NOTES

A SOLUTION TO A COUNTABLE SYSTEM OF EQUATIONS ARISING IN MARKOVIAN DECISION PROCESSES[1]

By Cyrus Derman and Arthur F. Veinott, Jr.

Columbia University and Stanford University

Let $\{X_n\}$, $n = 0, 1, \cdots$, be a Markov chain having a state space consisting of the non-negative integers and having stationary transition probabilities $\{p_{ij}\}$. Let $\{w_1\}$, $i = 0, 1, \cdots$, be a sequence of real numbers. Consider the system of equations

$$(1) \qquad g + v_i = w_i + \textstyle\sum_{j=0}^{\infty} p_{ij}v_j, \qquad i = 0, 1, \cdots,$$

in the unknown variables $\{g, v_0, v_1, \cdots\}$. In [2], the system (1) arises in connection with conditions for the existence and construction of optimal rules for controlling a Markovian decision process. For a finite state space, existence of solutions to (1) is guaranteed by the condition that the Markov chain have at most one ergodic class of states. (See [3].) In this note we give conditions ensuring the existence (Theorem 1) and uniqueness (Theorem 2) of solutions to (1).

For $i, j, n = 0, 1, \cdots$, let

$$Z_n(j) = 1, \quad \text{if} \quad X_n = j \quad \text{and if} \quad X_m \neq 0 \quad \text{for} \quad 0 < m \leqq n$$

$$= 0, \quad \text{otherwise,}$$

$$_0p_{ij}^* = E(\textstyle\sum_{n=0}^{\infty} Z_n(j) \mid X_0 = \mathrm{i}), \quad \text{and} \quad m_{i0} = \textstyle\sum_{j=0}^{\infty} {_0p_{ij}^*}.$$

If the last series converges absolutely, then m_{i0} is the mean first passage time from i to 0 and we say m_{i0} is finite. If the m_{i0} are all finite, as we assume throughout, the state 0 is positive recurrent and there is only one recurrent class.

Let $Y_n = \sum_{j=0}^{\infty} w_j Z_n(j)$ and $c_{i0} = E(\sum_{n=0}^{\infty} Y_n \mid X_0 = i)$. By an obvious generalization of Theorem 5 in [1], p. 81, we get $c_{i0} = \sum_{j=0}^{\infty} {_0p_{ij}^*} w_j$ provided the series is absolutely convergent. If the series is absolutely convergent we say c_{i0} is finite. In applications w_i is often the cost incurred when in state i so c_{i0} is then the expected cost during a first passage from i to 0.

Theorem 1. (*Existence.*) *If the numbers m_{i0} and c_{i0}, $i = 0, 1, \cdots$, are finite, then the numbers*

$$(2) \qquad g = c_{00}/m_{00} \quad \text{and} \quad v_i = c_{i0} - gm_{i0}, \qquad i = 0, 1, \cdots,$$

satisfy (1) *and $\sum_{j=0}^{\infty} p_{ij}v_j$ converges absolutely, $i = 0, 1, \cdots$.*

Received 14 July 1966.

[1] Work supported by the Army, Navy, Air Force and NASA under contract Nonr-225(53) (NR-042-002) and by the Office of Naval Research under Contract Nonr-225(77) (NR-347-010). Reproduction in whole or in part is permitted for any purpose of the United States Government.

PROOF. *Let* $w_i^* = w_i - g$ and $Y_n^* = \sum_{j=0}^{\infty} w_j^* Z_n(j)$. Then for $i = 0, 1, \cdots$,

$$\begin{aligned} v_i &= E(\textstyle\sum_{n=0}^{\infty} Y_n^* \mid X_0 = i) \\ &= w_i^* + \textstyle\sum_{n=1}^{\infty} \sum_{j=0}^{\infty} E(Y_n^* \mid X_0 = i, X_1 = j) p_{ij} \\ &= w_i^* + \textstyle\sum_{j=0}^{\infty} \sum_{n=1}^{\infty} E(Y_n^* \mid X_0 = i, X_1 = j) p_{ij} \\ &= w_i^* + \textstyle\sum_{j=0}^{\infty} p_{ij} v_j \end{aligned}$$

so (1) holds. The interchange of expectation and summation is justified since the finiteness of the m_{i0} and c_{i0} imply that $\sum_{n=0}^{\infty} E(|Y_n^*| \mid X_0 = i) < \infty$. This in turn implies that the series above are absolutely convergent so the interchange of summations is also justified.

