

Finite dimensional approximation in infinite dimensional mathematical programming

Irwin E. Schochetman

Department of Mathematical Sciences, Oakland University, Rochester, MI 48309, USA

Robert L. Smith*

Department of Industrial and Operations Engineering, The University of Michigan, Ann Arbor, MI 48109, USA

Received 18 April 1989

Revised manuscript received 20 August 1990

We consider the problem of approximating an optimal solution to a separable, doubly infinite mathematical program (P) with lower staircase structure by solutions to the programs (P(N)) obtained by truncating after the first N variables and N constraints of (P). Viewing the surplus vector variable associated with the N th constraint as a state, and assuming that all feasible states are eventually reachable from any feasible state, we show that the efficient set of all solutions optimal to all possible feasible surplus states for (P(N)) converges to the set of optimal solutions to (P). A tie-breaking algorithm which selects a nearest-point efficient solution for (P(N)) is shown (for convex programs) to converge to an optimal solution to (P). A stopping rule is provided for discovering a value of N sufficiently large to guarantee any prespecified level of accuracy. The theory is illustrated by an application to production planning.

Key words: Value convergence, reachability, solution set convergence, tie-breaking, stopping rule, infinite horizon optimization, production planning.

1. Introduction

Consider the following doubly infinite mathematical programming problem:

$$\begin{aligned} & \text{minimize} && \sum_{j=1}^{\infty} c_j(y_1, \dots, y_j) \\ & \text{subject to} && \sum_{j=1}^i a_{ij}(y_1, \dots, y_j) \geq b_i, \quad i = 1, 2, \dots, \\ & && y_j \in Y_j, \quad j = 1, 2, \dots, \end{aligned}$$

where $Y_j \in \mathbb{R}^{n_j}$, $j = 1, 2, \dots$ and $b_i \in \mathbb{R}^{m_i}$, $i = 1, 2, \dots$. Apart from the implicit structural assumption that each constraint contains at most finitely many variables, (P) is at

* The work of Robert L. Smith was partially supported by the National Science Foundation under Grant ECS-8700836.

this point completely general, including integer programs where each Y_j is discrete. Although there is an extensive literature on the solution of semi-infinite mathematical programs (see for example Anderson and Nash [1]), there is relatively little work on the doubly infinite case (Grinold [8, 9, 10], Jones, Zydiak and Hopp [12], and Flam and Wets [7] being notable exceptions). In this paper, we establish conditions on (P) which allow for arbitrarily good approximations to the solutions of (P) by solving finite dimensional approximations to (P) obtained by truncating beyond finitely many variables and constraints.

If we interpret the index j in (P) to correspond to the j th period in a multi-period planning problem, then the above truncation to (P) becomes a finite horizon approximation of an infinite horizon optimization problem. This so-called planning (or solution) horizon approach has an extensive literature (see for example Bean and Smith [3], Bès and Sethi [5], and Schochetman and Smith [19]). Our paper extends this work to a general mathematical programming framework where the requirement of [3, 5] that variables be discrete and the assumption of [19] that all finite horizon feasible solutions are extendable to infinite horizon feasible solutions do not hold in general.

Moreover, in the presence of a certain property, which we call *reachability*, we show how to enlarge the set of finite dimensional optimal solutions to guarantee Hausdorff convergence to the set of infinite dimensional optimal solutions as the dimensionality increases. The tie-breaking algorithm of Schochetman and Smith [20] is then employed to select a *sequence* of finite dimensional optimal solutions that converges to an infinite dimensional optimal solution as the dimensionality of the approximation increases. This differs fundamentally from [7]–[10] where it is shown only that convergent *subsequences* exist.

Throughout this paper we adopt the following.

Assumptions.

(A) The feasible activity vectors y_j lie in non-empty, compact regions, i.e., each Y_j is a non-empty, compact subset of \mathbb{R}^{n_j} , so that there exists $r_j > 0$ such that $|y_j|_j \leq r_j$, for all $y_j \in Y_j$ and $j = 1, 2, \dots$, where $|\cdot|_j$ is the Euclidean norm on \mathbb{R}^{n_j} .

(B) The constraint functions are continuous, i.e., for each $i = 1, 2, \dots$, a_{ij} is a continuous, real-valued function on $Y_1 \times \dots \times Y_j$, for $j = 1, \dots, i$.

(C) The objective function is continuous and absolutely convergent, i.e., for each $j = 1, 2, \dots$, c_j is a continuous, real-valued function on $Y_1 \times \dots \times Y_j$ and $\sum_{j=1}^{\infty} \|c_j\|_{\infty} < \infty$, where

$$\|c_j\|_{\infty} = \max\{|c_j(y_1, \dots, y_j)|: y_k \in Y_k, k = 1, \dots, j\}.$$

Note that this implies uniform convergence of the objective function.

In Section 2, we establish a Hilbert space context for (P) and note that, as a consequence of Assumption A, the product and metric topologies on $\prod_{j=1}^{\infty} Y_j$ are

identical. We then show that under Assumptions A, B and C, (P) has an optimal solution.

In Section 3, we formally introduce the finite dimensional approximations, (P(N)) to (P) effectively consisting of the first N variables and first N constraints of (P), $N = 1, 2, \dots$, and study their inter-relationships.

In Section 4, we prove that the sequence of optimal objective function values to the approximating problems (P(N)) converges to the optimal value of (P), as $N \rightarrow \infty$ (i.e., value convergence). We also show that for convex programs with strictly convex objective functions, the sequence of optimal solutions to the problems (P(N)) also converges to the (unique) optimal solution of (P), as $N \rightarrow \infty$ (i.e., solution convergence).

In Section 5, we consider the class of problems where the objective function and constraints are separable and each variable appears in at most finitely many constraints. Treating the surplus vector associated with the N th constraint of (P(N)) as a state variable, we form for each N , the closure of the set of all optimal solutions to *all* feasible states. In the presence of our reachability property, we show that this sequence of optimal sets converges to the set of optimal solutions of (P) in the underlying Hausdorff metric. Thus, under reachability we have solution *set* convergence. Also in this section, we discuss a natural tie-breaking rule for selecting solutions to the (P(N)) in such a way that the resulting solution sequence converges to an optimal solution of (P). We also discuss a forward algorithm which, together with a stopping rule for determining N , yields any desired level of accuracy. Finally, in Section 6, we discuss an application in production planning.

2. Mathematical preliminaries

We begin by constructing an infinite dimensional Hilbert space in which we embed (P).

Each \mathbb{R}^{n_j} is a Hilbert space with usual inner product, norm and metric denoted respectively by $\langle \cdot, \cdot \rangle_j$, $|\cdot|_j$ and d_j , $j = 1, 2, \dots$. If we let $Y = \prod_{j=1}^{\infty} Y_j$, then by Assumption A and the Tychonoff Theorem [13, p. 143], Y is a non-empty, compact subset of $\prod_{j=1}^{\infty} \mathbb{R}^{n_j}$ relative to the product of the metric topologies.

Fix $0 < \beta_j < 1$, each j , such that $\sum_{j=1}^{\infty} \beta_j^2 r_j^2 < \infty$ (for example, $\beta_j = 1/jr_j$). For each $j = 1, 2, \dots$, denote by H_j the Hilbert space given by \mathbb{R}^{n_j} as a (real) linear space with inner product, norm and metric given respectively as follows:

$$\begin{aligned} \langle x_j, y_j \rangle_j &= \beta_j^2 \langle x_j, y_j \rangle_j, \\ \|x_j\|_j &= \beta_j |x_j|_j, \\ \rho_j(x_j, y_j) &= \beta_j d_j(x_j, y_j), \quad x_j, y_j \in H_j. \end{aligned}$$

Of course, the metrics ρ_j and d_j are topologically equivalent for all j , so that Y is also a compact subset of $\prod_{j=1}^{\infty} H_j$.

Let H denote the Hilbert sum [2, p. 222] of the H_j , i.e.,

$$H = \left\{ (x_j) \in \prod_{j=1}^{\infty} H_j : \sum_{j=1}^{\infty} \|x_j\|_j^2 = \sum_{j=1}^{\infty} \beta_j^2 |x_j|_j^2 < \infty \right\},$$

with inner product given by

$$(x, y) = \sum_{j=1}^{\infty} (x_j, y_j)_j = \sum_{j=1}^{\infty} \beta_j^2 \langle x_j, y_j \rangle_j.$$

Hence, the corresponding norm and metric are given by

$$\|x\| = \left[\sum_{j=1}^{\infty} \|x_j\|_j^2 \right]^{1/2} = \left[\sum_{j=1}^{\infty} \beta_j^2 |x_j|_j^2 \right]^{1/2}$$

and

$$\rho(x, y) = \left[\sum_{j=1}^{\infty} \rho_j(x_j, y_j)^2 \right]^{1/2} = \left[\sum_{j=1}^{\infty} \beta_j^2 d_j(x_j, y_j)^2 \right]^{1/2}, \quad x, y \in H.$$

From the choice of the β_j , it follows that

$$Y \subseteq H.$$

Thus, Y inherits a metric structure from H via ρ .

Lemma 2.1. *The ρ -metric topology and the product metric topology on Y are the same.*

Proof. See pages 120 and 153 of [16]. \square

Thus, since Y is compact in the product topology, it is a compact metric space relative to ρ . Moreover, a sequence $\{y^n\}$ in Y converges to y relative to ρ if and only if for each j , $\{y_j^n\}$ converges to y_j in \mathbb{R}_j^n relative to the usual Euclidean metric.

