Sborník prací Přírodovědecké fakulty University Palackého v Olomouci. Matematika

Pavla Kunderová

On a mean reward from a Markov replacement process with only one isolated class of recurrent states

Sborník prací Přírodovědecké fakulty University Palackého v Olomouci. Matematika, Vol. 18 (1979), No. 1, 145--155

Persistent URL: `http://dml.cz/dmlcz/120076`

1979 — ACTA UNIVERSITATIS PALACKIANAE OLOMUCENSIS
FACULTAS RERUM NATURALIUM — TOM 61

Katedra matematické analýzy a numerické matematiky
přírodovědecké fakulty Univerzity Palackého v Olomouci
Vedoucí katedry: prof. RNDr. Miroslav Laitoch, CSc.

ON A MEAN REWARD FROM A MARKOV REPLACEMENT PROCESS WITH ONLY ONE ISOLATED CLASS OF RECURRENT STATES

PAVLA KUNDEROVÁ

(*Received March 31, 1978*)

Summary

The object of investigation in this paper is a Markov process with rewards under a stationary replacement policy as described in [3]. In Theorem 1 we derive a system of equations for establishing the mean reward from the process. The maximal reward is characterized by Theorem 2 and in its proof there is described the Howard's iteration method (see [1]) for finding the maximal reward and the corresponding optimal stationary replacement policy.

1. Basic definitions and notations

Let a homogeneous Markov process with rewards $\{X_t, t \geqq 0\}$ (see [3]) describing the evolution of a system in state space $I = \{1, 2, \dots, r\}$ be defined by exit intensities $(\mu(1), \dots, \mu(r))$, $0 < \mu(j) \leqq \infty$, $j = 1, \dots, r$ and by a stochastic matrix $\boldsymbol{P} = \| p(i,j) \|_{i,j=1}^r$, $p(i, i) = 0$ of transition probabilities in the moment of the exit. We constitute a matrix of the so called transition intensities $\boldsymbol{M} = \| \mu(i,j) \|_{i,j=1}^r$, where $\mu(i,j) = \mu(i)\, p(i,j)$ for $i \neq j$, $\mu(i, i) = -\mu(i)$,

$$-\mu(i, i) = \sum_{j \neq i} \mu(i, j) \tag{1}$$

The system being in state i at time t passes during the infinitesimal interval $(t, t + \mathrm{d}t)$ into state j with the probability $\mu(i, j)\, \mathrm{d}t$.

Consider a situation, where the development of the process can be influenced by an action called replacement (see [3]). Under a replacement of type $(i, +j)$ we mean the instantaneous shift of the system from state i into state j. The information on the evolution of the process up to the n-th state change is given by the sequence of states visited

$$i_0, i_1, i_2, \ldots, i_{n-1}, i_n = j, \tag{2}$$

by the corresponding sojourn times

$$t_0, t_1, t_2, \ldots, t_{n-1}, \tag{3}$$

and by the sequence

$$\delta_0, \delta_1, \delta_2, \ldots, \delta_{n-1}, \tag{4}$$

where $\delta_m = 0$ if the system was left i_m without interference and $\delta_m = 1$ if the passage from i_m into i_{m+1} was the result of a replacement.

For the history of the process up to the n-th state change we use the notation

$$\omega_n = [i_0, t_0, \delta_0; i_1, t_1, \delta_1; \ldots; i_{n-1}, t_{n-1}, \delta_{n-1}; i_n],$$

and the complete history of the process is given by a sequence

$$\omega = [i_0, t_0, \delta_0; i_1, t_1, \delta_1; \ldots].$$

A replacement policy (see [3]) is a decision, for all possible sequences (2)–(4) and all states j, on how long the system will be left in j without shifting (maximal sojourn time) and in what state it is to be shifted. Since we do not want to exclude the random choice of these quantities, we identify a replacement policy with a sequence of functions

$$F = \{{}^nF_k(t/\omega_n)\}, \qquad k = 1, 2, \ldots, r; n = 0, 1, 2, \ldots \tag{5}$$

${}^nF_k(t/\omega_n)$ is the probability that the maximal sojourn time in i_n will be less than t and that the eventual shift will be into $k \neq i_n$. We make

Assumption 1. *We consider only such replacement policies F where with probability 1*

a) there exists only a finite number of replacements in every finite interval,

b) there are not two or more replacements in the same moment.

