

GARY J. KOEHLER

A generalized Markov decision process

Revue française d'automatique, d'informatique et de recherche opérationnelle. Recherche opérationnelle, tome 14, n° 4 (1980), p. 349-354.

http://www.numdam.org/item?id=RO_1980__14_4_349_0

© AFCET, 1980, tous droits réservés.

L'accès aux archives de la revue « Revue française d'automatique, d'informatique et de recherche opérationnelle. Recherche opérationnelle » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/legal.php>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

A GENERALIZED MARKOV DECISION PROCESS (*)

by Gary J. KOEHLER ⁽¹⁾

Abstract. — In this paper we present a generalized Markov decision process that subsumes the traditional discounted, infinite horizon, finite state and action Markov decision process, Veinott's discounted decision processes, and Koehler's generalization of these two problem classes.

Résumé. — Nous présentons dans cet article un processus de Markov généralisé qui englobe le processus de décision markovien actualisé à l'horizon infini, avec état et action finis; les processus de décision actualisés de Veinott; et la généralisation de Koehler de ces deux classes de problèmes.

1. INTRODUCTION

In this paper we explore the extension of results obtained in [2] to a broader class of problems. The author's motivation is as follows. Many practical large-scale linear programs in energy and food allocation modelling can be easily partitioned to the form

$$\left. \begin{array}{l} \text{Max } c'x \\ \text{S.T.} \\ Bx = b \\ Ex = d \\ x \geq 0 \end{array} \right\} \quad (1.1)$$

where $b \geq 0$, B is essentially Leontief (to be defined later), and $Bx = b$ accounts for most of the constraints. Programs having Leontief constraint sets can be solved without inversion [2] and procedures taking advantage of this fact in solving (1.1) are of interest. In this paper we explore the results obtained in [2] on Leontief Systems and extend these to use on essentially Leontief Systems.

In the interest of brevity we borrow heavily from [2] in notation and preliminary results.

(*) Received September 1979.

(1) School of Industrial Engineering, Purdue University, West Lafayette, Indiana, U.S.A.

2. PROBLEM STATEMENT

Consider the problem

$$\left. \begin{array}{l} \text{Max } c'x \\ \text{S.T.} \\ Bx = b \\ x \geq 0 \end{array} \right\} \quad (2.1)$$

where B is an $m \times k$ essentially Leontief matrix and $b \geq 0$. An essentially Leontief matrix is a pre-Leontief matrix [4]—a matrix having at most one positive element per column—where, additionally, there is an $x \geq 0$ such that $Ax > 0$. We assume the following:

Assumption A

Problem 2.1 has a bounded objective and the columns of B are scaled so that the positive elements of B are not greater than one.

Let $A_i \subseteq \{1, 2, \dots, n\}$ for $i = 1, \dots, m$ such that if $B_{ij} > 0$ then $j \in A_i$ and $\bigcup_i A_i = \{1, 2, \dots, n\}$. Note that each $A_i \neq \emptyset$ since $Bx > 0$ for some

$x \geq 0$. Let $\Delta = \prod_{i=1}^m A_i$. For $\delta \in \Delta$ let B^δ be the corresponding submatrix of B and let $Q^\delta = I - B^\delta$ and $P_\delta = (Q^\delta)'$. Here $P_\delta \geq 0$.

From [1,4] there is some $\delta^* \in \Delta$ such that:

1. $v^* = [(B^{\delta^*})']^{-1} c^{\delta^*}$ solves the dual to (2.1);
2. $\rho(P_{\delta^*}) < 1$, where $\rho(P)$ is the spectral radius of the square matrix P .

Let

$$\mathcal{L}_\delta(v) = P_\delta v + c^\delta$$

and

$$\mathcal{L}(v) = \max_{\delta \in \Delta} \mathcal{L}_\delta(v).$$

Both operators are isotone. Define

$$F = \{v : v = \mathcal{L}(v)\} \quad \text{and} \quad C = \left\{v : \lim_{n \rightarrow \infty} \mathcal{L}^n(v) = v^*\right\}.$$

Note that $v^* \in F$ and $v^* \in C$. Finally, denote the dual feasible set of (2.1) as

$$D = \{v : (I - P_\delta)v \geq c^\delta, \delta \in \Delta\}.$$

Consider the following four conditions:

1. $P_\delta \geq 0, \quad \delta \in \Delta;$ (2.2)
2. $\rho(P_\delta) < 1$ for all $\delta \in \Delta;$
3. $D \neq \emptyset;$
4. $I - P_\delta$ has a positive diagonal for each $\delta \in \Delta.$

These conditions hold for discounted (semi-) Markov decisions and the discounted processes of Veinott [3]. Under these conditions $C = R^m$. In Koehler [2] condition 2 was relaxed to yield

$$2. \rho(P_\delta) < 1 \text{ for some } \delta \in \Delta. \tag{2.3}$$

1, 3 and 4 as in (2.2).