Define $\mathcal{H}(Y)$ to be the space of all compact, non-empty subsets of Y and let D denote the Hausdorff metric on $\mathcal{H}(Y)$ corresponding to ρ . Recall that for C, K in $\mathcal{H}(Y)$, we have

$$D(C, K) = \max \left(\max_{x \in C} \rho(x, K), \max_{y \in K} \rho(y, C) \right),$$

where $\rho(x, K) = \min_{y \in K} \rho(x, y)$, $x \in C$. In this manner, $\mathcal{H}(Y)$ becomes a compact metric space [11, 15]. This metric will prove useful in Sections 4 and 5 where we study solution set convergence.

There is an alternate characterization of set convergence which will also prove very useful to us later on. Let $K_N \subseteq Y$, for $N = 1, 2, \dots$. Define $\liminf K_N$ and $\limsup K_N$ as follows [4, 11, 15]:

(1) $y \in \liminf K_N$ if and only if $y \in Y$ and, for each N sufficiently large, there exists y^N in K_N such that $y^N \rightarrow y$, as $N \rightarrow \infty$.

(2) $y \in \limsup K_N$ if and only if $y \in Y$ and there exists a subsequence $\{K_{N_k}\}$ of $\{K_N\}$ and a corresponding sequence $\{y^k\}$ such that $y^k \in K_{N_k}$, all k , and $y^k \rightarrow y$, as $k \rightarrow \infty$.

If $K \subseteq Y$ and $K = \liminf K_N = \limsup K_N$, we write

$$\lim K_N = K.$$

In general, $\liminf K_N$ and $\limsup K_N$ are closed subsets of Y , which may be empty and which satisfy $\liminf K_N \subseteq \limsup K_N$. Now suppose K_N is not empty for N large. Since Y is compact, we have that $\limsup K_N \neq \emptyset$. Also, if $K \in \mathcal{H}(Y)$ and $K_N \in \mathcal{H}(Y)$, all N , then

$$K_N \rightarrow K \text{ in } \mathcal{H}(Y) \text{ relative to } D$$

if and only if

$$K = \liminf K_N = \limsup K_N,$$

i.e., $\limsup K_N \subseteq K$ and $K \subseteq \liminf K_N$ [11, 14, 15]. Thus, $\lim K_N = K$ also.

For each $i = 1, 2, \dots$, recall that b_i is an element of the Euclidean space \mathbb{R}^{m_i} , so that

$$a_{ij} : Y_1 \times \dots \times Y_j \rightarrow \mathbb{R}^{m_i}, \quad j = 1, 2, \dots, i.$$

Thus, the i th constraint is a vector inequality involving the first i variables at most. It may be interpreted as a system of m_i scalar constraints in the components of the $y_j, j = 1, 2, \dots, i$. By Assumption B, each a_{ij} is a continuous function, so that the function

$$\sum_{j=1}^i a_{ij} : Y_1 \times \dots \times Y_i \rightarrow \mathbb{R}^{m_i}$$

is also continuous, $i = 1, 2, \dots$. Hence, if we define

$$F_i = \left\{ y \in Y : \sum_{j=1}^i a_{ij}(y_1, \dots, y_j) \geq b_i \right\}, \quad i = 1, 2, \dots,$$

then F_i is a closed and hence compact subset of Y . The *feasible region* X of (P) is then given by

$$X = \bigcap_{i=1}^{\infty} F_i.$$

Thus, X , being an intersection of compact sets, is itself a compact subset of Y . To avoid the trivial case, we assume that $X \neq \emptyset$, so that each $F_i \neq \emptyset$, as well. A necessary and sufficient condition for this is that the sets $\{F_i\}$ have the finite intersection property [13, p. 136]. Hence, under our assumptions, we have that the feasible region X is in $\mathcal{H}(Y)$.

Turning to the objective function of (P), for each $j = 1, 2, \dots$, we have that

$$c_j : Y_1 \times \dots \times Y_j \rightarrow \mathbb{R}.$$

By Assumption C, each c_j is continuous. Since $Y_1 \times \dots \times Y_j$ is compact, we have that the sup norm of c_j ,

$$\|c_j\|_\infty = \max_{Y_1 \times \dots \times Y_j} |c_j(y_1, \dots, y_j)| < \infty, \quad j = 1, 2, \dots$$

By Assumption C, we have moreover that $\sum_{j=1}^\infty \|c_j\|_\infty < \infty$. A sufficient condition for Assumption C to hold is that each c_j be of the form $\alpha^j k_j$, where $0 < \alpha < 1$ is a discount factor, k_j is a continuous function on $Y_1 \times \dots \times Y_j$ and the k_j are uniformly bounded.

Theorem 2.2. *The correspondence*

$$C(y) = \sum_{j=1}^\infty c_j(y_1, \dots, y_j)$$

defines a continuous function $C : Y \rightarrow \mathbb{R}$ which is the uniform limit of the sequence of continuous functions $C(\cdot; N)$ given by

$$C(y; N) = \sum_{j=1}^N c_j(y_1, \dots, y_j), \quad y \in Y.$$

In particular, given $\epsilon > 0$, there exists $N_\epsilon > 0$ such that $N \geq N_\epsilon$ implies that

$$\left| \sum_{j=N+1}^\infty c_j(y_1, \dots, y_j) \right| < \epsilon, \quad y \in Y.$$

Proof. This follows from Assumptions A, B and C. \square

Since C is continuous on Y and X is a compact, non-empty subset of Y , C attains its minimum on X . Let

$$C^* = \min_{x \in X} C(x)$$

be the optimal objective value of (P) and let

$$X^* = \{x \in X : C(x) = C^*\}$$

denote the set of optimal solutions. Then X^* is also a compact, non-empty subset of Y , i.e., $X^* \in \mathcal{H}(Y)$, where

$$\emptyset \neq X^* \subseteq X \subseteq Y.$$

Our primary objective in this paper is to approximate the optimal objective value C^* and an optimal solution x^* in X^* via corresponding quantities computed from finite dimensional approximations of (P).

3. The finite dimensional approximations

We begin by defining finite dimensional approximations to (P) formed by retaining the first N variables and the first N constraints. For each $N = 1, 2, \dots$, let $(P(N))$ be the finite dimensional problem given by

$$\begin{aligned} (P(N)) \quad & \text{minimize} \quad \sum_{j=1}^N c_j(y_1, \dots, y_j) \\ & \text{subject to} \quad \sum_{j=1}^i a_{ij}(y_1, \dots, y_j) \geq b_i, \quad i = 1, \dots, N, \\ & \quad \quad \quad y \in Y. \end{aligned}$$

Note that $(P(N))$ has N constraints in the N vector variables y_1, \dots, y_N , with the remaining y_j being arbitrary. Thus, $(P(N))$ is a finite dimensional problem having the advantage that all solutions are comparable as elements of Y . Our intent is to ultimately approximate solutions of (P) by solutions of the $(P(N))$. Let $X(N)$ denote the feasible region of $(P(N))$, for $N = 1, 2, \dots$. Then $X(N)$ is a non-empty, compact subset of Y , since $X(N) = \bigcap_{i=1}^N F_i$, i.e., $X(N) \in \mathcal{K}(Y)$, $N = 1, 2, \dots$.

Lemma 3.1. *For each $N = 1, 2, \dots$, we have*

- (i) $X(N) = \left\{ y \in Y : \sum_{j=1}^i a_{ij}(y_1, \dots, y_j) \geq b_i, i = 1, 2, \dots, N \right\}$,
- (ii) $X(N+1) \subseteq X(N)$. □

Since X is the non-empty intersection of the sets $X(N)$ in $\mathcal{K}(Y)$, i.e., $X = \bigcap_{N=1}^{\infty} X(N)$, we have that

$$X(N) \rightarrow X \text{ in } \mathcal{K}(Y), \text{ as } N \rightarrow \infty,$$

relative to the Hausdorff metric D [14, p. 339].

Turning to the objective function of $(P(N))$, let

$$C(y; N) = \sum_{j=1}^N c_j(y_1, \dots, y_j), \quad N = 1, 2, \dots,$$

where the domain of $C(\cdot; N)$ is Y . Then $(P(N))$ may be abbreviated as

$$\min_{y \in X(N)} C(y; N), \quad N = 1, 2, \dots$$

Also, recall that the $C(\cdot; N)$, when viewed as functions on Y , converge uniformly to C (Theorem 2.2).

Since $C(\cdot; N)$ is continuous and $X(N)$ is compact as well as non-empty, it follows that the minimum in $(P(N))$ is attained. Denote this optimal value by $C^*(N)$, $N = 1, 2, \dots$. Also let $X^*(N)$ be the set of optimal solutions of $(P(N))$, so that $X^*(N)$ is a compact, non-empty subset of $X(N)$, $N = 1, 2, \dots$.

In the next section, we search for conditions which allow for approximating the infinite dimensional optimal objective value C^* and an infinite dimensional optimal solution x^* by finite dimensional optimal objective values $C^*(N)$ and finite dimensional optimal solutions $x^*(N)$ for N sufficiently large.

4. Optimal value and solution convergence

We first establish (optimal) value convergence for the finite dimensional programs.

Theorem 4.1 (value convergence). *The optimal values of the finite dimensional problems $(P(N))$ converge to the optimal value of the infinite dimensional problem (P) , i.e.,*

$$C^*(N) \rightarrow C^*,$$

as $N \rightarrow \infty$. Also, $\limsup X^*(N) \subseteq X^*$.