According to the assumption to nearly every ω is assigned the trajectory $\{Y_t, t \geqq 0\}$, being not left continuous at time of the transition and not right continuous at time of the replacement. In what follows we denote by

$$\sigma_0, \sigma_1, \sigma_2, \ldots$$

the moments in which the trajectory is not continuous,

$$Y_t^- = Y_{t-}, t > 0; Y_0^- = Y_0; Y_t^+ = Y_{t+}, t \geqq 0,$$

E_j the mathematical expectation in a process without replacements under the condition $i_0 = j$,

D the set of couples $(i, +j)$ meaning admissible replacements,

$$D_i = \{j: (i, +j) \in D\}.$$

The reward from the process is defined by the following sets of numbers: $\varrho(i)$, $i \in I$ the reward per a time unit in state i; $r(i,j)$, $i, j \in I$ the reward from transition (i,j), we set $r(i,i) = 0$; $v(i,j)$, $i, j \in I$ the reward from the replacement $(i, +j)$, wə set $v(i,i) = 0$.

A stationary replacement policy f is given by a function $f(j)$ defined on a subset $I_f \subset I$ and taking values in I such that $f(j) \in D_j$ for $j \in I_f$, $f(j) \neq j$. The replacement policy f is the prescription to realize instantaneously the replacement $j \to f(j)$ whenever the transition in state j occurs. No replacements are made in states $j \notin I_f$. Let us make yet

Assumption 2. $(i, +j) \in D, (j, +k) \in D \Rightarrow (i, +k) \in D$ or $i = k$, $v(i,j) + v(j,k) \leqq v(i,k)$.

2. The mean reward per a time unit from the process with only one isolated class of recurrent states

Let R_T be the mean reward from the process up to the time T, in accordance with the previous definitions

$$R_T = \int_0^T \varrho(Y_t)\,\mathrm{d}t + \sum_{n=0}^{N} [r(Y_{\sigma_n}^-, Y_{\sigma_n}) + v(Y_{\sigma_n}, Y_{\sigma_n}^+)], \qquad \sigma_N \leqq T < \sigma_{N+1}.$$

In the sequel we use the statements (6)–(8), given in [4]: If the state space of the Markov process contains only one recurrent class (eventually the transient class), then there exists the mean reward per a time unit

$$\lim_{T\to\infty} \frac{1}{T} E_j(R_T) = \Theta, \qquad j = 1, 2, \ldots, r, \tag{6}$$

independent of j. Moreover, the limits

$$\lim_{T\to\infty} [E_j(R_T) - \Theta T] = w(j), \tag{7}$$

are finite and

$$\mu(j)\,w(j) + \Theta = \varrho(j) + \sum_{k\neq j} \mu(j,k)\,[r(j,k) + w(k)], \qquad j = 1, \ldots, r, \tag{8}$$

Θ is uniquely determined by (8), $w(1), \ldots, w(r)$ up to the additive constant.

For the rest of the paper we assume the state space of the process under arbitrary considered replacement policy to contain only one recurrent class and eventually the transient class. If $j \in I_f$ then (8) takes the form

$$\mu(j)\,w(j) + \Theta = \varrho(j) + \mu(j)\,[v(j, f(j)) + w(f(j))],$$

which being modified to include $\mu(j) = \infty$,

$$w(j) = v(j, f(j)) + w(f(j)).$$

If $j \notin I_f$ then from (8)

$$w(j) \sum_{k \neq j} \mu(j, k) + \Theta = \varrho(j) + \sum_{k \neq j} \mu(j, k) [r(j, k) + w(k)].$$