Here $C \neq R^m$ in general. In this paper we further relax (2.2) to

$$1, 2 \text{ and } 3 \text{ as in (2.3)}. \tag{2.4}$$

Drop 4.

Some properties of $\mathcal{L}(\cdot)$ and C which were discovered in [2] which also hold for (2.4) are summarized below. Refer to [2] for notation.

PROPOSITION 2.5

1. $\liminf_{n \rightarrow \infty} \mathcal{L}^n(v) \geq v^*$ for all $v \in R^m$.
2. $L(C) \subseteq C$.
3. C is convex.
4. If $F = [v^*]$, then $L(D) \subseteq C$. In addition, if for some $d > 0$ $P_\delta d \leq d$ for all $\delta \in \Delta$, then $C = R^m$.
5. If $\rho(P_\delta) \leq 1$ for all $\delta \in \Delta$ and $c^\delta \notin \text{range}(I - P_\delta)$ whenever $\rho(P_\delta) = 1$, then $F = \{v^*\}$.
6. Assume that $C = R^m$. Then for every $\delta \in \Delta$, $\rho(P_\delta) \leq 1$ and for δ having $\rho(P_\delta) = 1$ there exists no $v \in R^m$ such that $v \leq P_\delta v + c^\delta$.

Notice that the isotonicity of $\mathcal{L}(\cdot)$ and $L(v^*) \subseteq C$ give an easily satisfied sufficient condition for picking a point, v^0 , of C . That is, let $v^0 = -Me$ where $M \geq 0$ and e is an m vector of ones.

3. SPECTRAL PROPERTIES OF THE P_δ 's AND GEOMETRIC PROPERTIES OF D

In this section we extend almost all of the properties found in [2] which relate the spectral radii of the P_δ 's to boundedness of D and the property that it has a non-empty interior. From [2] we have:

PROPOSITION 3.1: *Consider the following properties of the $\rho(P_\delta)$'s and D :*

- (a) *For some $d > 0$, $P_\delta d \leq d$ for all $\delta \in \Delta$.*
- (b) *$\rho(P_\delta) \leq 1$ for all $\delta \in \Delta$.*
- (c) *D is unbounded from above.*
- (d) *For some $d \geq 0$, $P_\delta d \leq d$ for all $\delta \in \Delta$.*

Then (a) \rightarrow (b) \rightarrow (c) \Leftrightarrow (d).

In [2] it was shown, under the conditions given in (2.3), that if D has a non-empty interior then $\rho(P_\delta) = 1$ implies $c^\delta \notin \text{range}(I - P_\delta)$. This is no longer the case under the conditions in (2.4). For example, let

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

$$c' = (-2 \quad 2 \quad 0)$$

D has a non-empty interior yet for $\delta = (3,2)$ we have $P_\delta = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $\rho(P_\delta) = 1$ and $c^\delta = \begin{pmatrix} 0 \\ 2 \end{pmatrix} \in \text{range}(I - P_\delta)$.

The following results from [2] can be established under the conditions of (2.4).

THEOREM 3.2: *Suppose for some $d > 0$, $P_\delta d \leq d$ for every $\delta \in \Delta$ and if $\rho(P_\delta) = 1$ then $c^\delta \notin \text{range}(I - P_\delta)$. Then D has a non-empty interior.*

Proof: Suppose $B_j \leq 0$. Then $B_j = e_i - (-B_j + e_i)$. For some δ the i -th row of P_δ looks like $(e_i - B_j)'$. Thus $(e_i - B_j)' d \leq d_i$ by the first part of our premise for some $d > 0$. Thus $d_i - B_j' d \leq d_i$ or $0 \leq B_j' d$. However, $B_j' d < 0$.

Thus $B_j \leq 0$ implies $B_j = 0$.

Let $B_j = 0$ and δ be a decision using δ^* except in the state where B_j may be used. It is readily demonstrated that $\rho(P_\delta) = 1$. So $c^\delta \notin \text{range}(I - P_\delta)$ implies that $c_j < 0$ since otherwise v^* would give c^δ .

Thus we may restrict our attention to just those columns of B having a positive element and proceed to use the proof of Theorem 4.3 in [2] to get the desired result. \square

Notice that under the conditions of Theorem 3.2 $F = \{v^*\}$ [see Prop. 2.5, No. (5)].

4. THE IRREDUCIBLE CASE

Refer to [2] for appropriate definitions. In this section we assume:

Assumption B

B of problem (2.3) is not permutable to an essentially dynamic Leontief matrix.

