J$ is a net in compact Y. Thus, there exists a subnet $\{x_k^*\}_K$ of $\{x_j^*\}_J$, with corresponding subnet $\{c_k^*\}_K$ of $\{c_j^*\}_J$, and $x \in Y$ such that $\lim_K x_k^* = x$ and $\lim_K c_k^* = \liminf_I c_i^*$. Since $\lim_I X_i = X$, i.e., $\limsup_I X_i \subseteq X$ and

$$x \in \limsup_{j \in J} X_j \subseteq \limsup_{i \in I} X_i \subseteq X$$

(i.e., x is $(\mathcal{P})$-feasible), we have that

$$\liminf_{i \in I} c_i^* = \lim_{k \in K} c_k^* = \lim_{k \in K} c_k(x_k^*).$$

But $\lim_K c_k(x_k^*) = c(x)$, since

$$|c_k(x_k^*) - c(x)| \leq |c_k(x_k^*) - c(x_k^*)| + |c(x_k^*) - c(x)|,$$

c is continuous at x and the subnet $\{c_k\}_K$ converges uniformly to c on Y. Finally, $c(x) \geq c^*$, which completes this part of the proof.

Next we show that $\limsup_I c_i^* \leq c^*$. Let $x^* \in X^*$, so that $c^* = c(x^*)$. Since $X \subseteq \liminf_I X_i$, we have that $x^* \in \liminf_I X_i$. Hence, there exists a net $\{x_i\}_I$ such that $x_i \in X_i, \quad \forall i \in I$, for which $\lim_I x_i = x^*$. By definition of $\limsup_I c_i^*$, there exists a subnet $\{c_j^*\}_J$ of $\{c_i^*\}_I$ such that $\lim_J c_j^* = \limsup_I c_i^*$. Necessarily, $\lim_J x_j = x^*$ also. Let $c_j^* = c_j(x_j^*)$, where $x_j^* \in X_j^* \subseteq X_j, \quad \forall j \in J$. Then

$$c^* = c(x^*) = \lim_{j \in J} c_j(x_j) \geq \lim_{j \in J} c_j^* = \limsup_{i \in I} c_i^*,$$

which completes the proof. □

Remark. Note that the second part of the previous proof does not require Y to be compact. Also, it is not true in general that $\lim_I X_i^* = X^*$ (see Schochetman and Smith [1992a], for example). However, we do have a partial result in this direction.

Theorem 2.9. *Every accumulation point of optima drawn from the (P_i) is optimal for (P), i.e.,*

$$\limsup_I X_i^* \subseteq X^*.$$

Proof. Let $x \in \limsup_I X_i^*$. Then there exists a subnet $\{X_j\}_J$ of $\{X_i\}_I$ and a corresponding net $\{x_j^*\}$ such that $x_j^* \in X_j^*, \quad \forall j \in J$, and $\lim_J x_j^* = x$. Therefore, $x \in \limsup_I X_i$ by definition, so that $x \in X$ by hypothesis, i.e., $c^* \leq c(x)$. Since $\lim_J x_j^* = x$, we have that

$$c(x) = \lim_{j \in J} c_j(x_j^*) = \lim_{j \in J} c_j^* = c^*,$$

by Theorem 2.8, so that $x \in X^*$. □

The following easy corollary to Theorem 2.9 is a fundamental result. It states that *any* optimal solution to an approximating problem (P_i) arbitrarily well approximates some optimal solution to (P) for sufficiently large i.

Theorem 2.10. *Let Q be a compact subset of T and $\delta > 0$. Then there exists $i_0 \in I$ with the following property. For each $i \in I$ with $i_0 \leq i$, and for each $x_i^* \in X_i^*$, there exists $y^* \in X^*$ such that $d(x_i^*(t), y^*(t)) < \delta, \forall t \in Q$.*

Proof. Suppose not. Then, for each $i \in I$, there exists $k_i \in I$ such that $i \leq k_i$, and there exists $y_i \in X_{k_i}^*$ with the following property. For each $y^* \in X^*$, there exists $t_0 \in Q$ (depending on y^* and i) such that $d(y_i(t_0), y^*(t_0)) \geq \delta$. Now $\{y_i\}_I$ is a net in the compact space Y. Thus, there exists a subnet $\{y_j\}_J$ of $\{y_i\}_I$ and $y \in Y$ such that $\lim_J y_j = y$ in Y. Consequently, $y \in \limsup_I X_i^* \subseteq X^*$. Moreover, by the topology on Y, there exists $j_0 \in J$ such that $d(y_j(t), y(t)) < \delta, \forall t \in Q$, for all $j \geq j_0$ in J. In particular, $d(y_{j_0}(t)), y(t) < \delta, \forall t \in Q$. But, for $y^* = y$ and $i = j_0$, there exists $t_0 \in Q$ such that $d(y_{j_0}(t_0)), y(t_0) \geq \delta$. Contradiction. □

Example 2.11. (Discrete-Time, Continuous-Action Space with Discrete-Time Approximation Index) Let $T = I = \mathbb{N}$, $A = \mathbb{R}$, and let Y_t be a non-empty compact subset of $\mathbb{R}$, $\forall t \in \mathbb{N}$, with $Y = E = \prod_{\mathbb{N}} Y_t$.

For each $y \in Y$, let $K_t(y)$ (resp. $R_t(y)$) denote the cumulative cost (resp. revenue) attributed to y through time $t = 1, 2, \ldots$. We assume that for each $y \in Y$, the real functions (sequences) $t \to K_t(y)$ and $t \to R_t(y)$ on $\mathbb{N}$ are non-negative, non-decreasing and uniformly bounded by some exponential function, i.e., without loss of generality, there exist $B > 0$ and $\beta > 1$ such that

$$\max(K_t(y),\ R_t(y)) \leq B\beta^t, \quad \forall y \in Y, \quad \forall t = 1, 2, \ldots .$$

For each $y \in Y$, also define $C_t(y) = K_t(y) - R_t(y)$ to be the cumulative net cost of y through time $t = 1, 2, \ldots$. For each t, we assume that all costs and revenues incurred at time t are discounted by the discount factor $\alpha = (1 + \rho)^{-1}$, where $\rho > 0$ is the interest rate. Then, c_t is defined by

$$c_t(y) = \sum_{s=1}^{t} \alpha^{s-1}[C_s(y) - C_{s-1}(y)], \quad \forall y \in Y,\ \forall t = 1, 2, \ldots .$$