We have thus established the system of equations for determining the mean reward Θ

$$\begin{aligned} v(j, f(j)) + w(f(j)) - w(j) = 0, \qquad & j \in I_f, \\ \varrho(j) + \sum_{k \neq j} \mu(j, k) [r(j, k) + w(k) - w(j)] - \Theta = 0, \qquad & j \notin I_f. \end{aligned} \tag{9}$$

Theorem 1. *The system* (9) *determines* Θ *uniquely*, $w(1), \ldots, w(r)$ *up to the additive constant.*

Proof. a) For simplicity assume $I_f = \{1, \ldots, j - 1\}, j \leqq r$. Let $\boldsymbol{M}$ denote the matrix of the system (9). The matrix $\overline{\boldsymbol{M}}$ constructed from the matrix $\boldsymbol{M}$ by leaving the last column out, is the quasistochastic matrix of rank $r - 1$. Thus the system of r homogeneous equations with r unknowns

$$\boldsymbol{x}'\overline{\boldsymbol{M}} = \boldsymbol{0}, \qquad \boldsymbol{x}' = (x_1, \ldots, x_r)$$

has non-zero solutions forming the vector modulus of rank 1. It follows from [5], page 194 that every solution $\boldsymbol{x}$ except for multiplying of a constant is the stationary distribution of the process with the matrix of transition intensities $\overline{\boldsymbol{M}}$. Not all states $j, j + 1, \ldots, r$ in this process can be transient ones. Therefore the system

$$\boldsymbol{x}'\boldsymbol{M} = \boldsymbol{0}, \qquad \boldsymbol{x}' = (x_1, \ldots, x_r)$$

has a single zero solution, since with respect to the above the $(r + 1)$-th equation

$$-x_j - x_{j+1} - \ldots - x_r = 0$$

cannot be fulfilled otherwise. This implies that the rank of the matrix $\boldsymbol{M}$ equals to r.

b) Let us suppose $\bar{\Theta}, \bar{w}(1), \ldots, \bar{w}(r)$ to be another solution of (9). Subtracting (9) from the corresponding equations we obtain

$$\begin{aligned} w(f(j)) - \bar{w}(f(j)) - w(j) + \bar{w}(j) = 0, \qquad & j \in I_f, \\ \sum_{k \neq j} \mu(j, k) [w(k) - \bar{w}(k) - w(j) + \bar{w}(j)] - \Theta + \bar{\Theta} = 0, \qquad & j \notin I_f. \end{aligned} \tag{10}$$

If, say, $-\Theta + \bar{\Theta} < 0$, then for $j \notin I_f$ from (10)

$$\max_k [w(k) - \bar{w}(k)] > w(j) - \bar{w}(j).$$

As by Assumption 1 $f(j) \notin I_f$, this relation holds for all $j \in I$, which is a contradiction. We proceed analogously in showing the impossibility of $-\Theta + \bar{\Theta} > 0$.

Thus

$$\bar{\Theta} = \Theta.$$

Denoting $w(k) - \bar{w}(k) = \tilde{w}(k)$, $k \in I$, we get from (10)

$$\begin{aligned} &\tilde{w}(f(j)) - \tilde{w}(j) = 0, \qquad j \in I_f \\ &\sum_{k \neq j} \mu(j, k)\, \tilde{w}(k) + \mu(j, j)\, \tilde{w}(j) = 0, \qquad j \notin I_f. \end{aligned} \tag{11}$$

In a matrix notation for $x_j = \tilde{w}(j)$, $j \in I$, the system (11) has the form

$$\bar{\boldsymbol{M}}\boldsymbol{x}' = \boldsymbol{0}, \qquad \boldsymbol{x} = (x_1, \ldots, x_r).$$

The solution of the system (11) forms a vector modulus of rank 1 containing vectors having all components equal. That is

$$\tilde{w}(j) = w(j) - w'(j) = c, \qquad j \in I,$$

where c is an arbitrary constant. The proof is thus complete.

