RAIRO. RECHERCHE OPÉRATIONNELLE

GARY J. KOEHLER

A complementarity approach for solving Leontief substitution systems and (generalized) Markov decision processes

RAIRO. Recherche opérationnelle, tome 13, nº 1 (1979), p. 75-80

<http://www.numdam.org/item?id=RO_1979__13_1_75_0>

R.A.I.R.O. Recherche opérationnelle/Operations Research
(vol. 13, n° 1, février 1979, p. 75 à 80)

A COMPLEMENTARITY APPROACH FOR SOLVING LEONTIEF SUBSTITUTION SYSTEMS AND (GENERALIZED) MARKOV DECISION PROCESSES (*)

by Gary J. KOEHLER [1]

Abstract. — *Leontief Substitution Systems and (generalized) Markov decision processes can be solved by value iteration, policy iteration, linear programming or hybrids of these. In this paper we present a new procedure based upon an equivalent complementarity formulation. This formulation can be solved by iterative methods as in value iteration.*

1. INTRODUCTION

In this paper we present a new algorithm for solving (generalized) Markov decision processes and Leontief Substitution Systems. Utilizing a result due to Cottle and Veinot [1] we formulate a nonstandard linear complementarity problem which is equivalent to the original problem. This resulting problem can be solved in a number of ways. In particular, Mangasarian's general class of iterative methods [5] for solving symmetric linear complementarity problems can be used.

2. NOTATION AND PRELIMINARY RESULTS

A matrix B is said to be Leontief [7] if it has exactly one positive element in each column and there is an $x \geqq 0$ such that $Bx > 0$. Consider the problem

$$\left.\begin{array}{c} \underset{\text{s. t.}}{\text{Max}} \quad c'x, \\ Bx = b, \\ x \geqq 0, \end{array}\right\} \qquad (2.1)$$

where B is an mxk Leontief matrix and $b > 0$. We impose the following on (2.1).

(*) Reçu mars 1978.

([1]) Department of Management, University of Florida, Gainesville.

R.A.I.R.O. Recherche opérationnelle/Operations Research, 0399/0842/1979/75/$ 1.00

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 $A_i = \{ j : B_{ij} > 0 \}$ for $i = 1, \ldots, m$ and $\Delta = \prod_{i=1}^{m} A_i$. For $\delta \in \Delta$ let B_δ be the corresponding submatrix of B and let $Q_\delta = I - B_\delta$ with $P_\delta = Q'_\delta$. Of course $P_\delta \geqq 0$.

Since the objective in (2.1) is bounded, we have that the problem has an optimal solution at an extreme point of the feasible set. By Veinott [7] this extreme point corresponds to a basis B_{δ^*} where $\delta^* \in \Delta$, $B_{\delta^*}^{-1}$ exists and is nonnegative and $p(P_{\delta^*}) < 1$ [$p(P_{\delta^*})$ is the spectral radius of P_{δ^*}].

Let $D = \{ y : B' y \geqq c \}$ represent the dual feasible set. From Cottle and Veinott [1] D has a least element v^* and $v^* = (B'_{\delta^*})^{-1} c_{\delta^*}$. Since $b \geqq 0$ and v^* is the least element of D, v^* solves the dual of (2.1).

Given Δ and P_δ and c_δ for each $\delta \in \Delta$ we refer to these as a generalized Markov decision process. Discounted Markov decision processes fit this format with each P_δ taking the form αR_δ where $0 < \alpha < 1$ and R_δ is stochastic. A more general notion of a discounted Markov decision process has $P_\delta \geqq 0$ and $p(P_\delta) < 1$ for all $\delta \in \Delta$ [6]. Both of these problems give rise to a totally Leontief system and these always have a bounded solution. The generalized Markov decision process we define here may have $p(P_\delta) \geqq 1$ for some $\delta \in \Delta$. We refer to [4] for several properties of this form of generalized Markov decision process and for properties of value iteration under this format.

We take as our goal in solving a generalized Markov decision process that of finding the least element of the fixed points of L where

$$L(v) = \operatorname*{Max}_{\delta \in \Delta} P_\delta v + c_\delta . \tag{2.2}$$

Since v^* is the least element of D and is also a fixed point of L and every fixed point of L must be in D, the determination of v^* is our goal.

3. SOLUTION METHODS

There are many methods available for solving (2.2) for the special case where $p(P_\delta) < 1$ for all $\delta \in \Delta$. Specifically, one can use value iteration, linear programming, policy iteration, or hybrids of these. For the general case value iteration is starting point dependent [4] and policy iteration may break down.

One can also formulate the problem of simultaneously finding v^* and a solution to (2.1) as a linear complementarity problem. Recently Eaves [2] has directly formulated the fixed point problem given in (2.2) as a complementary problem and provided an algorithm for finding v^*.

Below, we give a new linear complementarity formulation for finding v^*. The resulting problem can be solved by traditional linear complementarity methods, by quadratic programming methods, or by iterative methods. The method follows that given by Hildreth [3].