Similarly for the t-horizon discounted cost $k_t(y)$ and revenue $r_t(y)$, i.e.,

$$k_t(y) = \sum_{s=1}^{t} \alpha^{s-1}[K_s(y) - K_{s-1}(y)], \quad \forall y \in Y,\ \forall t = 1, 2, \ldots ,$$

and

$$r_t(y) = \sum_{s=1}^{t} \alpha^{s-1}[R_s(y) - R_{s-1}(y)], \quad \forall y \in Y,\ \forall t = 1, 2, \ldots ,$$

so that $c_t(y) = k_t(y) - r_t(y)$ in this case. Define the infinite horizon discounted net cost to be

$$c(y) = \sum_{s=1}^{\infty} \alpha^{s-1}[C_s(y) - C_{s-1}(y)] = \lim_{t\to\infty} c_t(y), \quad \forall y \in Y,$$

provided this limit exists. We similarly define the infinite horizon discounted cumulative cost and revenue functions $k(y)$ and $r(y)$, i.e.,

$$k(y) = \sum_{s=1}^{\infty} \alpha^{s-1}[K_s(y) - K_{s-1}(y)] = \lim_{t\to\infty} k_t(y), \quad \forall y \in Y,$$

and

$$r(y) = \sum_{s=1}^{\infty} \alpha^{s-1}[R_s(y) - R_{s-1}(y)] = \lim_{t\to\infty} r_t(y), \quad \forall y \in Y,$$

respectively. If $\rho > \beta - 1$, so that $0 < \alpha\beta < 1$, then for each $y \in Y$, the quantities $c(y)$, $k(y)$ and $r(y)$ do exist and $c(y) = k(y) - r(y)$. For each $y \in Y$, we also have that

$$|c(y) - c_t(y)| \leq a_t,$$

where a_t is defined by

$$a_t = \frac{4B}{\alpha(1 - \alpha\beta)}(\alpha\beta)^{t+1}, \quad \forall t = 1, 2, \ldots .$$

(See 2.2-2.5 of Schochetman and Smith [1989] for analogous results in the case where $T = \mathbb{R}$ and $\beta = e^{\gamma}$.) Consequently, the sequence $\{c_t\}_{\mathbb{N}}$ converges uniformly to c on Y. The same is true for k and r, i.e., $\forall t = 1, 2, \ldots,$

$$|k(y) - k_t(y)| \leq a_t,$$

and

$$|r(y) - r_t(y)| \leq a_t,$$

Finally, $c_t^* \leq c_s^* + a_s$, for all $s, t = 1, 2, \ldots$ such that $s \leq t$. If, in addition, we assume that for each $t = 1, 2, \ldots,$ the functions K_t, R_t, and hence C_t, are continuous real-valued functions on Y, then for each $t = 1, 2, \ldots,$ the real-valued functions k_t, r_t, c_t are continuous on Y. Hence, the real-valued functions k, r, c are also continuous on Y. (See 2.6-2.7 of Schochetman and Smith [1989].)

Example 2.12. (Continuous-Time, Continuous-Action Space with a Continuous-Time Approximation Index) Let $T = I = \mathbb{R}^+$ and $A = \mathbb{R}$, as in Examples 2.5 and 2.7. As in Schochetman and Smith [1989], let $K_t(y)$ (resp. $R_t(y)$) denote the cumulative cost (resp. revenue) attributed to $y \in Y$ through time $t \geq 0$. We assume that, for each $y \in Y$, the functions $t \to K_t(y)$ and $t \to R_t(y)$ on $\mathbb{R}^+$ are non-negative, non-decreasing and uniformly bounded by some exponential function, i.e., without loss of generality, there exist $B, \gamma > 0$ such that

$$\max(K_t(y),\ R_t(y)) \leq Be^{\gamma t}, \quad \forall y \in Y, \quad \forall t \geq 0.$$

Also define $C_t(y) = K_t(y) - R_t(y)$ to be the cumulative net cost of $y \in Y$ through time $t \geq 0$. Then, for each $y \in Y$, the function $t \to C_t(y)$ is of locally bounded variation on $\mathbb{R}^+$. For each $t \geq 0$, we assume that all costs and revenues incurred at time $0 \leq s \leq t$ are continuously discounted by $e^{-\rho s}$, where $\rho > 0$ is the interest rate. Thus, the t-horizon discounted net cost $c_t(y)$ for $y \in Y$ is given by the Stieltjes integral (see Widder [1946])

$$c_t(y) = \int_0^t e^{-\rho s} dC_s(y), \quad \forall t \geq 0.$$

The t-horizon discounted cost $k_t(y)$ and revenue $r_t(y)$ for y are obtained analogously, i.e.,

$$k_t(y) = \int_0^t e^{-\rho s} dK_s(y), \quad \forall t \geq 0,$$

and

$$r_t(y) = \int_0^t e^{-\rho s} dR_s(y), \quad \forall t \geq 0.$$

Obviously, $c_t(y) = k_t(y) - r_t(y)$, $\forall y \in Y$ and $\forall t \geq 0$. Define the infinite horizon discounted net cost $c(y)$ of $y \in Y$ to be the Laplace-Stieltjes transform (see Widder [1946])

$$c(y) = \int_0^\infty e^{-\rho s} dC_s(y) = \lim_{t \to \infty} c_t(y),$$

provided this limit exists. Similarly for the infinite horizon discounted cost and revenue functions $k(y)$ and $r(y)$, respectively, i.e.,

$$k(y) = \int_0^\infty e^{-\rho s} dK_s(y) = \lim_{t \to \infty} k_t(y), \quad \forall y \in Y,$$

and

$$r(y) = \int_0^\infty e^{-\rho s} dR_s(y) = \lim_{t\to\infty} r_t(y), \quad \forall y \in Y.$$

If $0 < \gamma < \rho$, then for each $y \in Y$, the quantities $c(y)$, $k(y)$ and $r(y)$ exist, and $c(y) = k(y) - r(y)$. For each $y \in Y$, and for each $t \geq 0$, we have $|c(y) - c_t(y)| \leq a_t$, where a_t is defined by

$$a_t = \frac{\rho B e^{(\gamma-\rho)t}}{\rho - \gamma}, \quad \forall t \geq 0.$$

(See 2.2-2.5 of Schochetman and Smith [1989].) Consequently, the net $\{c_t\}_{t\geq 0}$ of finite horizon net cost functions converges uniformly to the infinite horizon net cost function c on Y. The same is true for k and r. If, for each $t \geq 0$, the functions K_t, R_t, and hence C_t, are continuous real-valued functions on Y, then the real-valued functions k_t, r_t, c_t are also continuous on Y. Thus, the real-valued functions k, r, c are likewise continuous on Y. (See 2.6-2.7 of Schochetman and Smith [1989].) Moreover, the net $\{c_t^*\}_{t\geq 0}$ of optimal costs satisfies $c_t^* \leq c_s^* + a_s$, $\forall 0 \leq s \leq t$ and $c^* = \lim c_t^*$ (see 3.2 of Schochetman and Smith [1989]).