Proof. Let $I = \{1, 2, \dots\} \cup \{\infty\}$. Then I is a compact metric space under stereographic projection. Define

$$f: Y \times I \rightarrow \mathbb{R}$$

by

$$f(y, N) = \begin{cases} C(y; N), & y \in Y, N = 1, 2, \dots, \\ C(y), & y \in Y, N = \infty. \end{cases}$$

We claim that f is continuous. In particular, if $y^n \rightarrow y$ in Y and $N_n \rightarrow \infty$, as $n \rightarrow \infty$, then

$$\begin{aligned} |f(y, \infty) - f(y^n, N_n)| &= |C(y) - C(y^n; N_n)| \\ &\leq |C(y) - C(y^n)| + |C(y^n) - C(y^n; N_n)| \\ &= |C(y) - C(y^n)| + \left| \sum_{j=N_n+1}^{\infty} c_j(y_j^n) \right|, \end{aligned}$$

which goes to zero as $n \rightarrow \infty$, since C is continuous and uniformly convergent (Theorem 2.2). Since $\lim X(N) = X$, it follows that the mapping $N \rightarrow X(N)$ (with $X(\infty) = X$) is continuous from I into $\mathcal{H}(Y)$. Hence, by the Maximum Theorem [4, p. 116], it follows that $\min\{f(y, N): y \in X(N)\}$ converges to $\min\{f(y, \infty): y \in X\}$, i.e., $C^*(N) \rightarrow C^*$, as $N \rightarrow \infty$, and that $\limsup X^*(N) \subseteq X^*$. \square

Having established value convergence, we next address the question of solution convergence. That is, we seek conditions under which it is possible to choose solutions optimal for $(P(N))$ which converge to a solution optimal for (P) . As we will see shortly, for strictly convex programs (P) , this convergence occurs for all solutions optimal for $(P(N))$. However, in general, not all solution selections optimal for the $(P(N))$ will converge. In fact, we will find it necessary to enlarge the set of possible solution selections from the $(P(N))$ to include non-optimal feasible solution selections with the property that the limit of any convergent subsequence of these is optimal for (P) .

More formally, we define a *finite dimensional algorithm* A^* for (P) to be a sequence $\{A^*(N), N = 1, 2, \dots\}$ where each $A^*(N)$ is a closed, non-empty subset of $X(N)$ satisfying

$$A^*(\infty) \subseteq X^*,$$

where $A^*(\infty)$ is defined to be $\limsup A^*(N)$, the set of accumulation points of the $A^*(N)$. In this case of course, $\liminf A^*(N)$, although possibly empty, must also be a subset of X^* . For example, if each $A^*(N)$ is a closed, non-empty subset of $X^*(N)$, then it follows from Theorem 4.1 that $A^* = \{A^*(N), N = 1, 2, \dots\}$ is a finite dimensional algorithm for (P) . In what follows, we will see that there exist finite dimensional algorithms A^* for which $A^*(N) \not\subseteq X^*(N)$, for all N .

We seek a finite dimensional algorithm A^* which is itself convergent, i.e, such that $A^*(N) \rightarrow A^*(\infty)$ in $\mathcal{H}(Y)$ relative to the Hausdorff metric D , as $N \rightarrow \infty$. For such A^* , we are interested in selecting $x^*(N)$ in $A^*(N)$, for each N , so that the sequence $\{x^*(N)\}$ converges to a solution in X^* . This selection is made as follows. Let $p \in Y$ and $K \in \mathcal{H}(Y)$. Then the distance $\rho(p, K)$ from p to K is attained, possibly non-uniquely. If $s_p(K)$ is such a point for each K in $\mathcal{H}(Y)$, then we call s_p a *nearest-point selection* defined by p . If K is such that $\rho(p, K)$ is *uniquely* attained by a point in K , then p is called a *uniqueness point* for K . The set of all such uniqueness points will be denoted by $U(K)$. Clearly, $\emptyset \neq K \subseteq U(K) \subseteq Y$ in general.

Theorem 4.2. *Let A^* be a finite dimensional algorithm for (P) and A a closed subset of X^* containing $A^*(\infty)$. Then the following are equivalent:*

- (i) $A^*(N) \rightarrow A$, relative to the Hausdorff metric D , as $N \rightarrow \infty$.
- (ii) $\liminf A^*(N) = \limsup A^*(N) = A$.
- (iii) For all $x^* \in A$, there exists $x^*(N) \in A^*(N)$, each N , such that $x^*(N) \rightarrow x^*$, relative to the ρ metric, as $N \rightarrow \infty$.
- (iv) For each p in $U(A)$ and each nearest-point selection s_p defined by p , $s_p(A^*(N)) \rightarrow s_p(A)$, as $N \rightarrow \infty$.

In particular, if $A = X^$ is a singleton, where $X^* = \{x^*\}$, then all the above conditions are satisfied; in fact,*

$$x^*(N) \rightarrow x^*,$$

as $N \rightarrow \infty$, where $x^(N)$ is any element of $X^*(N)$, $N = 1, 2, \dots$*

Proof. This theorem follows immediately from Schochetman and Smith [20] in conjunction with our previous discussion. \square

Analogously, we have the following dual version of Theorem 4.2.

Theorem 4.3. *Suppose A^* is a finite dimensional algorithm for (P) with A as above. Then the following are equivalent:*

- (i) $\liminf A^*(N) \neq \emptyset$.
- (ii) *There exists an $x^* \in A$ for which there exists $x^*(N) \in A^*(N)$, each N , such that $x^*(N) \rightarrow x^*$, relative to the ρ metric, as $N \rightarrow \infty$.*

(iii) *There exists p in $U(A)$ such that $s_p(A^*(N)) \rightarrow s_p(A)$, as $N \rightarrow \infty$, for all nearest-point selections s_p defined by p . \square*

Theorem 4.3 gives the existence of limit points of A^* in (i) as a necessary condition for approximating a solution to (P) by suitable choices of solutions to the $(P(N))$ from the $A^*(N)$. In practice, the easier way to demonstrate this condition is to establish the stronger claim that A^* converges to some non-empty set A , the latter being condition (i) of Theorem 4.2. Under this condition, any nearest-point selection (defined by p) of solutions from the $A^*(N)$ is guaranteed to converge, provided that there is a unique point closest to p from A . Since $A \subseteq Y$, we also have $A \subseteq H$, where H is a Hilbert space. Hence, a sufficient condition for there to be a unique point in A closest to p , for any $p \in Y$, is that A be closed and convex [2, p. 15]. In particular, for (P) a convex program, X^* is a convex subset of H , where (P) is said to be *convex* if:

- (i) Y_j is convex, $j = 1, 2, \dots$
- (ii) $-a_{ij}$ is convex, $j = 1, \dots, i, i = 1, 2, \dots$
- (iii) c_j is convex, $j = 1, 2, \dots$

We therefore obtain the following corollary to Theorem 4.2.

Corollary 4.4. *Let A^* be a finite dimensional algorithm for (P) where (P) is a convex program. Then the following are equivalent:*

- (i) $A^*(N) \rightarrow X^*$, as $N \rightarrow \infty$.
- (ii) *For each $p \in Y$ and each nearest-point selection s_p defined by p , $s_p(A^*(N)) \rightarrow x_p^*$ as $N \rightarrow \infty$, where x_p^* is the unique point in X^* closest to p . \square*

In the next section, we will construct a finite dimensional algorithm A^* which, under a reachability condition, converges to the set of all infinite dimensional optimal solutions X^* . Hence, by the previous corollary, nearest-point finite dimensional solutions will converge to an optimal solution of (P).

Before leaving this section, we note that if (P) is a convex program with *strictly* convex objective function C , the optimal solution x^* will be uniquely attained. From Theorem 4.2, or as a corollary to Theorem 4.1, we obtain the following.

Theorem 4.5 (solution convergence). *Suppose (P) is a convex program with strictly convex objective function C , i.e., c_j is strictly convex on $\mathbb{R}^{n_1} \times \dots \times \mathbb{R}^{n_i}$, for some $j = 1, 2, \dots$. Then*

$$x^*(N) \rightarrow x^*,$$

where $x^*(N)$ is any solution to $(P(N))$, $N = 1, 2, \dots$, and x^* is the unique solution to (P). \square

The previous theorem extends a result established directly by McKenzie in [17].

5. Solution convergence through tie-breaking

In the previous section, we saw that tie-breaking by means of a nearest-point selection of solutions from a *convergent* finite dimensional algorithm generates a sequence of solutions which converges to an infinite dimensional optimal solution. Our objective in this section is to construct such a convergent finite dimensional algorithm.

We begin by assuming henceforth that (P) is *separable lower staircase*, i.e.,

$$c_j(y_1, \dots, y_j) = c_j(y_j), \quad j = 1, 2, \dots,$$

and

$$a_{ij}(y_1, \dots, y_j) = a_{ij}(y_j), \quad j = 1, \dots, i, \quad i = 1, 2, \dots,$$

with

$$a_{ij} = 0, \quad j = 1, \dots, i-2, \quad i = 3, 4, \dots$$

Under this assumption, (P) may be written in the form

$$\begin{aligned} \text{(P)} \quad & \text{minimize} \quad \sum_{j=1}^{\infty} c_j(y_j) \\ & \text{subject to} \quad a_{i,i-1}(y_{i-1}) + a_{i,i}(y_i) \geq b_i, \quad i = 1, 2, \dots, \\ & \quad \quad \quad y_j \in Y_j, \quad j = 1, 2, \dots, \end{aligned}$$

where a_{10} is identically zero.