Let $v \in D$ and $z < v^*$. Clearly

$$0 < v^* - z \leqq v - z$$

(since v^* is the least element of D – the dual feasible set) so that we can find v^* by finding the minimum norm vector from z to D. This can be formulated as:

$$\underset{\text{s. t.}}{\text{Min}} \quad (v-z)'(v-z), \qquad B'v \geqq c.$$

After discarding constants and forming the associated saddle-point problem we get

$$\underset{\lambda \geqq 0}{\text{Max}} \; \underset{v}{\text{Min}} \, v'v - 2z'v - \lambda'(B'v - c).$$

Then the inner problem is solved by

$$v = z + \frac{B\lambda}{2},$$

and this gives an equivalent problem

$$\underset{\lambda \geqq 0}{\text{Max}} \, (c - B'z)'\lambda - \frac{\lambda' B' B \lambda}{4}. \tag{3.1}$$

This problem is a concave quadratic programming problem.

Let λ^* solve (2.3). Then

$$\lambda^{*\prime}(B'v^* - c) = 0$$

and

$$v^* = z + \frac{B\lambda^*}{2},$$

which implies $B\lambda^* > 0$. Note then that λ^* solves (2.1) for $b = B\lambda^*$. Let $\delta \in \Delta$ where $\lambda_\delta^* > 0$ and B_δ^{-1} exists and is non-negative (there will be at least one such δ [7]). Then δ is an optimal set of actions for the generalized Markov decision process and B_δ is an optimal basis for problem (2.1).

Now problem (3.1) can be solved as a quadratic program. One iterative method for doing this is the one variable at a time method. That is, guess λ^0. Compute λ^1 by holding constant all components of λ^0 except λ_i^0 and find the λ_i maximizing the resulting expression with $\lambda_i \geqq 0$. This is the procedure given by Hildreth [3]. When this method is applied cyclically one obtains a Gauss-Seidel type iteration (with projections to the non-negative orthant) on the system

$$B' B \lambda = 2(c - B' z).$$

Hildreth showed that, providing (3.1) has a solution, the resulting sequence converges to a solution of (3.1). Under our assumptions (3.1) always has a solution.

Problem (3.1) can also be cast into a more formal complementarity format. Here we want to find $\lambda \geqq 0$ such that

$$\left.\begin{array}{c} B' B \lambda - 2(c - B' z) \geqq 0, \\ \lambda' (B' B \lambda - 2(c - B' z)) = 0. \end{array}\right\} \tag{3.2}$$

Recently Mangasarian [5] has generalized many of the iterative procedures for solving symmetric linear complementarity problems. A symmetric linear complementarity problem is a problem of the form: Determine $\lambda \geqq 0$ where

$$\left.\begin{array}{c} M \lambda + q \geqq 0, \\ \lambda' (M \lambda + q) = 0, \end{array}\right\} \tag{3.3}$$

and M is symmetric. Problem (3.2) is a symmetric linear complementarity problem. Mangasarian's algorithm in its full generality is: Let $\lambda^0 \geqq 0$. Then

$$\lambda^{n+1} = \gamma (\lambda^n - \omega E^n (M \lambda^n + q + K^n (\lambda^{n+1} - \lambda^n)))_+ + (1 - \gamma) \lambda^n, \tag{3.4}$$

where $0 < \gamma \leqq 1$, $\omega > 0$, $\{E^n\}$ and $\{K^n\}$ are bounded sequences of real matrices with each E^n being a positive diagonal matrix with $E^n_{ii} > \alpha$ for some $\alpha > 0$ and for some $\tau > 0$:

$$y' ((\gamma \omega E^n)^{-1} + K^n - M/2) y \geqq \tau \| y \|^2,$$

for all n and y. Here $(\ )_+$ means the projection onto the non-negative orthant. Mangasarian showed that all cluster points of (3.4) solve (3.3). He gave some sufficient conditions for guaranteeing the existence of a cluster point. One was that M be copositive plus with some solution to either $M \lambda > 0$ or $M \lambda + q > 0$ (clearly the first condition implies the second.) A square matrix M is copositive plus if $\lambda \geqq 0$ implies $\lambda' M \lambda \geqq 0$ and $\lambda \geqq 0$ giving $\lambda' M \lambda = 0$ implies $M \lambda = 0$.

In applying algorithm (3.4) to problem (3.2) we have that $B'B$ is copositive plus. In the following two results we show when $B'B\lambda>0$ and $B'B\lambda-2(c-B'z)>0$ have solutions.

THEOREM 1: *$B'B\lambda>0$ has a solution if and only if $p(P_\delta)<1$ for all $\delta\in\Delta$.*

Before proving theorem 1 we note that Veinott [7] has shown that $p(P_\delta)<1$ for all $\delta\in\Delta$ if and only if $Bx=0$, $x\geqq 0$ has no non-trivial solution. Also $p(P_\delta)<1$ for all $\delta\in\Delta$ if and only if there is a $y\geqq 0$ such that $y'B>0$.