§3 Approximation Algorithms and Policy Convergence

The problems $(\mathcal{P}_i)$, $\forall i \in I$, are viewed as approximating problems to $(\mathcal{P})$. In general, the directed set I is not countable. However, it is intuitively the case that an approximation algorithm for solving $(\mathcal{P})$ should be sequential in nature. (For example, consider the situation where an infinite horizon continuous-time optimization problem is being approximated by a sequence of finite horizon subproblems whose horizons are increasing without bound.) Consequently, for the remainder of this paper, we assume that N is a *countable* subset of the directed set I which is order-isomorphic to the positive integers and satisfies the property that for each $i \in I$, there exists $n_i \in N$ such that $i \preceq n_i$, i.e., N is a subnet (Kelley [1955, p.70]) of I. Formally, we are assuming that there exists an order-preserving, one-to-one mapping ϕ of $(\mathbb{N}, \leq)$ into $(I, \preceq)$ which satisfies: for each $i \in I$, there exists $n_i \in \mathbb{N}$ such that if $m \in \mathbb{N}$ is such that $n_i \leq m$, then $i \preceq \phi(m)$. We then have $N = \phi(\mathbb{N})$. For convenience, on N, we will use $\preceq$ and $\leq$ interchangeably.

Example 3.1. Let $(I, \preceq)$ be the positive reals $(\mathbb{R}^+, \leq)$ and N the positive integers $\mathbb{N}$. (More generally, N could be a strictly monotone sequence in $\mathbb{R}^+$ which is unbounded.) Alternately, let I denote the uncountable set of all finite subsets of $\mathbb{N}$ (with $\preceq$ given by $\subseteq$) with N the subset of I given by $N = \{\, \{1, 2, \ldots, n\} : n \in \mathbb{N}\}$.

In general, for each $n \in N$, let $\mathcal{A}_n^*$ be a closed, non-empty set of $(\mathcal{P}_n)$-optimal solutions, i.e., $\emptyset \neq \mathcal{A}_n^* \subseteq X_n^*$, so that $c_n^* = c_n(x)$, $\forall x \in \mathcal{A}_n^*$ and $\mathcal{A}_n^* \in \mathcal{K}(Y)$, $\forall n \in N$. Define

$$\mathcal{A}_\infty^* = \limsup_{n\in N} \mathcal{A}_n^*,$$

which is necessarily a closed subset of X (Klein and Thompson [1984, p.28]). It is also non-empty, i.e., $\mathcal{A}_\infty^* \in \mathcal{K}(Y)$, since the $\mathcal{A}_n^*$ are non-empty and Y is compact. Since $\mathcal{A}_n^* \subseteq X_n^*$, for all $n \in N$, we have that

$$\mathcal{A}_\infty^* = \limsup_{n\in N} \mathcal{A}_n^* \subseteq \limsup_{n\in N} X_n^* \subseteq \limsup_{i\in I} X_i^* \subseteq X^*,$$

by Theorem 2.12, i.e., the elements of $\mathcal{A}_\infty^*$ are optimal for $(\mathcal{P})$. In this context, as in Schochetman and Smith [1989], we define an *approximation algorithm* $\mathcal{A}^*$ for $(\mathcal{P})$ to be such a sequence $\{\mathcal{A}_n^*\}_N$, for some choice of N as above. If $\mathcal{A}^*$ is an approximation algorithm, then the elements of $\mathcal{A}_\infty^*$ will be called *$\mathcal{A}^*$-algorithmically* *optimal solutions.* The algorithm $\mathcal{A}^*$ cannot approximate the *other* elements of X^*. We will also say that the approximation algorithm $\mathcal{A}^*$ *converges* if

$$\liminf_{n \in N} \mathcal{A}_n^* = \limsup_{n \in N} \mathcal{A}_n^*,$$

in Y, i.e., if

$$\lim_{n \in N} \mathcal{A}_n^* = \mathcal{A}_\infty^*,$$

in $\mathcal{K}(Y)$.

Proposition 3.2. *Suppose $\mathcal{A}^*$ is an approximation algorithm for $(\mathcal{P})$ which admits a* unique *$\mathcal{A}^*$-algorithmically optimal solution, i.e., $\mathcal{A}_\infty^* = \{x^*\}$, for some $x^* \in X^*$. Then:*
(i) The algorithm $\mathcal{A}^*$ converges to $\{x^*\}$ in $\mathcal{K}(Y)$, i.e., $\lim_N \mathcal{A}_n^* = \{x^*\}$.
(ii) Every selection from the $\mathcal{A}_n^*$ converges to x^*, i.e., if x_n is any *element of $\mathcal{A}_n^*$, $\forall n \in N$, then $\lim_N x_n = x^*$ in Y.*

Proof. Both parts follow from Corollary 2.2 of Schochetman and Smith [1991]. □

Remark. If $(\mathcal{P})$ admits a unique optimal solution, i.e., if $X^* = \{x^*\}$, then $\emptyset \neq \mathcal{A}_\infty^* \subseteq \{x^*\}$, so that $\mathcal{A}_\infty^* = \{x^*\}$. Such optimization problems are said to be well-posed (Dontchev and Zolezzi [1993]).

Example 3.3. (Strictly Convex Programs) Consider the particular case where A is also a (real) topological vector space, X is a closed, non-empty, *convex* feasible region and c is a *strictly convex* objective function. In this case, $(\mathcal{P})$ is well-known to admit a unique solution.

Recall our fundamental notion of nearness of solutions in Y as near agreement over a compact subset of T. This guides our definition of neighborhoods of Y. Toward this end, let Q be a compact subset of T, $x \in Y$ and $\delta > 0$. Define

$$U_Q(x, \delta) = \{y \in Y : d(x(t), y(t)) < \delta, \quad \forall t \in Q\}$$

and

$$U_Q(G, \delta) = \bigcup_{x \in G} U_Q(x, \delta),$$

for $G \subseteq Y$. Note that $U_Q(x, \delta)$ is a basic open neighborhood of x in the relative topology of Y, so that $U_Q(G, \delta)$ is open in Y. In this context, we are able to give necessary and sufficient conditions for $\mathcal{A}_\infty^*$ to be a singleton.