THEOREM 2. (*Uniqueness.*) *If the numbers* m_{i0} *and* c_{i0}, $i = 0, 1, \cdots$, *are finite, if* $\sum_{j=0}^{\infty} {}_0p_{ij}^*(c_{j0} - (c_{00}/m_{00})m_{j0})$, $i = 0, 1, \cdots$, *converges absolutely, and if* $\{g, v_0, v_1, \cdots\}$ *is a sequence with* $\sum_{j=0}^{\infty} {}_0p_{ij}^* v_j$, $i = 0, 1, \cdots$, *converging absolutely, then* $\{g, v_0, v_1, \cdots\}$ *satisfies* (1) *if and only if there is a real number r such that*

$$g = c_{00}/m_{00} \quad \textit{and} \quad v_i = c_{i0} - g m_{i0} + r, \qquad i = 0, 1, \cdots. \tag{3}$$

PROOF. It is immediate from the hypotheses and Theorem 1 that $\{g, v_0, v_1, \cdots\}$ defined in (3) satisfies (1) and $\sum_{j=0}^{\infty} {}_0p_{ij}^* v_j$ converges absolutely as well as $\sum_{j=0}^{\infty} p_{ij} v_j$. Let $\{g', v_0', v_1', \cdots\}$ be any other solution to (1) with $\sum_{j=0}^{\infty} {}_0p_{ij}^* v_j'$ converging absolutely for $i = 0, 1, \cdots$. Hence $\sum_{k=0}^{\infty} p_{ik} v_k'$ is absolutely convergent. Now premultiplying both sides of (1) by $\pi_i \equiv {}_0p_{0i}^*/m_{00}$, summing over $i = 0, 1, \cdots$, using the relations $\sum_{i=0}^{\infty} \pi_i = 1$ and $\pi_j = \sum_{k=0}^{\infty} p_{kj} \pi_k$, $j = 0, 1, \cdots$, and the fact that the interchange of summations is justified, we get $g' = \sum_{i=0}^{\infty} \pi_i w_i$ which is independent of $\{v_0', v_1', \cdots\}$. Thus since $\{g, v_0, v_1, \cdots\}$ satisfies (1) we must have $g = g'$.

Letting $\Delta_i = v_i' - v_i$, $i = 0, 1, \cdots$, we get from (1) subtracting one system from the other that

$$\Delta_i = \sum_{j=0}^{\infty} p_{ij} \Delta_j, \qquad i = 0, 1, \cdots. \tag{4}$$

Let $p_{ij}^n = \Pr(X_n = j \mid X_0 = i)$. Evidently for $N = 1, 2, \cdots$,

$$\sum_{n=1}^{N} p_{ij}^n \leqq {}_0p_{ij}^* + (N-1)\, {}_0p_{0j}^*, \qquad j = 0, 1, \cdots,$$

so that

$$N^{-1} \sum_{n=1}^{N} p_{ij}^n |\Delta_j| \leqq [{}_0p_{ij}^* + {}_0p_{0j}^*]|\Delta_j|, \qquad j = 0, 1, \cdots. \tag{5}$$

Since the series on the right side of (5) converges absolutely by hypothesis, and $\lim_{N\to\infty} N^{-1} \sum_{n=1}^{\infty} p_{ij}^n = \pi_j$, we get from the dominated convergence theorem that

$$\lim_{N\to\infty} \sum_{j=0}^{\infty} N^{-1} \sum_{n=1}^{\infty} p_{ij}^n \Delta_j = \sum_{j=0}^{\infty} \pi_j \Delta_j. \tag{6}$$

Since from (5), $\sum_{j=0}^{\infty} p_{ij}^n \Delta_j$ converges absolutely we can iterate (4), yielding

$$\Delta_i = \sum_{j=0}^{\infty} p_{ij}^n \Delta_j, \quad i = 0, 1, \cdots; n = 1, 2, \cdots. \tag{7}$$

Hence on substituting (7) into (6) $\Delta_i = \sum_{j=0}^{\infty} \pi_j \Delta_j$, $i = 0, 1, \cdots$. Thus Δ_i is independent of i, which completes the proof.

EXAMPLE. If the sequences $\{m_{i0}\}$ and $\{w_i\}$, $i = 0, 1, \cdots$, are bounded, then so is the sequence $\{c_{i0}\}$, $i = 0, 1, \cdots$, since $|c_{i0}| \leqq \sup_{k,j} m_{k0} |w_j|$. Thus Theorem 1 applies and, in addition, the solution to (1) given in (2) is bounded. This result is used in [2].

We remark that since

$$\sum_{j=0}^{\infty} {}_0p_{0j}^* |u| \geqq {}_0f_{0k} \sum_{j=0}^{\infty} {}_0p_{kj}^* |u_j|$$

where

$${}_0f_{0k} = \Pr\left(\sum_{n=0}^{\infty} Z_n(k) > 0 \mid X_0 = 0\right) > 0,$$

$\sum_{j=0}^{\infty} {}_0p_{kj}^* |u_j|$ is absolutely convergent for every recurrent state k provided that $\sum_{j=0}^{\infty} {}_0p_{0j}^* |u_j|$ is absolutely convergent. Thus under the assumption of one recurrent class the hypotheses of Theorems 1 and 2 could have been stated only for state 0 and the transient states.

REFERENCES

[1] CHUNG, K. L. (1960). *Markov Chains with Stationary Transition Probabilities*. Springer, Berlin.

[2] DERMAN, CYRUS (1966). Denumerable state Markovian decisio processes—average cost criterion. *Ann. Math. Statist.*) **37** 1545–1554

[3] HOWARD, RONALD (1960). *Dynamic Programming and Markov Processes*. Wiley, New York.