Let Θ_f denote the mean reward per a time unit from the replacement process under the stationary replacement policy f. Let us introduce the maximal reward

$$\bar{\Theta} = \max_f \{\Theta_f\}.$$

The replacement policy $\bar{f}$ is called optimal, if $\bar{\Theta} = \Theta_{\bar{f}}$. The maximal reward is characterized by Theorem 2.

Theorem 2. *$\bar{\Theta}$ is the single possible number to which $\bar{w}(1), \ldots, \bar{w}(r)$ is to find so that*

$$\begin{aligned} &\max \{v(j, k) + \bar{w}(k) - \bar{w}(j),\ k \in D_j;\ \varrho(j) + \\ &+ \sum_{k \neq j} \mu(j, k) [r(j, k) + \bar{w}(k) - \bar{w}(j)] - \bar{\Theta}\} = 0, \qquad j \in I. \end{aligned} \tag{12}$$

If $\bar{f}$ is such a replacement policy that the maximum in the compound bracket is reached for $j \in I_{\bar{f}}$ by the expression $v(j, \bar{f}(j)) + \bar{w}(\bar{f}(j)) - \bar{w}(j)$ and for $j \notin I_{\bar{f}}$ by the expression $\varrho(j) + \sum_{k \neq j} \mu(j, k) [r(j, k) + \bar{w}(k) - \bar{w}(j)] - \bar{\Theta}$, then $\bar{f}$ is the optimal stationary replacement policy.

Proof. We prove first the existence of the solution of (12) by the Howard's iteration procedure. Choosing an arbitrary stationary replacement policy f_0 we successively determine the stationary replacement policies $f_1, \ldots, f_n$ as follows:

a) We solve the system of equations

$$\begin{aligned} &v(j, f_n(j)) + w_n(f_n(j)) - w_n(j) = 0, \qquad j \in I_{f_n}, \\ &\varrho(j) + \sum_{k \neq j} \mu(j, k) [r(j, k) + w_n(k) - w_n(j)] - \Theta_n = 0, \qquad j \notin I_{f_n}. \end{aligned} \tag{13}$$

We put thereby $w_n(j_0) = 0$, where j_0 belongs to the (single as assumed) recurrent class I_1 with respect to the matrix $\boldsymbol{M}_n = \| \mu_n(i, j) \|_{i,j=1}^r$ of the transition intensities under the policy f_n. By Theorem 1 Θ_n is uniquely determined by the system (13), $w_n(1), \ldots, w_n(r)$ up to the additive constant.

b) For all $j \in I$ we successively determine

$$\max \{v(j, k) + w_n(k) - w_n(j), k \in D_j; \varrho(j) + \sum_{k \neq j} \mu(j, k) [r(j, k) + w_n(k) - w_n(j)] - \Theta_n\}$$

The policy f_{n+1} is determined as follows:
If the maximum for a fixed $j \in I$ is reached by the expression

$$\varrho(j) + \sum_{k \neq j} \mu(j, k) [r(j, k) + w_n(k) - w_n(j)] - \Theta_n,$$

we choose

$$j \notin I_{f_{n+1}}.$$

In the contrary, when the maximum is obtained by the expression

$$v(j, k) + w_n(k) - w_n(j) \qquad \text{for any } k \in D_j, \text{ we choose}$$

$$j \in I_{f_{n+1}}, \qquad f_{n+1}(j) = k.$$

Here the choice of $k = f_n(j)$ is preferred.

c) If the policy f_{n+1} does not possess the property required by Assumption 1, namely, that $f_{n+1}(j) \notin I_{f_{n+1}}$ for all $j \in I_{f_{n+1}}$, we change it to the policy f'_{n+1} as follows: in those states $j \in I_{f_{n+1}}$, where $f_{n+1}(j) \in I_{f_{n+1}}$, we take $f'_{n+1}(j) = f_{n+1}(f_{n+1}(j))$, in others $j \in I_{f_{n+1}}$ we have $f'_{n+1}(j) = f_{n+1}(j)$.