Proof of theorem 1: ($\Rightarrow$) If there is a λ giving $B'B\lambda>0$ then by duality there is no non-trivial solution to $B'B\lambda=0$, $\lambda\geqq 0$. Thus $p(P_\delta)<1$ for all $\delta\in\Delta$.

($\Leftarrow$) Suppose $p(P_\delta)<1$ for all $\delta\in\Delta$. Then there is a $y\geqq 0$ such that $y'B>0$. Let $\delta\in\Delta$ and $x_\delta=B_\delta^{-1}y$ and $x_{\bar\delta}=0$ where $\bar\delta$ is the set of indices not listed in δ. Then

$$B'Bx=B'B_\delta B_\delta^{-1}y=B'y>0. \quad \square$$

REMARK: Theorem 1 gives that a cluster point exists and gives some incentive for using the Mangasarian method for solving discounted Markov decisions processes.

For the following, recall that $D=\{v : B'v\geqq c\}$.

THEOREM 2: *$B'B\lambda-2(c-B'z)>0$ has a solution if and only if D has an interior.*

Proof: ($\Rightarrow$) Suppose $B'B\lambda-2(c-B'z)>0$ has a solution. Then let

$$v=z+\frac{B\lambda}{2}.$$

Then $B'B\lambda=2B'(v-z)$. Hence $B'B\lambda-2(c-B'z)=2(B'Bv-c)$. Thus $B'v>c$.

($\Leftarrow$) Let v be a solution to $B'v>c$. Then choose $\delta\in\Delta$ where B_δ is nonsingular (there is at least one such δ). Let $\lambda_\delta=2(B_\delta)^{-1}(v-z)$ and $\lambda_{\bar\lambda}=0$. Then

$$B'B\lambda-2(c-B'z)=2(B'v-c)>0. \quad \square$$

Note that if the conditions of theorem 1 are satisfied [i. e., $p(P_\delta)<1$ for all $\delta\in\Delta$] then D has an interior and the conditions of theorem 2 are satisfied [To *see* this merely add $B'y>0$ and $B'v^*\geqq c$ together. Then $B'(v^*+y)>c$.].The converse is not true as can be seen by the following example:

$$c'=(-2 \quad 1 \quad 0),$$

$$B=\begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 1 \end{bmatrix}.$$

Here $\Delta=\{(1,2), (1,3)\}$, $p(P_{(1,2)})=1$ but $v'=(1,2.5)$ gives $v'B=(-1.5, 1.5, 2.5)$.

4. CONCLUDING REMARKS

Finding an optimal basis of a Leontief Substitution System or finding v^* and an optimal set of activities for a generalized Markov decision process can be accomplished by solving (3.1). This can be solved in a number of ways. One iterative method which always works is Hildreth's procedure. Under rather mild conditions (i. e., that D has an interior), the family of algorithms proposed by Mangasarian can be used.

All the iterative methods for solving (3.1) work in the space of problem (2.1) rather than in its dual space (like value iteration and policy iteration). This presents a major drawback since one must maintain at least one working vector of size k where $k=\sum_{i=1}^{m}|A_i|$ and $k \geqq m$. Usually, $k \gg m$. Since value iteration requires at least one working vector of size m, the iterative methods suggested for solving (3.1) may not be as computationally attractive as methods for solving (2.2). However, with the large flexibility inherent in the class of algorithms given by Mangasarian this cannot be definitely answered now. Also when $k \approx m$, the procedure may be quite competitive with value iteration.

As a final question, what are the effects of the choice of z on the solution methodology? Usually (in value iteration [4]) one chooses $z=-M\,1$ where $M \gg 0$ and 1 is a vector of ones. This is done to insure that $z<v^*$.

ACKNOWLEDGEMENTS

The author would like to acknowledge the suggestion given by Professor Jean-Louis Goffin to investigate Hildreth's procedure.

REFERENCES

1. R. W. Cottle and A. F. Veinott, Jr., *Polyhedral Sets Having a Least Element*, Mathematical Programming, Vol. 3, 1972, pp. 238-249.
2. B. C. Eaves, *Complementary Pivot Theory and Markovian Decision Chains*, in Fixed Points, Algorithms and Applications, S. Karamardian, Éd., Academic Press, 1977, pp. 59-85.
3. C. Hildreth, *A Quadratic Programming Procedure*, Naval Research Logistic Quarterly, Vol. 4, 1957, pp. 79-85.
4. G. J. Koehler, *Value Convergence in a Generalized Markov Decision Process*, University of Florida, August, 1977 (forthcoming) S.I.A.M., Journal on Control and Optimization.
5. O. L. Mangasarian, *Solution of Symmetric Linear Complementarity Problems by Iterative Methods*, J. Optimization Theory and Applications, Vol. 22, No. 4, 1977, pp. 265-485.
6. E. L. Porteus, *Bounds and Transformations for Discounted Finite Markov Decision Chains*, Operations Research, Vol. 23, No. 4, 1975, pp. 761-784.
7. A. F. Veinott, Jr., *Extreme Points of Leontief Substitution Systems*, Linear Algebra and Its Applications, Vol. 1, 1968, pp. 181-194.