In order to construct a finite dimensional algorithm $\{A^*(N), N = 1, 2, \dots\}$ which converges, for example, to X^* , we need by Theorem 4.2 to choose each $A^*(N)$ sufficiently large to insure that for each $x^* \in X^*$, there exists a choice $x^*(N)$ in $A^*(N)$, all N , such that $x^*(N) \rightarrow x^*$. On the other hand, we also need to be sure that each $A^*(N)$ is sufficiently small so as to insure that each convergent subsequence $\{x^*(N_k), k = 1, 2, \dots\}$ drawn from the $A^*(N)$ converges to a point x^* in X^* . The set $X^*(N)$ of all finite dimensional optimal solutions for (P(N)) is too small a choice for $A^*(N)$ in general. Instead, viewing N as a discrete time parameter, we enlarge $X^*(N)$ to include all solutions optimal to some feasible "state" ending period N . We begin by formalizing the notion of state.

Fix N . If y is feasible for (P(N)), then the only connection between the constraints of (P(N)) and the remaining constraints of (P) is the value $a_{N+1,N}(y_N)$. We call this value the *state* $s(N) \in \mathbb{R}^{m_{N+1}}$ associated with the solution y in $X(N)$. Let $s(N)$ be an arbitrary element of $\mathbb{R}^{m_{N+1}}$ and define

$$X(N, s(N)) = \{y \in X(N) : a_{N+1,N}(y_N) = s(N)\}.$$

Then $X(N, s(N))$ is the set of (P(N))-feasible solutions y having state $s(N)$ at stage N . Since the constraint functions are continuous, it is a closed, although possibly empty, subset of $X(N)$. Let

$$\begin{aligned} S_0(N) &= \{s(N) \in \mathbb{R}^{m_{N+1}} : X(N, s(N)) \neq \emptyset\} \\ &= \{s(N) \in \mathbb{R}^{m_{N+1}} : a_{N+1,N}(x_N) = s(N), \text{ for some } x(N) \in X(N)\} \end{aligned}$$

be the set of all states feasible for $(P(N))$ and let

$$S^*(N) = \{s(N) \in \mathbb{R}^{m_{N+1}}: a_{N+1,N}(x_N^*) = s(N), \text{ for some } x^* \in X^*\}$$

be the set of all states $s(N)$ optimal for (P) . We have $S^*(N) \subseteq S_0(N)$, all N . Moreover, one can easily show that, for each N , $S^*(N)$ and $S_0(N)$ are compact, non-empty subsets of $\mathbb{R}^{m_{N+1}}$.

For each N and each $s(N)$ in $S_0(N)$, consider the mathematical program $(P(N, s(N)))$ given by:

$$\begin{aligned} (P(N, s(N))) \quad & \text{minimize} \quad C(y(N); N) \\ & \text{subject to} \quad y(N) \in X(N, s(N)). \end{aligned}$$

Since the minimum $C^*(N, s(N))$ is attained, we may define

$$X^*(N, s(N)) = \{y(N) \in X(N, s(N)): C(y(N); N) = C^*(N, s(N))\},$$

which represents the set of all solutions optimal to state $s(N) \in S_0(N)$ for problem $(P(N))$. Each such $X^*(N, s(N))$ is a compact, non-empty subset of $X(N) \subseteq Y$, i.e., an element of $\mathcal{H}(Y)$.

For each N , and $S(N)$ satisfying $S^*(N) \subseteq S(N) \subseteq S_0(N)$, let

$$\chi^*(N, S(N)) = \bigcup_{s(N) \in S(N)} X^*(N, s(N)).$$

For technical reasons, we also define $\overline{\chi^*}(N, S(N))$ to be the closure of $\chi^*(N, S(N))$ in Y , so that $\overline{\chi^*}(N, S(N)) \in \mathcal{H}(Y)$, for all N .

The set $\chi^*(N, S(N))$, being the set of all solutions optimal to *some* feasible state in $S(N)$ for $(P(N))$, is called an *efficient set* (Ryan, Bean and Smith [18]). Note that

$$X^*(N) \subseteq \chi^*(N, S_0(N))$$

is a *strict* inclusion in general, since $\chi^*(N, S_0(N))$ is the union of the $X^*(N, s(N))$ over *all* feasible states $s(N) \in S_0(N)$, not just those states corresponding to optimal solutions of $(P(N))$.

Remark. One can show that for fixed N , the mapping $s(N) \rightarrow X(N, s(N))$ of $S_0(N)$ into $\mathcal{H}(Y)$ is upper semi-continuous [4]. If this mapping is also continuous (as for example when Y_N is discrete), then from the Maximum Theorem [4, p. 116], the resulting mapping $s(N) \rightarrow X^*(N, s(N))$ of *closed* $S(N)$ into $\mathcal{H}(Y)$ is upper semi-continuous. In this case, $\chi^*(N, S(N))$ is automatically closed, thus, eliminating the necessity of taking its closure.

We turn now to establishing conditions under which the efficient sets $\chi^*(N, S(N))$ converge to the set X^* of solutions optimal for (P) . By definition, it is necessary and sufficient to establish that

- (1) $X^* \subseteq \liminf \chi^*(N, S(N))$ and
- (2) $\limsup \chi^*(N, S(N)) \subseteq X^*$.

The next lemma establishes (1).

Lemma 5.1. *Let $S^*(N) \subseteq S(N) \subseteq S_0(N)$, for each N . Then*

$$X^* \subseteq \liminf \chi^*(N, S(N)).$$

Proof. Let $x^* \in X^*$, and thus $x^* \in X(N)$, all N . Fix N and define $s(N) = a_{N+1,N}(x_N^*)$, so that $s(N) \in S^*(N)$. We next verify that $x^* \in X^*(N, s(N))$, i.e., x^* is optimal for $(P(N, s(N)))$.

Suppose not. Then there exists x in $X^*(N, s(N))$ such that $x \neq x^*$ and $C(x; N) < C(x^*; N)$. For such x , we have

$$s(N) = a_{N+1,N}(x_N) = a_{N+1,N}(x_N^*).$$

Now define

$$z = (x_1, \dots, x_{N-1}, x_N, x_{N+1}^*, x_{N+2}^*, \dots).$$

Then

$$\begin{aligned} a_{N+1,N}(z_N) + a_{N+1,N+1}(z_{N+1}) &= a_{N+1,N}(x_N) + a_{N+1,N+1}(x_{N+1}^*) \\ &= a_{N+1,N}(x_N^*) + a_{N+1,N+1}(x_{N+1}^*) \\ &\geq b_{N+1}. \end{aligned}$$

Thus, z is feasible for (P) , i.e., $z \in X$. Moreover,

$$\begin{aligned} C(z) &= C(z; N) + \sum_{j=N+1}^{\infty} c_j(z_j) \\ &= C(x; N) + \sum_{j=N+1}^{\infty} c_j(x_j^*) \\ &< C(x^*; N) + \sum_{j=N+1}^{\infty} c_j(x_j^*) \\ &= C(x^*), \end{aligned}$$

i.e., x^* is not optimal. Contradiction. Hence, $x^* \in X^*(N, s(N))$, so that $x^* \in \chi^*(N, S^*(N)) \subseteq \chi^*(N, S(N))$. Since N is arbitrary, we have that

$$x^* \in \bigcap_{N=1}^{\infty} \chi^*(N, S(N)) \subseteq \liminf \chi^*(N, S(N)). \quad \square$$

In order to establish (2), we require that a state reachability property be satisfied.

Definition (reachability). Suppose $S^*(N) \subseteq S(N) \subseteq S_0(N)$, for each N . Let k be a positive integer and $s(k)$ a feasible state in $S_0(k)$. We say that the sequence $\{S(N), N = 1, 2, \dots\}$ is *reachable from state $s(k)$* at k if, given any sequence $\{t(N), N = 1, 2, \dots\}$ of feasible states with $t(N) \in S(N)$, all N , there exists $N_k > k$ sufficiently large such that for each $N \geq N_k$, there exists a sequence of decisions $y^N = (y_1^N, \dots, y_N^N, \dots)$ which

- (i) is feasible for $(P(N))$ and achieves state $t(N)$, i.e., $y^N \in X(N, t(N))$ and
- (ii) whose first k decisions are feasible for $(P(k))$ and achieve state $s(k)$, i.e., $y^N \in X(k, s(k))$.

We say that the sequence $\{S(N), N = 1, 2, \dots\}$ is *reachable from all feasible states* if it is reachable from all states $s(k) \in \underline{S}(k)$, for all k , where

$$\underline{S}(k) = \{s(k) \in \mathbb{R}^{m_{k+1}}: a_{k+1,k}(x_k) = s(k), \text{ for some } x \in X\}$$

is the set of states at k which are feasible for (P), $k = 1, 2, \dots$. Note that for each N , $S^*(N) \subseteq \underline{S}(N) \subseteq S_0(N)$, in general.

Roughly speaking, a sequence of feasible state sets is reachable if any sequence of states drawn from the sequence of sets can be eventually reached (i.e., attained) from any feasible state of (P) at any stage. This is a key property which decouples current and future decisions, allowing for finite dimensional approximation of an infinite dimensional problem. This decoupling effect is formalized in the following theorem.