Theorem 3.4. *The following are equivalent for an approximation algorithm $\mathcal{A}^* = \{\mathcal{A}_n^*\}_N$:*
(i) $\mathcal{A}_\infty^*$ is a singleton $\{x^*\}$.
(ii) Solution indices exist for $\mathcal{A}^*$, i.e., there exists $x^* \in \mathcal{A}_\infty^*$ such that, for each compact $Q \subseteq T$ and $\delta > 0$, there exists $n_0 \in N$ satisfying

$$\mathcal{A}_n^* \subseteq U_Q(x^*, \delta),$$

for all $n \in N$ such that $n \geq n_0$.
(iii) Policy convergence takes place for all approximate solution subsequences generated by $\mathcal{A}^*$, i.e., there exists $x^* \in \mathcal{A}_\infty^*$, such that for all subsequences $\{\mathcal{A}_{n_j}^*\}$ of $\{\mathcal{A}_n^*\}$, and corresponding sequences $\{x_j\}$ for which $x_j \in \mathcal{A}_{n_j}^*$, $\forall j = 1, 2 \dots$, we have $\lim x_j = x^*$ in Y.

Proof. (i) $\implies$ (ii). Suppose $\mathcal{A}_\infty^* = \{x^*\}$. Then $\lim_N \mathcal{A}_n^* = \{x^*\}$ in $\mathcal{K}(Y)$ by Proposition 3.2. Suppose (ii) is false for x^*. Then, there exist Q and δ as in (ii) such that for each $n \in N$, there exists $j_n \in N$

such that $j_n \geq n$ and $\mathcal{A}_{j_n}^* \not\subseteq U_Q(x^*, \delta)$, i.e., the intersection of $\mathcal{A}_{j_n}^*$ with the complement of $U_Q(x^*, \delta)$ in Y is non-empty. Thus, there exists a subsequence $\{\mathcal{A}_{j_n}^*\}$ of $\{\mathcal{A}_n^*\}_N$, and a corresponding sequence $\{y_n\}_N$, such that $y_n \in \mathcal{A}_{j_n}^*$, but $y_n \notin U_Q(x^*, \delta)$, $\forall n \in N$. Since $U_Q(x^*, \delta)$ is open, its complement is closed and hence, compact. Hence, passing to a subsequence if necessary, we may assume that $\lim_N y_n = y$, for some $y \notin U_Q(x^*, \delta)$. In particular, $y \neq x^*$. But $y_n \in \mathcal{A}_{j_n}^*$, $\forall n \in N$, so that $y \in \limsup_N \mathcal{A}_n^* = \mathcal{A}_\infty^*$, i.e., $y = x^*$, by hypothesis. Contradiction.

(ii) $\Longrightarrow$ (iii). Suppose x^* is as in (ii). Let $\{\mathcal{A}_j^*\}_J$ be a subsequence of $\{\mathcal{A}_n^*\}_N$ and $\{x_j\}_J$ a corresponding sequence satisfying $x_j \in \mathcal{A}_j^*$, $\forall j \in J$. Let Q be a compact subset of T and $\delta > 0$. By (ii), there exists $n_0 \in N$ such that $\mathcal{A}_n^* \subseteq U_Q(x^*, \delta)$, for all $n \in N$ satisfying $n \geq n_0$. By definition of subsequence, there exists $j_0 \in J$ such that whenever $j \in J$ and $j \geq j_0$, the image of j in N is at least n_0. Consequently, for $j \geq j_0$ we have

$$d(x_j(t), x^*(t)) < \delta, \quad \forall t \in Q,$$

so that $x_j \in U_Q(x^*, \delta)$, $\forall j \geq j_0$, i.e., $\lim_J x_j = x^*$.

(iii) $\Longrightarrow$ (i). By hypothesis, $\{x^*\} \subseteq \mathcal{A}_\infty^*$. Suppose y^* in $\mathcal{A}_\infty^*$. Then there exists a subsequence $\{\mathcal{A}_j^*\}_J$ of $\{\mathcal{A}_n^*\}_N$ and a corresponding sequence $\{x_j\}_J$ such that $x_j \in \mathcal{A}_j^*$, $\forall j \in J$, and $\lim_J y_j = y$. By (iii), $\lim_J y_j = x^*$. Since Y is Hausdorff, it follows that $y = x^*$. □

§4 Stopping Rule for a Finite Algorithm

Let $\mathcal{A}^*$ be an approximation algorithm for $(\mathcal{P})$ as in section 3. In this section, under suitable additional assumptions, we present a Stopping Rule for this algorithm, as well as sufficient conditions for this Stopping Rule to be satisfied at some $m \in N$, given compact $Q \subseteq Y$ and $\delta > 0$. To do this, we require some additional notation and ideas. For the remainder of the paper, we adopt the following assumptions.

Assumptions

(1) The X_n are nested downward, i.e., if $n \geq j$ in N, then $X_n \subseteq X_j$. Note that since $\lim_N X_n = X$, we also have that $X = \cap_N X_n$ (Klein and Thompson [1984, p.28]).

(2) Costs are monotonically increasing, i.e., if $n \geq j$ in N, then $c_n(y) \geq c_j(y)$, $\forall y \in X_n$, where $X_n \subseteq X_j$ by (1).

Example 4.1. In Examples 2.11 and 2.12, simply set $R = 0$ in order to satisfy both of these assumptions.

Lemma 4.2. *The sequence $\{c_n^*\}_N$ is monotonically non-decreasing and bounded above by c^*, i.e., $c_n^* \uparrow c^*$, as $n \to \infty$.*

Proof. For $n \geq j$ in N, let $x_n^* \in X_n^*$, so that $c_n^* = c_n(x_n^*)$, i.e., $c_n^* \geq c_j(x_n^*)$, by assumption (2). Since $x_n^* \in X_j$ by assumption (1), it follows that $c_j(x_n^*) \geq c_j^*$, which proves the first part. For the second part, recall Theorem 2.8. □

Let $\{a_n\}_N$ be a sequence of real numbers satisfying $a_n \geq c^* - c_n^*$, so that $a_n \geq 0$, $\forall n \in N$. Note that in general, the sequence $\{a_n\}_N$ need not converge.

Example 4.3. If we know some $b \geq c^*$, then we can let $a_n = b - c_n^*$, $\forall n \in N$. As an example, set $b = c(x)$, for any $x \in X$. Thus, there exist many such sequences. In particular, recall examples 2.11 and 2.12.

Our next Lemma establishes upper and lower bounds on the optimal costs c_n^* as we increase n.