We shall now show the correctness of the procedure in c). Let us suppose that $f_n(j) \notin I_{f_n}$ for all $j \in I_{f_n}$ and the policy f_{n+1} being constructed in the above described way. Further let

$$j \in I_{f_{n+1}}, \qquad f_{n+1}(j) = k \in I_{f_{n+1}}, \qquad f_{n+1}(k) = k'. \tag{14}$$

By the construction of the replacement policy f_{n+1} it implies that

$$v(k, k') + w_n(k') - w_n(k) \geqq 0$$

and therefore by Assumption 2

$$v(j, k) + w_n(k) - w_n(j) \leqq v(j, k) + v(k, k') + w_n(k') - w_n(j) \leqq$$
$$\leqq v(j, k') + w_n(k') - w_n(j).$$

We see that the equality must hold here, because the expression $v(j, k) + w_n(k) - w_n(j)$ is maximal (replacement $j \to k$ under the policy f_{n+1} in the state j) from all expressions

$$v(j, i) + w_n(i) - w_n(j), \qquad i \in D_j.$$

We are thus led to the conclusion that k' is equivalent to k for the state j, moreover

$$v(k, k') + w_n(k') - w_n(k) = 0.$$

We can argue by contradiction that also $k \in I_{f_n}$, $k' = f_n(k)$. Therefore there cannot occur the situation

$$f_{n+1}(j) = k, \qquad f_{n+1}(k) = k', \qquad f_{n+1}(k') = k'',$$

and therefrom also

$$f_n(k) = k', \qquad f_n(k') = k'',$$

which contradicts the assumption of the replacement policy f_n. For this it suffices to change the constructed policy in the way described in c).

For thus constructed replacement policy f_{n+1} then

$$\begin{gathered} v(j, f_{n+1}(j)) + w_n(f_{n+1}(j)) - w_n(j) \geqq 0, \qquad j \in I_{f_{n+1}}, \\ \varrho(j) + \sum_{k \neq j} \mu(j, k) \left[r(j, k) + w_n(k) - w_n(j) \right] - \Theta_n \geqq 0, \qquad j \notin I_{f_{n+1}}. \end{gathered} \tag{15}$$

By Theorem 1

$$\begin{gathered} v(j, f_{n+1}(j)) + w_{n+1}(f_{n+1}(j)) - w_{n+1}(j) = 0, \qquad j \in I_{f_{n+1}}, \\ \varrho(j) + \sum_{k \neq j} \mu(j, k) \left[r(j, k) + w_{n+1}(k) - w_{n+1}(j) \right] - \Theta_{n+1} = 0, \qquad j \notin I_{f_{n+1}}. \end{gathered} \tag{16}$$

Subtracting (15) from (16) we obtain

$$\begin{gathered} w_{n+1}(f_{n+1}(j)) - w_n(f_{n+1}(j)) - w_{n+1}(j) + w_n(j) \leqq 0, \qquad j \in I_{f_{n+1}}, \\ \sum_{k \neq j} \mu(j, k) \left[w_{n+1}(k) - w_n(k) - w_{n+1}(j) + w_n(j) \right] - \Theta_{n+1} + \Theta_n \leqq 0, \qquad j \notin I_{f_{n+1}}. \end{gathered} \tag{17}$$

In analogy with the proof of Theorem 1 we can verify the impossibility of $\Theta_n - \Theta_{n+1} > 0$. Hence

$$\Theta_1 \leqq \Theta_2 \leqq \ldots \leqq \Theta_n \leqq \ldots$$

Because of only a finite number of the stationary replacement policies, there exists n_0 so that

$$\Theta_{n_0} = \Theta_{n_0+1} = \ldots$$

Let $\boldsymbol{P}_{n+1} = \| p_{n+1}(i, j) \|_{i,j=1}^{r}$ denote the matrix of the transition probabilities under policy f_{n+1}. Then $p_{n+1}(j, k) = p(j, k)$ for $j \notin I_{f_{n+1}}$. For $n \geqq n_0$ is from (17) for $j \in I_{f_{n+1}}$

$$\begin{aligned} w_n(j) - w_{n+1}(j) &= r(j) + w_n(f_{n+1}(j)) - w_{n+1}(f_{n+1}(j)) = \\ &= r(j) + \sum_{k \in I} p_{n+1}(j, k) \left[w_n(k) - w_{n+1}(k) \right], \qquad \text{where } r(j) \leqq 0. \end{aligned}$$