Theorem 5.2. *Let $S^*(N) \subseteq S(N) \subseteq S_0(N)$, all N . If $\{S(N), N = 1, 2, \dots\}$ is reachable from all feasible states, then*

$$X = \lim_{N \rightarrow \infty} X(N, t(N)),$$

for all selections $\{t(N), N = 1, 2, \dots\}$ with $t(N) \in S(N)$, all N .

Proof. Since $X(N, t(N)) \subseteq X(N)$, all N , it follows that

$$\limsup X(N, t(N)) \subseteq \limsup X(N) = X,$$

since

$$X = \bigcap_{N=1}^{\infty} X(N) = \lim_{N \rightarrow \infty} X(N).$$

Now suppose (P) has the reachability property to $\{S(N), N = 1, 2, \dots\}$ from all feasible states and $t(N) \in S(N)$, all N . Let $z \in X$ and define $s(k) = a_{k+1,k}(z_k)$, $k = 1, 2, \dots$. Then $z \in X(k, s(k))$ and $s(k) \in \underline{S}(k)$, all k . By reachability for $k=1$ (and $s(1)$), there exists $N_1 > 1$ such that for all $N \geq N_1$, there exists $y^{1,N}$ which is (P(N))-feasible with state $t(N)$ and whose first decision is (P(1))-feasible with state $s(1)$. By reachability for $k=2$, there exists $N_2 > N_1$ such that for all $N \geq N_2$, there exists $y^{2,N}$ which is (P(N))-feasible with state $t(N)$ and whose first two decisions are (P(2))-feasible with state $s(2)$. For $N_1 \leq N < N_2$, define z^N in Y as follows:

$$z_j^N = \begin{cases} z_1, & j = 1, \\ y_j^{1,N}, & j = 2, \dots, N, \end{cases}$$

where z_j^N is chosen arbitrarily in Y_j , for $j > N$. We leave it to the reader to verify that z^N is (P(N))-feasible with state $t(N)$, i.e., $z^N \in X(N, t(N))$, $N_1 \leq N < N_2$.

Now suppose $1 < N_1 < N_2 < \dots < N_k$ have been found for $1, 2, \dots, k$ respectively by reachability and $z^{N_1}, \dots, z^{N_{k-1}}$ have been constructed so that $z^N \in X(N, t(N))$,

$N_1 \leq N < N_k - 1$ and $z_j^N = z_j$, for $j = 1, \dots, l$, $N_l \leq N < N_{l+1}$, $l = 1, \dots, k - 1$. Recall that N_k has the property that for each $N \geq N_k$, there exists $y^{k,N}$ which is $(P(N))$ -feasible with state $t(N)$ and whose first k decisions are $(P(k))$ -feasible with state $s(k)$. Applying reachability to $k + 1$, there exists $N_{k+1} > N_k$ such that for all $N \geq N_{k+1}$, there exists $y^{k+1,N}$ which is $(P(N))$ -feasible with state $t(N)$ and whose first $k + 1$ decisions are $(P(k + 1))$ -feasible with state $s(k + 1)$. For $N_k \leq N < N_{k+1}$, define z^N in Y as follows:

$$z_j^N = \begin{cases} z_j, & j = 1, \dots, k, \\ y_j^{k,N}, & j = k + 1, \dots, N, \end{cases}$$

where z_j^N is chosen arbitrarily in Y_j , for $j > N$. Once again, one can show that $z^N \in X(N, t(N))$, for $N_k \leq N < N_{k+1}$. In this way, we obtain a sequence $\{z^N\}$ in Y such that $z^N \in X(N, t(N))$, all $N \geq N_1$, and

$$z_j^N = z_j, \quad j = 1, \dots, l, \quad N_l \leq N < N_{l+1}, \quad l = 1, \dots, k - 1, \quad k = 1, 2, \dots$$

Therefore, $z^N \rightarrow z$, as $N \rightarrow \infty$, in the product topology of Y and hence, relative to the metric ρ as well (Lemma 2.1). Consequently, $z \in \liminf X(N, t(N))$. \square

Theorem 5.2 states that whenever a state sequence is reachable from any feasible state, the feasible decisions which reach those states will eventually cover the entire feasible region X .

We return to the task of showing, under reachability, that the efficient sets $\chi^*(N, S(N))$ converge to the set of optimal solutions X^* . The next lemma establishes the inclusion in (2) above.

Lemma 5.3. *Suppose $S^*(N) \subseteq S(N) \subseteq S_0(N)$, all N . If $\{S(N), N = 1, 2, \dots\}$ is reachable from all feasible states, then*

$$\limsup \chi^*(N, S(N)) \subseteq X^*.$$

Proof. Fix x in $\limsup \chi^*(N, S(N))$. Then there exists a sequence

$$1 < N_1 < N_2 < N_3 < \dots,$$

and a corresponding sequence $\{x^{N_i}\}$ such that $x^{N_i} \in \chi^*(N_i, S(N_i))$, all i , and $x^{N_i} \rightarrow x$, as $i \rightarrow \infty$. Then, for each i , there exists $t(N_i) \in S(N_i)$ such that $x^{N_i} \in X^*(N_i, t(N_i))$, so that $x \in \limsup_i X^*(N_i, t(N_i))$. For each N different from N_i , let $t(N) \in S(N)$ be arbitrary, so that we get $t(N) \in S(N)$, all N . Arguing as in the proof of Theorem 4.1, we may show that $\limsup X^*(N, t(N)) \subseteq X^*$. This completes the proof since

$$\limsup_i X^*(N_i, t(N_i)) \subseteq \limsup_N X^*(N, t(N)). \quad \square$$

We are now ready to prove our main result in this section.

Theorem 5.4 (efficient set convergence). *Suppose (P) is separable lower staircase. Let $S^*(N) \subseteq S(N) \subseteq S_0(N)$, for all N , and suppose $\{S(N), N = 1, 2, \dots\}$ is reachable from any feasible state. Then:*

(i) *Setting $A^*(N) = \overline{\chi^*(N, S(N))}$, all N , we have*

$$A^*(N) \rightarrow X^*, \quad \text{as } N \rightarrow \infty,$$

i.e., A^ is a convergent finite dimensional algorithm for (P).*

(ii) *For every p in $U(X^*)$, we have that*

$$s_p(\overline{\chi^*(N, S(N))}) \rightarrow s_p(X^*), \quad \text{as } N \rightarrow \infty,$$

where $s_p(\overline{\chi^(N, S(N))})$ is any point in $\overline{\chi^*(N, S(N))}$ nearest to p and $s_p(X^*)$ is the unique point in X^* nearest to p .*

Proof. By Lemmas 5.1 and 5.3, we have that $X^* = \lim \chi^*(N, S(N))$. Now apply Theorem 4.2 and [14, p. 337] to complete the proof. \square

Theorem 5.4 generalizes (to continuous variable mathematical programs) a result proven for discrete decision sets by Ryan, Bean and Smith [18]. For Y_j finite and uniformly bounded, all j , and β sufficiently small (where $\beta_j = \beta^j, 0 < \beta < 1$), they showed that the uniqueness set $U(K)$ for any K in $\mathcal{K}(Y)$ contains the origin. Moreover, letting p be the origin, they showed that the nearest point selection determined by p corresponds to the lexicographically minimum solution.

Since the uniqueness set of X^* is all of Y for convex programs, we obtain the following consequence of Theorem 5.4.

Corollary 5.5. *Suppose the mathematical program (P) in the preceding theorem is convex. Then, for any p in Y ,*

$$s_p(\overline{\chi^*(N, S(N))}) \rightarrow s_p(X^*), \quad \text{as } N \rightarrow \infty.$$

Proof. Apply Corollary 4.4. \square

The use of efficient set convergence in approximating solutions to (P) via solutions to the (P(N)) is complicated by two issues. The first is the task of computing $s_p(\overline{\chi^*(N, S(N))})$, which is a solution to a quadratic program. The difficulty lies in determining its feasible region, which is the efficient set $\overline{\chi^*(N, S(N))}$. What is required is a parametric right hand side solution of (P(N)). For the discrete case, as in Ryan, Bean and Smith [18], a dynamic programming solution of (P(N)) will automatically generate $\overline{\chi^*(N, S(N))} = \chi^*(N, S(N))$ for all $S^*(N) \subseteq S(N) \subseteq S_0(N)$. Also, for the continuous case with one-dimensional surplus states $s(N)$, as in inventory and production planning, the efficient sets are obtainable through conventional parametric programming. However, as the dimensionality of the surplus states

grows, the task of computing $\overline{\chi^*(N, S(N))}$ becomes more difficult, since multi-dimensional variation of the right hand side must be considered.

The second issue in implementing efficient set convergence in approximating solutions to (P) via those of (P(N)) involves the choice of N . Under the conditions of Theorem 5.4 or Corollary 5.5, a nearest point to p in $\overline{\chi^*(N, S(N))}$ will be arbitrarily close to the infinite dimensional optimal solution in X^* nearest to p for sufficiently large N . We next provide a stopping rule for discovering how large N must be to insure a desired level of accuracy.

The procedure described below is to solve a sequence of ever higher dimensional approximations to (P) until a stopping criterion is met that insures the desired level of accuracy. Specifically, using the finite dimensional algorithm A^* , employment of the stopping rule will approximate the first k decisions of the infinite dimensional optimal solution $s_p(X^*)$ within an error at most $\delta > 0$. The first part of the procedure is the stopping rule.