Lemma 4.4. *For $j \leq n$ in N, we have*

$$c_j^* \leq c_n^* \leq c_j^* + a_j.$$

Proof. By the previous lemma,

$$c_j^* \leq c_n^* \leq c^* = (c^* - c_j^*) + c_j^* \leq a_j + c_j^*,$$

by the choice of a_j. □

We next define a measure of error in value from optimal associated with solutions close to optimal for problem (P_m). Define

$$M_Q(\delta, m) = \inf\{c_m(x) : x \in X_m \backslash U_Q(\mathcal{A}_m^*, \delta)\},$$

where the slash denotes set difference (in Y). Since $U_Q(\mathcal{A}_m^*, \delta)$ is an open subset of Y, it follows that $X_m \backslash U_Q(\mathcal{A}_m^*, \delta)$ is a closed (hence, compact) subset of Y, because it is equal to the intersection of X_m and the complement of $U_Q(\mathcal{A}_m^*, \delta)$ in Y. Thus, $M_Q(\delta, m)$ is attained, since c_m is continuous. Recall that $c_m^* = c_m(x)$, $\forall x \in \mathcal{A}_m^*$.

Stopping Rule. Fix compact $Q \subseteq T$ and $\delta > 0$. Let $m \in N$. Then stop at m if
(Policy Criterion)

$$d(x(t), y(t)) < \delta/3, \quad \forall t \in Q, \quad \forall x, y \in \mathcal{A}_m^*,$$

and
(Value Criterion)

$$M_Q(\delta/3, m) - c_m^* > 2a_m.$$

We will say that the algorithm $\mathcal{A}^*$ *terminates* at m if the Policy and Value Criteria are satisfied for m, which we will call a *solution index of tolerance δ and support Q.*

Solution Index Algorithm

1. Choose compact $Q \subseteq T$, $\delta > 0$ and set $m = 1$.
2. Solve (P_m) to get $\mathcal{A}_m^*$ and c_m^*, which is equal to $c_m(x)$, for any $x \in \mathcal{A}_m^*$.
3. If the Stopping Rule is not satisfied, set $m = m + 1$ and go to step 2.
4. Otherwise, stop. In this event, m is a solution index of tolerance δ and support Q.

Lemma 4.5. *If $\mathcal{A}^*$ terminates at m, then for all $n \geq m$, we have $X_n^* \subseteq U_Q(\mathcal{A}_m^*, \delta/3)$.*

Proof. If not, then there exists $n \geq m$ such that X_n^* is not a subset of $U_Q(\mathcal{A}_m^*, \delta/3)$. Hence, there exists $x^* \in X_n^* \backslash U_Q(\mathcal{A}_m^*, \delta/3)$. Then, since the X_n are nested downward,

$$X_n^* \subseteq X_n \subseteq X_m$$

and

$$x^* \in X_m \backslash U_Q(\mathcal{A}_m^*, \delta/3).$$

Thus, by definition,

$$M_Q(\delta/3, m) \le c_m(x^*).$$

Also,

$$c_m(x^*) - c_m^* \ge M_Q(\delta/3, m) - c_m^* > 2a_m,$$

by the Value Criterion, i.e.,

$$c_m(x^*) - c_m^* > 2a_m.$$

On the other hand,

$$c_n^* = c_n(x^*) \ge c_m(x^*) \ge c_m(x^*) - a_m,$$

by Assumption (2). Consequently, adding the previous two inequalities together, we obtain that

$$c_n^* - c_m^* > a_m,$$

i.e.,

$$c_n^* > c_m^* + a_m.$$

By Lemma 4.4, this is a contradiction. Therefore, $X_n^* \subseteq U_Q(\mathcal{A}_m^*, \delta/3), \quad \forall n \ge m.$ □

Theorem 4.6. *If $\mathcal{A}^*$ terminates at m, then for each $x^* \in \mathcal{A}_\infty^*$, and each $n \ge m$, $\mathcal{A}_n^* \subseteq U_Q(x^*, \delta)$, i.e., for each $x \in \mathcal{A}_n^*$, we have*

$$d(x(t), x^*(t)) < \delta, \quad \forall t \in Q.$$

Proof. Fix $x^* \in \mathcal{A}_\infty^*$ and $\epsilon > 0$. Note that $U_Q(x^*, \epsilon)$ is an open neighborhood of x^* in the topology of Y. By definition of $\mathcal{A}_\infty^*$ (Klein and Thompson [1984, p.24]), there exists $n \ge m$ such that $\mathcal{A}_n^* \cap U_Q(x^*, \epsilon) \ne \emptyset$, i.e., there exists $x_n \in \mathcal{A}_n^*$ such that

$$d(x_n(t), x^*(t)) < \epsilon, \quad \forall t \in Q.$$

Consequently, $x^* \in U_Q(x_n, \epsilon)$. We have

$$x_n \in \mathcal{A}_n^* \subseteq U_Q(\mathcal{A}_m^*, \delta/3),$$

by Lemma 4.5. Therefore, there exists $x_m \in \mathcal{A}_m^*$ such that $x_n \in U_Q(x_m, \delta/3)$, i.e.,

$$d(x_n(t), x_m(t)) < \delta/3, \quad \forall t \in Q.$$

Hence, by the Triangle Inequality,

$$d(x^*(t), x_m(t)) < \delta/3 + \epsilon, \quad \forall t \in Q,$$

so that

$$x^* \in U_Q(x_m, \delta/3 + \epsilon), \quad \forall \epsilon > 0.$$

We next show that there exists $y \in \mathcal{A}_m^*$ such that

$$d(y(t), x^*(t)) \le \delta/3, \quad \forall t \in Q.$$

For each positive integer k, by the previous argument, there exists $y_k \in \mathcal{A}_m^*$ such that $x^* \in U_Q(y_k, \delta/3+1/k)$. Then $\{y_k\}$ is a sequence in $\mathcal{A}_m^*$ which is compact. Hence, there exists an accumulation point y of this sequence in $\mathcal{A}_m^*$. This point x^* must have the property that

$$d(y(t), x^*(t)) \leq \delta/3, \quad \forall t \in Q.$$

Now let $n \geq m$ and $x \in \mathcal{A}_n^*$. By Lemma 4.5, $x \in U_Q(\mathcal{A}_m^*, \delta/3)$. Hence, there exists $x_m \in \mathcal{A}_m^*$ such that $x \in U_Q(x_m, \delta/3)$, i.e.,

$$d(x(t), x_m(t)) < \delta/3, \quad \forall t \in Q.$$

By the Policy Criterion,

$$d(x_m(t), y(t)) < \delta/3, \quad \forall t \in Q.$$

Thus, by the Triangle Inequality,

$$d(x(t), x^*(t)) < d(x(t), x_m(t)) + d(x_m(t), y(t)) + d(y(t), x^*(t)) < \delta, \quad \forall t \in Q,$$

which implies that $x \in U_Q(x^*, \delta)$, so that $\mathcal{A}_n^* \subseteq U_Q(x^*, \delta)$, $\forall n \geq m$. This completes the proof. □

Remark. The conclusion of Theorem 4.6 is true for all of X_n^*. However, we are assuming that our algorithm $\mathcal{A}^*$ yields only that portion of X_n^* given by $\mathcal{A}_n^* \subseteq X_n^*$.