From (17) we have further for $j \notin I_{f_{n+1}}$, $n \geqq n_0$,

$$\sum_{k \neq j} \mu(j) \, p(j, k) \left[w_n(k) - w_{n+1}(k) \right] - \left[w_n(j) - w_{n+1}(j) \right] \mu(j) \geqq 0,$$

from where

$$\sum_{k \neq j} p(j, k) [w_n(k) - w_{n+1}(k)] + r(j) = w_n(j) - w_{n+1}(j), \qquad \text{where } r(j) \leqq 0.$$

Hence for all $j \in I$, $n \geqq n_0$,

$$w_n(j) - w_{n+1}(j) = r(j) + \sum_{k \in I} p_{n+1}(j, k) [w_n(k) - w_{n+1}(k)], \qquad r(j) \leqq 0. \tag{18}$$

We now show that it follows from (18)

$$w_n(j) \leqq w_{n+1}(j), \qquad j \in I, n \geqq n_0. \tag{19}$$

Introduce the shortened notation

$$\| p_{n+1}(i, j) \|_{i,j=1}^r = \| p(i, j) \|_{i,j=1}^r = \mathbf{P}, \qquad \mathbf{P}^n = \| p^{(n)}(i, j) \|_{i,j=1}^r,$$

$\mathbf{P}^0$ the unit matrix. Let I_1 be the isolated class of recurrent states, I' the transient class with respect to the matrix $\mathbf{P}$. Let us write next

$$\pi(j) = \lim_{n \to \infty} p^{(n)}(i, j).$$

Then $\pi(j) > 0$ for $j \in I_1$, $\pi(j) = 0$ for $j \in I'$.

Multiplying (18) by the numbers $\pi(j)$, adding for all $j \in I$ and applying the relation

$$\pi(k) = \sum_j \pi(j) p(j, k)$$

we obtain

$$\sum_j \pi(j) [w_n(j) - w_{n+1}(j)] = \sum_j \pi(j) r(j) + \sum_k \pi(k) [w_n(k) - w_{n+1}(k)].$$

We see now that there must be $r(j) = 0$ for $j \in I_1$.
Then (18) yields

$$w_n(j) - w_{n+1}(j) = \sum_{k \in I} p(j, k) [w_n(k) - w_{n+1}(k)], \qquad j \in I_1,$$

and on successive substitution we come to

$$w_n(j) - w_{n+1}(j) = \sum_k p^{(m)}(j, k) [w_n(k) - w_{n+1}(k)], \qquad j \in I_1, m = 1, 2, \ldots$$

Letting $m \to \infty$ yields

$$w_n(j) - w_{n+1}(j) = \sum_k \pi(k) [w_n(k) - w_{n+1}(k)] = \text{constant}, \qquad j \in I_1.$$

Further $j_0 \in I_1$ and $w_n(j_0) = w_{n+1}(j_0) = 0$, from where

$$w_n(j) = w_{n+1}(j) \qquad \text{for } j \in I_1. \tag{20}$$

It is obvious from (18) and (20) that we can write for $j \in I'$

$$w_n(j) - w_{n+1}(j) = r(j) + \sum_{k \in I'} p(j, k) [w_n(k) - w_{n+1}(k)].$$

Then

$$w_n(j) - w_{n+1}(j) = r(j) + \sum_{k \in I'} p(j, k) r(k) + \sum_{k \in I'} p^{(2)}(j, k) [w_n(k) - w_{n+1}(k)],$$

and in the same manner further gives

$$w_n(j) - w_{n+1}(j) = \sum_{m=0}^{N} \sum_{k \in I'} p^{(m)}(j, k) r(k) + \sum_{k \in I'} p^{(N+1)}(j, k) [w_n(k) - w_{n+1}(k)].$$

If k is the transient state of the chain formed by the states of the process considered, then the serie $\sum_{n=0}^{\infty} p^{(n)}(j, k)$ converges for $j \in I$ (see [2]).