We begin by fixing $p \in U(X^*)$ and a corresponding nearest-point selection s_p . Let $S^*(N) \subseteq S(N) \subseteq S_0(N)$, all N . Let x_p^* denote the unique element of X^* such that $\rho(p, X^*) = \rho(p, x_p^*)$, i.e., $s_p(X^*) = x_p^*$. In what follows, it will be convenient to write $x_p^*(N, S(N))$ for $s_p(\overline{\chi^*(N, S(N))})$, all N . Recall that by Theorem 5.4, $x_p^*(N, S(N)) \rightarrow x_p^*$, as $N \rightarrow \infty$. Likewise, for $s(N) \in S(N)$, it will also be convenient to write $x_p(N, s(N))$ for $s_p(X^*(N, s(N)))$, $N = 1, 2, \dots$. Let k be a positive integer and $\delta > 0$. Assume A^* is the algorithm given in Theorem 5.4, i.e.,

$$A^*(N) = \overline{\chi^*(N, S(N))}, \quad \text{all } N.$$

Stopping Rule. Stop at N for algorithm A^* if $N \geq k$ and

$$d_j(x_p(N, s(N)), x_p^*(N, S(N)))_j \leq \delta, \quad j = 1, \dots, k,$$

for all $s(N) \in S(N)$.

The rule states that the algorithm A^* can be terminated when the nearest-point selections to all feasible states in $S(N)$ at N differ from the efficient set selection by at most δ in the first k decisions. Application of the stopping rule in practice requires exploitation of special structure in (P). For example, for discrete problems, $S(N)$ is finite so that there are only a finite number of selections to compute and compare for each N . Similarly, in the linear programming case there are finitely many candidate optimal bases.

Forward Algorithm.

Step 1. Choose k and δ . Set $N = k$.

Step 2. Solve (P(N)) using algorithm A^* to obtain $x_p(N, s(N))$, for all $s(N) \in S(N)$.

Step 3. If the stopping rule is not satisfied, set $N = N + 1$ and go to Step 2.

Step 4. Stop.

Theorem 5.6. *If the Forward Algorithm terminates at N , then $x_p^*(N, S(N))$ approximates x_p^* within error at most δ in the first k decisions, i.e.,*

$$d_j(x_p^*(N, S(N))_j, (x_p^*)_j) \leq \delta, \quad j = 1, \dots, k.$$

Proof. For each N , let $s^*(N) = a_{N+1, N}((x_p^*)_N)$ so that $s^*(N) \in S^*(N)$. By the triangle inequality, we have that

$$d_j(x_p^*(N, S(N))_j, (x_p^*)_j) \leq d_j(x_p^*(N, S(N))_j, x_p(N, s^*(N))_j) + d_j(x_p(N, s^*(N))_j, (x_p^*)_j), \quad j = 1, \dots, k.$$

By hypothesis, it follows that

$$d_j(x_p^*(N, S(N))_j, x_p(N, s^*(N))_j) \leq \delta, \quad j = 1, \dots, k,$$

since $S^*(N) \subseteq S(N)$. Thus, the proof will be complete if we show that $x_p(N, s^*(N))_j = (x_p^*)_j, j = 1, \dots, k$.

Suppose for the moment that

$$\sum_{j=1}^N \rho_j((x_p^*)_j, p_j)^2 < \sum_{j=1}^N \rho_j(x_p(N, s^*(N))_j, p_j)^2$$

Define

$$z_j = \begin{cases} (x_p^*)_j, & j = 1, \dots, N, \\ x_p(N, s^*(N))_j, & j = N + 1, \dots \end{cases}$$

Then z is $(P(N))$ -feasible with state $s^*(N)$, i.e., $z \in X(N, s^*(N))$. In fact, by the principle of optimality, z must also be optimal, i.e., $z \in X^*(N, s^*(N))$. Moreover, by our assumption,

$$\begin{aligned} \rho(z, p)^2 &= \sum_{j=1}^N \rho_j(z_j, p_j)^2 + \sum_{j=N+1}^{\infty} \rho_j(z_j, p_j)^2 \\ &= \sum_{j=1}^N \rho_j((x_p^*)_j, p_j)^2 + \sum_{j=N+1}^{\infty} \rho_j(x_p(N, s^*(N))_j, p_j)^2 \\ &< \sum_{j=1}^{\infty} \rho_j(x_p(N, s^*(N))_j, p_j)^2 \\ &= \rho(x_p(N, s^*(N)), p)^2, \end{aligned}$$

which implies that z in $X^*(N, s^*(N))$ is closer to p than $x_p(N, s^*(N))$. Contradiction. Hence, it must be that

$$\sum_{j=1}^N \rho_j((x_p^*)_j, p_j)^2 \geq \sum_{j=1}^N \rho_j(x_p(N, s^*(N))_j, p_j)^2.$$

Now define y by

$$y_j = \begin{cases} x_p(N, s^*(N))_j, & j = 1, \dots, N, \\ (x_p^*)_j, & j = N + 1, \dots \end{cases}$$

Since y has state $s^*(N)$ at N , it is easy to see that $y \in X$. Moreover, since $x_p^* \in X^*(N, s^*(N))$, we have that

$$\begin{aligned} C(y) &= \sum_{j=1}^N c_j(y_j) + \sum_{j=N+1}^{\infty} c_j(y_j) \\ &= \sum_{j=1}^N c_j(x_p(N, s^*(N))_j) + \sum_{j=N+1}^{\infty} c_j((x_p^*)_j) \\ &= \sum_{j=1}^N c_j((x_p^*)_j) + \sum_{j=N+1}^{\infty} c_j((x_p^*)_j) \\ &= C(x_p^*), \end{aligned}$$

i.e., y is optimal for (P), so that $y \in X^*$.

Finally, by above

$$\begin{aligned} \rho(y, p)^2 &= \sum_{j=1}^N \rho_j(y_j, p_j)^2 + \sum_{j=N+1}^{\infty} \rho_j(y_j, p_j)^2 \\ &= \sum_{j=1}^N \rho_j(x_p(N, s^*(N))_j, p_j)^2 + \sum_{j=N+1}^{\infty} \rho_j((x_p^*)_j, p_j)^2 \\ &\leq \sum_{j=1}^N \rho_j((x_p^*)_j, p_j)^2 + \sum_{j=N+1}^{\infty} \rho_j((x_p^*)_j, p_j)^2 \\ &= \rho(x_p^*, p)^2, \end{aligned}$$

which implies that $y = x_p^*$, since x_p^* is the unique point in X^* nearest to p . Consequently,

$$x_p(N, s^*(N))_j = (x_p^*)_j, \quad j = 1, \dots, N,$$

and the proof is complete, since $N \geq k$ by the Stopping Rule. \square

We turn next to the question of when the stopping criterion will ultimately be met.

Theorem 5.7. *The Forward Algorithm will finitely terminate for any choice of k and δ whenever*

$$X^*(N, s(N)) \rightarrow X^* \text{ in } \mathcal{H}(Y),$$

as $N \rightarrow \infty$, for all choices $s(N) \in S(N)$, $N = 1, 2, \dots$

Proof. If not, then the Stopping Rule is not satisfied eventually for some positive integer k and $\delta > 0$. Thus, for each $N \geq k$, there exists $s(N) \in S(N)$ and $j_N \in \{1, \dots, k\}$ such that

$$d_{j_N}(x_p(N, s(N))_{j_N}, x_p^*(N, S(N))_{j_N}) > \delta.$$

Since $\{1, \dots, k\}$ is a finite set, there exists $j_0 \in \{1, \dots, k\}$ for which $j_N = j_0$ infinitely often. Thus, there exists a subsequence $\{N_l\}$ of $\{N : N \geq k\}$ such that $N_l \geq k$ and

$$d_{j_0}(x_p(N_l, s(N_l))_{j_0}, x_p^*(N_l, S(N_l))_{j_0}) > \delta, \quad l = 1, 2, \dots$$

By Theorem 5.4, we have that $x_p^*(N, S(N)) \rightarrow x_p^*$, as $N \rightarrow \infty$, so that $x_p^*(N_l, S(N_l)) \rightarrow x_p^*$, as $l \rightarrow \infty$. Hence, by Lemma 2.1.,

$$x_p^*(N_l, S(N_l))_{j_0} \rightarrow (x_p^*)_{j_0}, \quad \text{as } l \rightarrow \infty.$$

By our hypothesis, $X^*(N, s(N)) \rightarrow X^*$ in $\mathcal{H}(Y)$, as $N \rightarrow \infty$. Thus, $X^*(N_l, s(N_l)) \rightarrow X^*$, and consequently $x_p(N_l, s(N_l)) \rightarrow x_p^*$, [19, Theorem 3.4] so that

$$x_p(N_l, s(N_l))_{j_0} \rightarrow (x_p^*)_{j_0}, \quad \text{as } l \rightarrow \infty.$$

Hence, by the triangle inequality,

$$d_{j_0}(x_p(N_l, s(N_l))_{j_0}, x_p^*(N_l, S(N_l))_{j_0}) \rightarrow 0, \quad \text{as } l \rightarrow \infty,$$

which is a contradiction. \square

If the convergence condition of Theorem 5.7 holds, then the Stopping Rule will be satisfied eventually. What remains is to give a sufficient condition for this convergence to take place.