Assumption B can be tested by discarding each non-positive column of B and using the procedures mentioned in [2] on the remaining matrix.

LEMMA 4.1: *If Assumption B holds, $d \geq 0$ and $P_\delta d \leq d$ for every $\delta \in \Delta$, then $d > 0$ and $B_j \leq 0$ implies $B_j = 0$.*

Proof: Without loss of generality we may partition B as $B = (B_1, B_2)$ with $B_2 \leq 0$ and B_1 Leontief. Then $d'(B_1, B_2) \geq 0$ gives that $d'B_2 = 0$. From [2] (Lemma 5.1) $d'B_1 \geq 0$ and B_1 not dynamic Leontief gives that $d > 0$. Hence, $d'B_2 = 0$ implies $B_2 = 0$. \square

From the Proposition 3.1 and Lemma 4.1 we get the following.

PROPOSITION 4.2: *If Assumption B holds, then the following are equivalent:*

- (a) D is unbounded from above;
- (b) $P_\delta d \leq d$ for some $d > 0$;
- (c) $\rho(P_\delta) \leq 1$ for all $\delta \in \Delta$.

Finally, pulling together several of the above results, we get:

THEOREM 4.3: *If Assumption B holds, then consider the following:*

- (a) $C = R^m$;
- (b) $\rho(P_\delta) \leq 1$ for all $\delta \in \Delta$ and $c^\delta \notin \text{range}(I - P_\delta)$ whenever $\rho(P_\delta) = 1$;
- (c) D is unbounded from above and has a non-empty interior.

Then (a) \Leftrightarrow (b) \rightarrow (c).

Proof: (a) \rightarrow (b): This follows from Proposition 2.5 part (6).

(b) \rightarrow (a): From Proposition 2.5 part (5) we have $F = \{v^*\}$. From Proposition 3.1 and Lemma 4.1 there is a $d > 0$ such that $P_\delta d \leq d$ for all $\delta \in \Delta$. Thus from Proposition 2.5 part (4) we have that $C = R^m$.

(b) \rightarrow (c): As in the previous part of the proof we have some $d > 0$ such that $P_\delta d \leq d$ for all $\delta \in \Delta$. By Proposition 3.1 D is unbounded from above. The rest follows from Proposition 4.2 and Theorem 3.2. \square

We cannot complete the implication in this theorem as can be seen by the following example. Let

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

$$c' = (1 \quad 1 \quad 0)$$

Here D has a non-empty interior and D has a positive direction of recession ($d' = (1,1)$) but

$$\left\{ \begin{pmatrix} 1 \\ \alpha \end{pmatrix}, \begin{pmatrix} \alpha \\ 1 \end{pmatrix} : \alpha \geq 1 \right\} \subseteq F,$$

which implies $C \neq R^m$.

Theorem 4.3 can be relaxed to complete the implications. This is done below.

THEOREM 4.4: *If Assumption B holds and if $B_j = 0$ implies that $c_j < 0$, then the conditions (a), (b), and (c) of Theorem 4.3 are equivalent.*

Proof: From Theorem 4.3 we have (a) \Leftrightarrow (b) \rightarrow (c).

(c) \rightarrow (a): If D is unbounded from above then from Proposition 3.1 and Lemma 4.1 there is a $d > 0$ such that $P_\delta d \leq d$ for all $\delta \in \Delta$ and $B_j \leq 0$ implies $B_j = 0$. By Proposition 4.2, $\rho(P_\delta) \leq 1$ for all $\delta \in \Delta$. Suppose $\rho(P_\delta) = 1$ for some $\delta \in \Delta$. From Theorem 4.3 of [2] $c^\delta \notin \text{range}(I - P_\delta)$ if some row has no positive element since then this row is a vector of zeroes and the corresponding element of c^δ is negative by assumption. Thus $c^\delta \notin \text{range}(I - P_\delta)$. The rest follows from Proposition 2.5 parts (5) and (4). \square

REFERENCES

1. R. W. COTTLE, and A. F. VEINOTT Jr., *Polyhedral Sets Having a Least Element*, Mathematical Programming, 3, 1972, pp. 238-249.
2. G. J. KOEHLER, *Value Convergence in a Generalized Markov Decision Process*, S.I.A.M. J. Optimization and Control, 17, 2, 1979, pp. 180-186.
3. A. F. VEINOTT Jr., *Discrete Dynamic Programming with Sensitive Discount Optimality Criteria*, Ann. Math. Stat., 40, 1969, pp. 1635-1660.
4. A. F. VEINOTT Jr., *Extreme Points of Leontief Substitution Systems*, Linear Algebra and Its Application, 1, 1968, pp. 181-194.