We next give sufficient conditions for the algorithm $\mathcal{A}^*$ to terminate.

Theorem 4.7. *Suppose $X^* = \{x^*\}$ and the a_n can be chosen so that $\lim_N a_n = 0$. Then for each compact $Q \subseteq T$, and for each $\delta > 0$, the algorithm $\mathcal{A}^*$ terminates at some $m \in N$ (which depends on Q and δ).*

Proof. Let Q and δ be as above.
(*Policy Criterion*) Since $X^* = \{x^*\}$ and $\emptyset \neq \mathcal{A}_\infty^* \subseteq X^*$, we have that

$$\limsup_N \mathcal{A}_n^* = \lim_N \mathcal{A}_n^* = \mathcal{A}_\infty^* = \{x^*\},$$

by (i) of Proposition 3.2. Thus, by Theorem 3.4, there exists n_0 sufficiently large such that for $n \geq n_0$, we have $\mathcal{A}_n^* \subseteq U_Q(x^*, \delta/6)$, i.e.,

$$d(x_n(t), x^*(t)) < \delta/6, \quad \forall t \in Q,$$

and for all $x_n \in \mathcal{A}_n^*$. Therefore, for each $n \geq n_0$ and each x_n, $y_n \in \mathcal{A}_n^*$, we have by the Triangle Inequality that

$$d(x_n(t), y_n(t)) \leq d(x_n(t), x^*(t)) + d(x^*(t), y_n(t))$$

$$< \delta/3, \quad \forall t \in Q.$$

This establishes the Policy Criterion for every $n \in N$ satisfying $n \geq n_0$.
(*Value Criterion*) Given Q and δ, we obtain n_0 as above. Hence, for each $n \geq n_0$,

$$x^* \in U_Q(x_n, \delta/6), \quad \forall x_n \in \mathcal{A}_n^*,$$

because

$$d(x_n(t), x^*(t)) < \delta/6), \quad \forall t \in Q.$$

Since $\mathcal{A}_n^* \neq \emptyset$, $\forall n$, we have that

$$x^* \in U_Q(\mathcal{A}_n^*, \delta/6), \quad \forall n \geq n_0. \tag{$*$}$$

Therefore, if, for some $m \geq n_0$, we have that

$$M_Q(\delta/3, m) - c_m^* > 2a_m,$$

then the Value Criterion is satisfied at m, as is the Policy Criterion, i.e., $\mathcal{A}^*$ terminates at m.

Thus, suppose that for each $n \geq n_0$, the Value Criterion does *not* hold, i.e.,

$$M_Q(\delta/3, n) - c_n^* \leq 2a_n,$$

for such n. Since $M_Q(\delta/3, n)$ is attained, for each $n \geq n_0$, there exists $x^n \in X_n \backslash U_Q(\mathcal{A}_n^*, \delta/3)$ such that $c_n(x^n) = M_Q(\delta/3, n)$. Hence,

$$0 \leq c_n(x^n) - c_n^* \leq 2a_n,$$

for all $n \geq n_0$. But $\{x^n\}$ is a sequence in compact Y. Thus, there exists (Kelley [1955, p.138]) a subsequence $\{x^{n_k}\}$ of $\{x^n\}$ and $x \in Y$ such that $\lim_k x^{n_k} = x$. Necessarily,

$$0 \leq c_{n_k}(x^{n_k}) - c_{n_k}^* \leq 2a_{n_k},$$

for all k sufficiently large such that $n_k \geq n_0$. Let k_0 be sufficiently large so that $k \geq k_0$ implies $n_k \geq n_0$, and hence,

$$0 \leq c_{n_k}(x^{n_k}) - c_{n_k}^* \leq 2a_{n_k}.$$

Since $\lim_k a_{n_k} = 0$, we have that

$$\lim c_{n_k}(x^{n_k}) = \lim c_{n_k}^* = c^*,$$

by Theorem 2.9. Now $\lim_k x^{n_k} = x$ and $x^n \in X_n$, $\forall n$, so that $x \in \limsup_n X_n = \lim_n X_n$, i.e., $x \in X$. But, by the Triangle Inequality,

$$|c_{n_k}(x^{n_k}) - c(x)| \leq |c_{n_k}(x^{n_k}) - c(x^{n_k})| + |c(x^{n_k}) - c(x)|, \quad \forall k \geq k_0,$$

which converges to 0, since the function c is continuous and the c_{n_k} converge uniformly to c on Y. Therefore, we have that $c(x) = c^*$, so that x is optimal. Since $X^* = \{x^*\}$, it must be that $x = x^*$.

Now let $0 < \epsilon < \delta/6$. Since, $\lim_k x^{n_k} = x^*$ in Y, there exists k_1 (which we may assume is at least k_0) such that for $k \geq k_1$,

$$d(x^{n_k}(t), x^*(t)) < \epsilon < \delta/6, \quad \forall t \in Q,$$

i.e.,

$$x^{n_k} \in U_Q(x^*, \epsilon) \subseteq U_Q(x^*, \delta/6).$$

Moreover, recall that $x^* \in U_Q(\mathcal{A}_n^*, \delta/6)$, for $n \geq n_0$, by $(*)$ above. Then, for each $k \geq k_1$ (so that $n_k \geq n_0$), there exists $x_{n_k} \in \mathcal{A}_{n_k}^*$ for which

$$d(x^*(t), x_{n_k}(t)) < \delta/6, \quad \forall t \in Q.$$

Thus, by the Triangle Inequality, for $k \geq k_1$,

$$d(x^{n_k}(t), x_{n_k}(t)) \leq d(x^{n_k}(t), x^*(t)) + d(x^*(t), x_{n_k}(t))$$

$$< \delta/3, \quad \forall t \in Q,$$

i.e., $x^{n_k} \in U_Q(\mathcal{A}_{n_k}, \delta/3)$. However, $k \geq k_1 \geq k_0$ implies $n_k \geq n_0$, so that $M_Q(\delta/3, n_k)$ is attained as $c_{n_k}(x^{n_k})$ at

$$x^{n_k} \in X_{n_k} \backslash U_Q(\mathcal{A}_{n_k}, \delta/3),$$

i.e.,

$$x^{n_k} \notin U_Q(\mathcal{A}_{n_k}, \delta/3).$$

Contradiction. Consequently, there exists $n \geq n_0$ for which the Value Criterion holds. □

In the next section, we present an application where the hypotheses of Theorem 4.7 hold.