Theorem 5.8. *Let $s(N) \in S(N)$, all N . If (P) has a unique solution, then $X^*(N, s(N)) \rightarrow X^*$ in $\mathcal{H}(Y)$, as $N \rightarrow \infty$ and hence the Forward Algorithm will finitely terminate for any choice of k and δ .*

Proof. By the proof of Lemma 5.3, we have that $\limsup X^*(N, s(N)) \subseteq X^*$. On the other hand, $X^* \subseteq \liminf X^*(N, s(N))$ by Corollary 2.2 of Schochetman and Smith [19]. \square

6. An application to production planning

The theoretical development of (P) in the previous sections proceeded from three major assumptions: separability, lower staircase structure and reachability. Sequential decision making over an unbounded horizon constitutes an important class of problems which typically meet these assumptions. In particular, the lower staircase structure can be obtained by introducing inventory-like surplus variables which summarize the effect of past decisions on current and future decisions. Examples include production and inventory planning and control, capacity expansion and equipment replacement. We will illustrate the previous development by an application to production planning.

Consider the problem of scheduling production to meet non-stationary, deterministic demand over an infinite horizon. The objective is to optimally balance the

scale of production against the cost of carrying inventory. The problem may be formulated by the following mathematical program [6, p. 87]:

$$\begin{aligned}
 \text{(Q)} \quad & \text{minimize} \quad \sum_{j=1}^{\infty} (k_j(P_j) + h_j(I_j)) \alpha^{j-1} \\
 & \text{subject to} \quad I_{i-1} + P_i - I_i = D_i, \quad i = 1, 2, \dots, \\
 & \quad \quad \quad 0 \leq P_j \leq P, \quad j = 1, 2, \dots, \\
 & \quad \quad \quad -B \leq I_j \leq I, \quad j = 1, 2, \dots,
 \end{aligned}$$

where I_j is the net inventory ending period j with $I_0 = 0$, P_j is the production in period j , D_j is the demand for production in period j , k_j is the production cost and h_j is the inventory holding/backorder cost for period j , $j = 1, 2, \dots$. The factor α is the discount factor reflecting the time value of money, where $0 < \alpha < 1$. We require that $P > 0$, $I > 0$, $B \geq 0$ and $P \geq D_j \geq 0$, all j . If $B > 0$, then backlogging is allowed. We impose the following assumptions on (Q).

Assumptions.

(I) For each j , the production cost k_j and inventory holding cost h_j are continuous, convex and uniformly bounded by an exponential function with rate at most $1/\alpha$, i.e., there exists $G > 0$ and $0 < \gamma < 1/\alpha$ such that

$$\max \left[\sup_{0 \leq P_j \leq P} k_j(P_j), \sup_{-B \leq I_j \leq I} h_j(I_j) \right] \leq G \gamma^j, \quad j = 1, 2, \dots$$

(II) $\liminf D_j < P$ and $\limsup D_j < 0$.

Assumption I provides for existence of an optimal solution and for flexibility in choosing p in Y for nearest-point selection. Assumption II is a regularity condition required to guarantee reachability for (Q).

Our first task is to show that (Q) may be reformulated as a special case of (P). For each j , define

$$y_j = \begin{bmatrix} I_j \\ P_j \end{bmatrix} \quad \text{and} \quad Y_j = [-B, I] \times [0, P].$$

Then $y_j \in \mathbb{R}^2$, all j , and the Y_j are the same compact, convex subset of \mathbb{R}^2 . The production planning problem (Q) may now be expressed in the form (P) as follows:

$$\begin{aligned}
 \text{(Q)} \quad & \text{minimize} \quad \sum_{j=1}^{\infty} c_j(y_j) \\
 & \text{subject to} \quad a_{i,i-1}(y_{i-1}) + a_{ii}(y_i) \geq b_i, \quad i = 1, 2, \dots, \\
 & \quad \quad \quad y_j \in Y_j, \quad j = 1, 2, \dots,
 \end{aligned}$$

where

$$c_j(y_j) = [k_j(P_j) + h_j(I_j)]\alpha^{j-1}, \quad j = 1, 2, \dots,$$

$$a_{ii}(y_i) = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} I_i \\ P_i \end{bmatrix}, \quad i = 1, 2, \dots,$$

$$a_{i,i-1}(y_{i-1}) = \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} I_{i-1} \\ P_{i-1} \end{bmatrix}, \quad i = 2, 3, \dots,$$

a_{i0} is identically zero and

$$b_i = \begin{bmatrix} D_i \\ -D_i \end{bmatrix}, \quad i = 1, 2, \dots$$

By construction the program (Q) is convex and separable lower staircase. Moreover, since

$$y_j = \begin{bmatrix} 0 \\ D_j \end{bmatrix}, \quad j = 1, 2, \dots,$$

yields a feasible solution for (Q), we have that the feasible region $X \neq \emptyset$.

For each $N = 1, 2, \dots$, the finite dimensional approximation (Q(N)) to (Q) becomes a finite horizon approximation given by

$$(Q(N)) \text{ minimize } \sum_{j=1}^N c_j(y_j)$$

subject to $a_{i,i-1}(y_{i-1}) + a_{ii}(y_i) \geq b_i, \quad i = 1, 2, \dots, N,$

$y \in Y.$

Since the inventory levels I_1, \dots, I_N are defined by the production schedule P_1, \dots, P_N via

$$I_j = \sum_{i=1}^j P_i - \sum_{i=1}^j D_i, \quad j = 1, 2, \dots, N,$$

we will say that (P_1, \dots, P_N, \dots) is feasible for (Q(N)) if $0 \leq P_j \leq P$, all j , and

$$-B + \sum_{i=1}^j D_i \leq \sum_{i=1}^j P_i \leq I + \sum_{i=1}^j D_i, \quad j = 1, \dots, N.$$

Suppose (P_1, \dots, P_N, \dots) is (Q(N))-feasible. The corresponding state $s(N)$ is then given by

$$s(N) = a_{N+1,N}(y_N) = \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} I_N \\ P_N \end{bmatrix} = \begin{bmatrix} I_N \\ -I_N \end{bmatrix},$$

i.e., $s(N)$ can be identified with the feasible inventory I_N at the end of period N . Moreover, if we let $X(N)$ denote the non-empty set of (Q(N))-feasible production schedules, i.e.,

$$X(N) = \left\{ (P_1, \dots, P_N, \dots) : 0 \leq P_j \leq P, \text{ all } j, \right. \\ \left. -B \leq \sum_{j=1}^n P_j - \sum_{j=1}^n D_j \leq I, 1 \leq n \leq N \right\},$$

then $X(N)$ is clearly a convex polytope. Also, since the mapping

$$(P_1, \dots, P_N, \dots) \rightarrow \sum_{j=1}^N P_j - \sum_{j=1}^N D_j = I_N$$

is continuous and affine on $X(N)$, its (non-empty) image is a compact, convex subset of $[-B, I]$. Thus, the set $S_0(N)$ of all $(Q(N))$ -feasible states may be identified with a closed subinterval of $[-B, I]$ i.e.,

$$S_0(N) = \left\{ \begin{bmatrix} x \\ -x \end{bmatrix} : -\underline{B} \leq x \leq \bar{I} \right\},$$

for some \underline{B} and \bar{I} satisfying $-B \leq -\underline{B} < \bar{I} \leq I$.

For the case of no backlogging, i.e., $B = 0$, and where either k_j or h_j is strictly increasing on its respective domain, $1 \leq j \leq N$, it is clear that if $(P_1^*, \dots, P_N^*, \dots)$ is optimal for $(Q(N))$, then the inventory I_N ending period N must be zero. It then follows that $X^*(N) \subseteq X^*(N, \theta)$, where $\theta = [0, 0]^t$. Since the opposite inclusion is also true, we have that $X^*(N) = X^*(N, \theta)$. Therefore, when it is feasible to obtain a strictly positive ending inventory for period N (e.g., when $B = 0, P = 2, I = 1, D_j = 1$, all j), we must have

$$X^*(N) \subset \chi^*(N, S_0(N)),$$

i.e., *strict* inclusion. Thus, in general, $\chi^*(N, S_0(N))$ is strictly larger than $X^*(N)$, all N .

We turn next to the problem of establishing reachability for (Q) .

Lemma 6.1. *Under Assumption II, we have*

(i) *There exists $0 \leq p < P$ and a subsequence $\{D_{j_m} : m = 1, 2, \dots\}$ of $\{D_j : j = 1, 2, \dots\}$ such that $D_{j_m} \leq p < P$, for all m .*

(ii) *There exists $0 < \delta \leq P$ and a subsequence $\{D_{i_n} : n = 1, 2, \dots\}$ of $\{D_i : i = 1, 2, \dots\}$ such that $0 < \delta \leq D_{i_n}$, all n .*

Remark. Assumption II is necessary as well as sufficient for (i) and (ii) to hold.

Proof. Omitted. \square

By Lemma 6.1, since $0 \leq D_j \leq P$, all j , it follows that inventory can be stockpiled by the amount $P - D_j \geq 0$ during periods of maximum production. In particular, this amount will be at least $\sigma = P - p > 0$ during the periods $j_m, m = 1, 2, \dots$. Analogously, inventory can be depleted by the amount D_i during periods of no production, where this amount will be at least δ during the periods $i_n, n = 1, 2, \dots$. This is enough to guarantee reachability as the next lemma asserts.