Because of this statement

$$w_n(j) - w_{n+1}(j) = \sum_{m=0}^{\infty} \sum_{k \in I'} p^{(m)}(j, k) r(k) \leqq 0 \qquad \text{for } j \in I'. \tag{21}$$

(21) together with (20) gives (19).

It follows from (19) and from the finiteness of the set of the stationary replacement policies that $m \geqq n_0$ can be found where

$$w_m(j) = w_{m+1}(j), \qquad j \in I.$$

Since $\Theta_m = \Theta_{m+1}$, we can write by using (13) and the method of determining the policy f_{m+1}

1. for $j \in I_{f_{m+1}}$

$$\begin{aligned} &\max \{v(j, k) + w_m(k) - w_m(j),\ k \in D_j;\ \varrho(j) + \\ &+ \sum_{k \neq j} \mu(j, k) [r(j, k) + w_m(k) - w_m(j)] - \Theta_m\} = \\ &= v(j, f_{m+1}(j)) + w_m(f_{m+1}(j)) - w_m(j) = \\ &= v(j, f_{m+1}(j)) + w_{m+1}(f_{m+1}(j)) - w_{m+1}(j) = 0. \end{aligned}$$

2. for $j \notin I_{f_{m+1}}$

$$\begin{aligned} &\max \{v(j, k) + w_m(k) - w_m(j),\ k \in D_j;\ \varrho(j) + \\ &+ \sum_{k \neq j} \mu(j, k) [r(j, k) + w_m(k) - w_m(j)] - \Theta_m\} = \\ &= \varrho(j) + \sum_{k \neq j} \mu(j, k) [r(j, k) + w_m(k) - w_m(j)] - \Theta_m = \\ &= \varrho(j) + \sum_{k \neq j} \mu(j, k) [r(j, k) + w_{m+1}(k) - w_{m+1}(j)] - \Theta_{m+1} = 0. \end{aligned}$$

We see that

$$\bar{\Theta} = \Theta_m, \qquad \bar{w}(j) = w_m(j), \qquad j \in I,$$

is the solution of equation (12).

We verify now that (12) determines $\bar{\Theta}$ uniquely.

Assume there exists $\Theta \neq \bar{\Theta}$ establishing (12), that is

$$\max \{v(j, k) + w(k) - w(j), k \in D_j;\ \varrho(j) + \\ + \sum_{k \neq j} \mu(j, k) [r(j, k) + w(k) - w(j)] - \Theta\} = 0, \qquad j \in I. \tag{22}$$

Let, say, $\bar{\Theta} - \Theta > 0$ and let $\bar{f}$ be the replacement policy defined in Theorem 2. Then

$$v(j, \bar{f}(j)) + \bar{w}(\bar{f}(j)) - \bar{w}(j) = 0, \qquad j \in I_{\bar{f}}, \tag{23}$$
$$\varrho(j) + \sum_{k \neq j} \mu(j, k) [r(j, k) + \bar{w}(k) - \bar{w}(j)] - \bar{\Theta} = 0, \qquad j \notin I_{\bar{f}}.$$

According to (22)

$$v(j, \bar{f}(j)) + w(\bar{f}(j)) - w(i) \leqq 0, \qquad j \in I_{\bar{f}},$$
$$\varrho(j) + \sum_{k \neq j} \mu(j, k) [r(j, k) + w(k) - w(j)] - \Theta \leqq 0, \qquad j \notin I_{\bar{f}}.$$