§5 An Application to Production Planning

Consider a production facility which produces one product over continuous time subject to a maximum production rate $M > 0$. Suppose that at any time $t \geq 0$, the cumulative demand through time t is given by $D(t)$, where the demand function $D : \mathbb{R}^+ \to \mathbb{R}^+$ is non-decreasing (therefore Riemann integrable over any bounded interval) and satisfies $D(0) = 0$.

Let $T = \mathbb{R}^+$ and $A = \mathbb{R}$, so that A^T is the set $\mathbb{R}^{\mathbb{R}^+}$ of all functions from $\mathbb{R}^+$ to $\mathbb{R}$ and $C(T, A)$ is the set of all such functions which are continuous (we view $\mathbb{R}^{\mathbb{R}^+}$ as a topological vector space of functions under pointwise operations and convergence). Define E and Y as in Example 2.5, with $Y_t = [0, \ Mt]$, $\forall t \geq 0$, so that $Y = E \cap \prod_{t \geq 0} Y_t$, i.e., $y \in Y$ if and only if $y : \mathbb{R}^+ \to \mathbb{R}$, $y(0) = 0$ and $y \in E$, that is,

$$0 \leq y(t) - y(s) \leq M(t - s), \quad \forall 0 \leq s \leq t.$$

Recall that Y is compact in $C(\mathbb{R}^+, \mathbb{R})$.

For each $y \in Y$, $y(t)$ denotes the cumulative production through time $t \geq 0$. The non-empty, compact Hausdorff space Y is also an equicontinuous family of real-valued functions on $\mathbb{R}^+$. In practical terms, Y consists of all (non-decreasing) cumulative production functions which do not exceed the maximum production rate M, and which reflect the fact that production begins at time $t = 0$. Also, let $\mathcal{K}(Y)$ be as in section 2.

In order to ensure that it is possible to satisfy demand at all times, we must assume that the demand function D and the production rate M are such that $D(t) \leq Mt$, $\forall t \geq 0$. Consequently, as in Example 2.7, the feasible region X is then the *convex* subset of $\mathbb{R}^{\mathbb{R}^+}$ given by

$$X = \{y \in Y : y(t) \geq D(t), \ \forall t \geq 0\}.$$

(If $D(t_0) > Mt_0$, for some $t_0 \geq 0$, then $X = \emptyset$.) Since the function $y(t) = Mt$ is in X, it is non-empty. Moreover, X is pointwise-closed in Y, so that X is compact, i.e., $X \in \mathcal{K}(Y)$. Also as in Example 2.7, let $(I, \preceq) = (\mathbb{R}^+, \leq)$ and define

$$X_t = \{y \in Y : y(s) \geq D(s), \quad \forall 0 \leq s \leq t\}, \quad \forall t \geq 0,$$

so that $\{X_t\}_{t \geq 0}$ is a net in $\mathcal{K}(Y)$ which is nested downward and satisfies

$$\lim_{t \to \infty} X_t = \cap_{t \geq 0} X_t = X.$$

In order to introduce the cost structure, for $t \geq 0$, let:
$h(t)$ = the unit *holding* cost at time t,
$p(t)$ = the unit *production* cost at time t, and
$q(t)$ = the unit *revenue* at time t.
We assume that the functions $h, p, q : \mathbb{R}^+ \to \mathbb{R}^+$ are continuous and bounded, i.e.,

$$\sup_{t \geq 0}\{h(t), \ p(t), \ q(t)\} < \infty.$$

Thus, for *any* choice of $\gamma > 0$, there exists $B > 0$ sufficiently large such that

$$\max\{h(t), \ p(t), \ q(t)\} \leq Be^{\gamma t}, \quad \forall t \geq 0.$$

Given this data, we will construct an infinite horizon optimization model for the production planning facility along the lines we have developed in sections 2 and 3.

Before we can specify the cost structure, we need to define the *inventory* and *sales* functions θ and σ, respectively. Let $\theta, \ \sigma : Y \times \mathbb{R}^+ \to \mathbb{R}^+$ be given by

$$\theta(y, t) = \max\{y(t) - D(t), 0\}, \quad \forall y \in Y, \quad \forall t \geq 0,$$

and

$$\sigma(y, t) = \min\{y(t), D(t)\}, \quad \forall y \in Y, \quad \forall t \geq 0.$$

Thus, if we follow production strategy $y \in Y$ over all time, then $\theta(y, t)$ represents the inventory on hand at time t, and $\sigma(y, t)$ represents the cumulative sales through time t. It is not difficult to verify that for any production strategy in Y, the inventory at time t is equal to total production through time t, less cumulative sales through time t, i.e.,

$$\theta(y, t) = y(t) - \sigma(y, t), \quad \forall y \in Y, \ \forall t \geq 0.$$

In particular, if $y \in X$, then

$$\sigma(y, t) = D(t), \ \forall t \geq 0,$$

so that $\sigma(y, t)$ is independent of y. Thus, for $y \in X$, we have

$$\theta(y, t) = y(t) - D(t), \quad \forall t \geq 0.$$

The following additional properties of θ and σ will be required in what follows.

(i) For all $t \geq 0$ and $y \in Y$, we have $0 \leq \theta(y,t), \sigma(y,t) \leq Mt$.

(ii) For all $t \geq 0$ and $x, y \in Y$, we have

$$|\theta(x,t) - \theta(y,t)|, \ |\sigma(x,t) - \sigma(y,t)| \leq |x(t) - y(t)|.$$

(iii) For each $y \in Y$, the function $t \to \theta(y,t)$ is Riemann integrable in t.

(iv) For each $y \in Y$, the function $t \to \sigma(y,t)$ is non-decreasing in t.

We are now ready to introduce a cost structure as in Example 2.12. Let $0 < \rho < \infty$ be a specified interest rate and choose $0 < \gamma < \rho$. Define the holding cost, production cost and revenue functions (respectively)

$$H, P, R : Y \times \mathbb{R}^+ \to \mathbb{R}^+$$

as follows. For each $y \in Y$ and $t \geq 0$, let:

$$H_t(y) = \int_0^t h(s)\theta(y,s)^\lambda ds$$

for any fixed $\lambda > 1$,

$$P_t(y) = \int_0^t p(s)dy(s)$$

and

$$R_t(y) = \int_0^t q(s)d\sigma(y,s).$$

Then $H_t(y)$ represents the cumulative holding cost for production strategy y through time t. Similarly for the cumulative production cost $P_t(y)$ and the cumulative revenue $R_t(y)$. Note that for $y \in X$, R_t is independent of y, i.e., it is constant on X. The reason for taking the λ-th power of θ in the definition of $H_t(y)$ will become clear shortly.