Lemma 6.2. *Suppose $S^*(N) \subseteq S(N) \subseteq S_0(N)$, all N . Then, under Assumption II, $\{S(N), N = 1, 2, \dots\}$ is reachable from all feasible states.*

Proof. See the Appendix. \square

Under Assumptions I and II, we then have from Theorem 5.4 that efficient set convergence takes place for (Q) and, moreover, from Corollary 5.5 that any nearest-point selection of points from the efficient sets will converge to an optimal solution of (Q). More specifically, let

$$P^*(N) = (P_1^*(N), \dots, P_N^*(N), \dots)$$

be the production schedule for (Q(N)) corresponding to a nearest selection to the origin θ in Y for the efficient set $\overline{\chi^*(N, S_0(N))}$, consisting of the closure of all optimal production schedules to all possible feasible inventories ending period N . Let

$$P^* = (P_1^*, P_2^*, \dots)$$

be the production schedule corresponding to the unique optimal solution to (Q) nearest θ .

Theorem 6.3. *Suppose (Q) satisfies Assumptions I and II. Then $P^*(N) \rightarrow P^*$, as $N \rightarrow \infty$, i.e.,*

$$P_j^*(N) \rightarrow P_j^*, \quad \text{as } N \rightarrow \infty,$$

for all $j = 1, 2, 3, \dots$ \square

Since the states, i.e., inventories, are one-dimensional, it is not difficult to calculate the efficient sets and hence, the nearest-point selections $P^*(N)$, $N = 1, 2, \dots$. The previous theorem assures us that these approximations to the infinite horizon optimal production schedule will converge to P^* as the horizon N lengthens. The Forward Algorithm of Section 5 may, moreover, be used to attempt discovery of a horizon N sufficiently far off to generate the desired accuracy in our estimate of the optimal first production decision P_1^* .

It is interesting to note that in the discrete case of no backlogging corresponding to $B = 0$, the monotonicity of the optimal production schedule $P^*(N)$ in the horizon N [6, p. 105] yields, for all selections, that $P_j^*(N) \rightarrow P_j^*$, as $N \rightarrow \infty$, for all j , and hence, $P^*(N) \rightarrow P^*$, as $N \rightarrow \infty$. One can show from this same monotonicity property that the efficient set point closest to the origin is the closest amongst those with zero ending inventory, i.e. the closest from the finite horizon optima $X^*(N)$. Therefore, the selection procedure reduces, in the no backordering case, to the classical result that $P^*(N) \rightarrow P^*$, as $N \rightarrow \infty$. However, in the presence of backordering, the monotonicity property no longer holds [6, p. 105]) and nearest-point selections are thus employed to ensure convergence.

Appendix

In this appendix, we give the proof of Lemma 6.2.

Proof of Lemma 6.2. We will show more generally that $\{S_0(N): N=1, 2, \dots\}$ is reachable from any state $s(k)$ in $S_0(k)$ at any period $k=1, 2, \dots$

Accordingly, suppose we are given period k with feasible ending inventory J and a sequence of feasible ending inventories $K_1, K_2, \dots, K_N, \dots$. Then there exists $(Q(k))$ -feasible (y_1, \dots, y_k, \dots) with $I_k = J$.

Suppose for the moment that $J = -B$ and we wish to reach inventory level I at some time in the future. By Lemma 6.1(i), there exists unique N'_k such that

$$-B + \sum_{j=k+1}^{k+N'_k} (P - D_j) \leq I,$$

while

$$-B + \sum_{j=k+1}^{k+N'_k+1} (P - D_j) > I.$$

Analogously, suppose that $J = I$ and we wish to reach inventory $-B$. By Lemma 6.1(ii), there exists N''_k such that

$$I - \sum_{j=k+1}^{k+N''_k} D_j \geq -B,$$

while

$$I - \sum_{j=k+1}^{k+N''_k+1} D_j < -B.$$

Now set $N_k = \max(k + N'_k + 1, k + N''_k + 1)$ and let $N \geq N_k$. We will show that we can feasibly reach from inventory level J at the end of period k to inventory level K_N at the end of period N . Recall that we have $Q(k)$ -feasible productions P_1, \dots, P_k, \dots with inventories satisfying $I_k = J$.

Fix $N \geq N_k$. There are three cases to consider: $J = K_N$, $J < K_N$ and $J > K_N$. If $J = K_N$, define $P_j = D_j$, $j = k+1, \dots, N$ and zero, say, thereafter. Then (P_1, \dots, P_N, \dots) is $(P(N))$ -feasible with ending inventory

$$I_N = \sum_{j=1}^N P_j - \sum_{j=1}^N D_j = \sum_{j=1}^k P_j - \sum_{j=1}^k D_j = J = K_N.$$

If $J < K_N$, then as above, there exists unique n' such that $k + n' \leq k + N'_k$, i.e., $k + n' + 1 \leq N$, and

$$J + \sum_{j=k+1}^{k+n'} (P - D_j) \leq K_N,$$

while

$$J + \sum_{j=k+1}^{k+n'+1} (P - D_j) > K_N.$$

Define

$$P_j = \begin{cases} P, & j = k + 1, \dots, k + n', \\ D_{k+n'+1} + K_N - J - \sum_{j=k+1}^{k+n'} (P - D_j), & j = k + n' + 1, \\ D_j, & j = k + n' + 2, \dots, N, \\ 0, & j > N. \end{cases}$$

Then (P_1, \dots, P_N, \dots) is $(P(N))$ -feasible with ending inventory

$$\begin{aligned} I_N &= \sum_{j=1}^N P_j - \sum_{j=1}^N D_j \\ &= \sum_{j=1}^{k+n'+1} P_j - \sum_{j=1}^{k+n'+1} D_j \\ &= J + \sum_{j=k+1}^{k+n'+1} (P_j - D_j) \\ &= J + \sum_{j=1}^{k+n'} (P - D_j) + P_{k+n'+1} - D_{k+n'+1} \\ &= K_N. \end{aligned}$$

Finally, if $J > K_N$, then there exists unique n'' such that $k + n'' \leq k + N''_k$, i.e., $k + n'' + 1 \leq N$, and

$$J - \sum_{j=k+1}^{k+n''} D_j \geq K_N,$$

while

$$J - \sum_{j=k+1}^{k+n''+1} D_j < K_N.$$

In this case, define

$$P_j = \begin{cases} 0, & j = k + 1, \dots, k + n'', \\ D_{k+n''+1} - J + \sum_{j=k+1}^{k+n''} D_j + K_N, & j = k + n'' + 1, \\ D_j, & j = k + n'' + 2, \dots, N, \\ 0, & j > N. \end{cases}$$

Then (P_1, \dots, P_N, \dots) is $(P(N))$ -feasible with ending inventory

$$\begin{aligned} I_N &= \sum_{j=1}^N P_j - \sum_{j=1}^N D_j \\ &= \sum_{j=1}^{k+n''+1} P_j - \sum_{j=1}^{k+n''+1} D_j \\ &= J - \sum_{j=k+1}^{k+n''} D_j + P_{k+n''+1} - D_{k+n''+1} \\ &= K_N, \end{aligned}$$

as required. \square

Acknowledgement

We are indebted to an anonymous referee for some helpful suggestions.

References

- [1] E.J. Anderson and P. Nash, *Linear Programming in Infinite Dimensional Spaces* (Wiley, New York, 1987).
- [2] J.-P. Aubin, *Applied Functional Analysis* (Wiley, New York, 1979).
- [3] J.C. Bean and R.L. Smith, "Conditions for the existence of planning horizons," *Mathematics of Operations Research* 9 (1984) 391-401.
- [4] C. Berge, *Topological Spaces* (Oliver and Boyd, London, 1963).
- [5] C. Bès and S.P. Sethi, "Concepts of forecast and decision horizons: applications to dynamic stochastic optimization problems," *Mathematics of Operations Research* 13 (1988) 295-310.
- [6] E.V. Denardo, *Dynamic Programming: Models and Applications* (Prentice-Hall, Englewood Cliffs, NY, 1982).
- [7] S.D. Flam and R.J.-B. Wets, "Existence results and finite horizon approximates for infinite horizon optimization problems," *Econometrica* 55 (1987) 1187-1209.
- [8] R. Grinold, "Infinite horizon programs," *Management Science* 18 (1971) 157-170.
- [9] R. Grinold, "Finite horizon approximations of infinite horizon linear programs," *Mathematical Programming* 12 (1977) 1-17.
- [10] R. Grinold, "Convex infinite horizon programs," *Mathematical Programming* 25 (1983) 64-82.
- [11] F. Hausdorff, *Set Theory* (Chelsea, New York, 1962, 2nd ed.).
- [12] P.C. Jones, J.L. Zydiak and W.J. Hopp, "Stationary dual prices and depreciation," *Mathematical Programming* 41 (1988) 357-366.
- [13] J.L. Kelley, *General Topology* (Van Nostrand, New York, 1955).
- [14] K. Kuratowski, *Topologie I* (Academic Press, New York, 1966).
- [15] K. Kuratowski, *Topologie II* (Academic Press, New York, 1968).
- [16] K. Kuratowski, *Introduction to Set Theory and Topology* (Pergamon Press, New York, 1972).
- [17] L.W. McKenzie, "Turnpike theory," *Econometrica* 44 (1976) 841-865.
- [18] S. Ryan, J.C. Bean and R.L. Smith, "A tie-breaking algorithm for discrete infinite horizon optimization," *Operations Research* 40, Supplement 2 (1992).
- [19] I.E. Schochetman and R.L. Smith, "Infinite horizon optimization," *Mathematics of Operations Research* 14 (1989) 1-16.
- [20] I.E. Schochetman and R.L. Smith, "Convergence of selections with applications in optimization," *Journal of Mathematical Analysis and Applications* 155 (1991) 278-292.