Subtracting this from (23) we obtain

$$\bar{w}(\bar{f}(j)) - w(\bar{f}(j)) - \bar{w}(j) + w(j) \geqq 0, \qquad j \in I_{\bar{f}},$$
$$\sum_{k \neq j} \mu(j, k) [\bar{w}(k) - w(k) - \bar{w}(j) + w(j)] - \bar{\Theta} + \Theta \geqq 0, \qquad j \notin I_{\bar{f}}.$$

From this we deduce a contradiction

$$\max_k [\bar{w}(k) - w(k)] > \bar{w}(j) - w(j), \qquad j \notin I_{\bar{f}},$$

since $\bar{f}(j) \notin I_{\bar{f}}$, it is

$$\max_k [\bar{w}(k) - w(k)] > \bar{w}(\bar{f}(j)) - w(\bar{f}(j)) \geqq \bar{w}(j) - w(j), \qquad j \in I_{\bar{f}},$$

which refutes $\bar{\Theta} - \Theta > 0$. We proceed analogously in disproving $\bar{\Theta} - \Theta < 0$, hence $\bar{\Theta} = \Theta$.

It is obvious that for an arbitrary stationary replacement policy f is $\Theta_f \leqq \bar{\Theta}$ because in Howard's iteration method we started from an arbitrary stationary replacement policy and the mean rewards $\Theta_n = \Theta_{f_n}$ constructed a non-decreasing succesion. It still remains to verify that the policy $\bar{f}$ is an optimal stationary one.

Theorem 1 tells us that the system

$$v(j, \bar{f}(j)) + w(\bar{f}(j)) - w(j) = 0, \qquad j \in I_{\bar{f}},$$
$$\varrho(j) + \sum_{k \neq j} \mu(j, k) [r(j, k) + w(k) - w(j)] - \Theta_{\bar{f}} = 0, \qquad j \notin J_{\bar{f}}, \tag{24}$$

determines uniquely the reward $\Theta_{\bar{f}}$. Comparing (23) and (24) we obtain $\Theta_{\bar{f}} = \bar{\Theta}$.

Finally, I should like to express my gratitude to dr. P. Mandl, DrSc., for providing me with valuable expert advice and helpful criticism in writting this article.

REFERENCES

[1] R. A. Howard: *Dynamic programming and Markov processes,* New York—London 1960.

[2] P. Mandl: *Řízené Markovovy řetězce.* Příloha časopisu Kybernetika, roč. 5, 1969, Academia Praha.

[3] P. Mandl: *Some results on Markovian replacement processes.* J. Appl. Prob. 8, 357—365 (1971).

[4] P. Mandl: *An identity for Markovian replacement processes.* J. Appl. Prob. 6, 348—354 (1969).

[5] T. A. Sarymsakov: *Osnovy teorii processov Markova.* Moskva 1954.

Souhrn

PRŮMĚRNÝ VÝNOS Z MARKOVA PROCESU S OBNOVAMI S JEDINOU IZOLOVANOU TŘÍDOU REKURENTNÍCH STAVŮ

PAVLA KUNDEROVÁ

Uvažuje se Markovův proces s obnovami popsaný v [3] se stacionární strategií obnovy. Ve větě 1 je odvozena soustava rovnic pro určování průměrného výnosu z procesu. Maximální výnos je charakterisován větou 2, v jejímž důkaze je popsána Howardova iterační metoda (viz [1]) nacházení maximálního výnosu a odpovídající optimální stacionární strategie.

Резюме

СРЕДНИЙ ДОХОД ИЗ ПРОЦЕССА МАРКОВА С ВОССТАНОВЛЕНИЯМИ С ЕДИНСТВЕННЫМ КЛАССОМ ВОЗВРАТНЫХ СОСТОЯНИЙ

РАВЛА КУНДЕРОВА

В работе рассмотрен процесс Маркова с восстановлениями определённый в [3] при использовании стационарной стратегии восстановления. В теореме 1 введена система уравнений для определения среднего дохода за единицу времени. Максимальный доход характеризуется теоремой 2, в доказательстве которой описан итерационный метод Ховарда для нахождения максимального дохода и отвечающей оптимальной стационарной стратегии.