Also define the cumulative cost function $K : Y \times \mathbb{R}^+ \to \mathbb{R}^+$ by $K = H + P$, so that the cumulative net cost function $C : Y \times \mathbb{R}^+ \to \mathbb{R}^+$ is given by $C = K - R = H + P - R$. Then $K_t(y)$, $R_t(y)$ and $C_t(y)$ are as in Example 2.12. We leave it to the interested reader to verify that, for each $y \in Y$, the functions $t \to K_t(y)$ and $t \to R_t(y)$ are non-negative, non-decreasing and uniformly bounded by an exponential function, as required in Example 2.12. Furthermore, for each $t \geq 0$, the functions $y \to K_t(y)$ and $y \to R_t(y)$, and hence, $y \to C_t(y)$, are continuous in y. Thus, our model satisfies the hypotheses of Examples 2.5, 2.7 and 2.11. Consequently, we obtain a net $\{c_t\}_{t>0}$ of continuous cost functions and a continuous cost function c such that

$$\lim_{t \to \infty} c_t = c$$

uniformly on Y, where, for $t \geq 0$,

$$c_t(y) = \int_0^t e^{-\rho s} dC_s(y), \quad \forall y \in Y,$$

and

$$c(y) = \int_0^\infty e^{-\rho s} dC_s(y) = \lim_{t \to \infty} c_t(y), \quad \forall y \in Y.$$

Similarly for $h_t(y)$, $p_t(y)$, $r_t(y)$ and $h(y)$, $p(y)$, $r(y)$, so that $c_t = h_t + p_t - r_t$, $\forall t \geq 0$, and $c = h + p - r$, where c is a strictly convex function because of our choice of λ in the definition of $H_t(y)$. Furthermore,

$$|c(y) - c_t(y)| \leq \rho B e^{(\gamma - \rho)t} / (\rho - \gamma) = a_t, \quad \forall y \in Y, \quad \forall t \geq 0,$$

so that

$$\lim_{t \to \infty} a_t = 0.$$

Thus, $(\mathcal{P})$ is given by $\min_{x \in X} c(x)$, and for each $t \geq 0$, $(\mathcal{P}_t)$ is given by $\min_{x \in X_t} c_t(x)$. Consequently, X^* and X_t^* are non-empty, closed subsets of Y, i.e., $X^* \in \mathcal{K}(Y)$, $X_t^* \in \mathcal{K}(Y)$, $\forall t \geq 0$, and

$$\limsup_{t \geq 0} X_t \subseteq X^*.$$

Note that X^* is a singleton $\{x^*\}$, since c is strictly convex as in Example 3.3. Moreover, the optimal solution values c^* and c_t^*, $\forall t \geq 0$, satisfy

$$c_t^* \leq c_s^* + a_s, \quad \forall 0 \leq s \leq t,$$

and $c^* = \lim_{t \to \infty} c_t^*$.

From the Remark following Proposition 3.2, we have that $\mathcal{A}_\infty^* = \{x^*\}$ for all approximating algorithms $\mathcal{A}$. Then by Proposition 3.2 (ii), $x_t^* \to x^*$ as $t \to \infty$, for all t-horizon optima $x_t^*, t = 1, 2, 3, \ldots$, since $\mathbb{N}$ is a linearly ordered subset of $\mathbb{R}^+$. In particular, by (ii) of Theorem 3.4, we conclude that for all times $\tau < \infty$ and positive numbers δ, there exists a $t \in \mathbb{N}$ satisfying $|x_u^*(s) - x^*(s)| < \delta$, for all $0 \leq s \leq \tau$, and for all $u > t, u \in \mathbb{N}$. That is, given τ, all sufficiently distant finite horizon optima uniformly well approximate all components of the unique infinite horizon optimum x^* over $[0, \tau]$.

Turning to how large the horizon must be to approximate x^* for a given component error δ and given interval $[0, \tau]$, we need to assume $R = 0$ to conclude that the optimal costs C_t are monotonically increasing in t, so that Assumptions 1 and 2 in section 4 are satisfied as well (see Example 4.1). We can now invoke the stopping criterion in section 4 to finitely terminate the forward procedure of solving the problem for horizons $T = 1, 2, 3, \ldots$. We are guaranteed by Theorem 4.7 that the corresponding algorithm finitely converges.

REFERENCES

1. Bean, J. C., R. L. Smith (1984). Conditions for the Existence of Planning Horizons. *Mathematics of Operations Research* **9** 391–401.
2. ______(1993). Conditions for the Discovery of Solution Horizons. *Mathematical Programming* **59** 215-229.
3. Bes, C., S. Sethi (1988). Concepts of Forecast and Decision Horizons: Applications to Dynamic Stochastic Optimization Problems. *Mathematics of Operations Research.* **13** 295-310.
4. Carlson, D. A., A. B. Haurie, A. Leizarowitz (1987). *Infinite Horizon Optimal Control: Deterministic and Stochastic Systems, 2nd Edition*, Springer-Verlag, Berlin.
5. Dontchev, A. L., T. Zolezzi (1993). *Well-posed optimization problems*, Springer-Verlag, Berlin.
6. Dugundji, J., (1966). *Topology*, Allyn and Bacon, Boston.
7. Fiacco, A. V. (1974). Convergence properties of local solutions of sequences of mathematical programming problems in general spaces *Journal of Optimization Theory and Applications* **13** 1-12.
8. Kelley, J. L. (1955). *General Topology*, Van Nostrand, New York.
9. Klein, E., A. C. Thompson (1984) *Theory of Correspondances: Including Applications to Mathematical Economics*, John Wiley and Sons, New York.
10. Luenberger, D. G. (1969). *Optimization by Vector Space Methods*, John Wiley, New York.
11. Schochetman, I. E., R. L Smith (1989). Infinite Horizon Optimization. *Mathematics of Operations Research.* **14** 559–574.
12. ______(1991). Convergence of Selections with Applications in Optimization. *Journal of Mathematical Analysis and Applications.* **155** 278-292.
13. ______(1992a). Finite Dimensional Approximation in Infinite Dimensional Mathematical Programming. *Mathematical Programming.* **54** 307–333.
14. ______(1992b). Convergence of Best Approximations from Unbounded Sets. *Journal of Mathematical Analysis and Applications.* **166** 112-128.
15. ______(1997). Existence and Discovery of Average-Cost Optimal Solutions in Deterministic Infinite Horizon Optimization, forthcomimg in *Mathematics of Operations Research.*.
16. Semple, J. (1996). Infinite Positive-Definite Quadratic Programming in a Hilbert Space. *Journal of Optimization Theory and Applications* **88** 743–749..
17. Widder, D. V. (1946). *The Laplace Transform*, Oxford University Press, Oxford.