ACTA UNIVERSITATIS PALACKIANAE OLOMUCENSIS
FACULTAS RERUM NATURALIUM
1985 MATHEMATICA XXIV Vol. 82

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

THE EXPECTED DISCOUNTED REWARD FROM A MARKOV REPLACEMENT PROCESS

PAVLA KUNDEROVÁ

(*Received March 21, 1984*)

1. Basic definitions and notations

Let a homogeneous Markov process with rewards $\{X_t, t \geqq 0\}$ describing the evolution of a system in state space $I = \{1, 2, \ldots, r\}$ be defined by exit intensities $(\mu(1), \ldots, \mu(r))$, $0 < \mu(j) \leqq \infty$, $j = 1, \ldots, r$, and by a stochastic matrix $\mathbf{P} = \| p(i,j) \|_{i,j=1}^{r}$, $p(i,i) = 0$ of transition probabilities in the moment of exit. We constitute a matrix of so called transition intensities $\mathbf{M} = \| \mu(i,j) \|_{i,j=1}^{r}$, 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 [2]. Under a *replacement of type* $(i, +j)$ we mean the instantaneous shift of the system from state i into state j. The information of the evolution of the process up to the n-th state change is given by the sequence of states visited

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

by the corresponding sojourn times

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

and by the sequence

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

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 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 we note the complete history of the process (according to [2])

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

A replacement policy (see [2]) 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 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

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

where ${}^nF_k(t/\omega_n)$ 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

$\sigma_0 = 0, \sigma_1, \sigma_2, \ldots$ 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 $i_0 = j$,

E_j^F the mathematical expectation in a replacement process under the replacement policy F and under the condition $i_0 = j$,

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

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

The reward from the process (see [2]) 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 $(i, +j)$, we set $v(i,i) = 0$.

A stationary replacement policy f is given by 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 $j \in I_f$ occurs. No replacements are made in states $j \notin I_f$.

Let us make yet

Assumption 2.

$$(i, +j) \in D, (j, +k) \in D \Rightarrow (i, +k) \in D \text{ or } i = k,$$
$$v(i,j) + v(j,k) \leqq v(i,k).$$

2. The expected discounted reward from the process

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

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

The Laplace–Stieltjes transform

$$R = \int_0^\infty e^{-\lambda T}\,dR(T), \qquad \lambda > 0$$

is *the discounting of the reward,* λ is so called discount factor (see [3]).

In the sequel we use the following statement given in [2], page 349, formula (7): For $\lambda > 0$ holds

$$(\mu(j) + \lambda)\,E_j R = \varrho(j) + \sum_{k \neq j} \mu(j, k)\,[r(j, k) + E_k R], \qquad j = 1, 2, \ldots, r, \qquad (6)$$

moreover the expected discounted rewards $E_j R$, $j = 1, 2, \ldots, r$ are uniquely determined by (6).

We confine our study of discounted reward from the replacement process to the stationary replacement policies f only.

Let us denote for simplicity $E_j^f R = \Theta_f(j)$.

If $j \in I_f$ then (6) takes the form

$$(\mu(j) + \lambda)\,\Theta_f(j) = \varrho(j) + \mu(j, f(j))\,[v(j, f(j)) + \Theta_f(f(j))]$$

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

$$\Theta_f(j) = v(j, f(j)) + \Theta_f(f(j)).$$

If $j \notin I_f$ then from (6)

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

We have thus established a system of equations for determining the expected discounted reward from the process under the stationary replacement policy f:

$$v(j, f(j)) + \Theta_f(f(j)) - \Theta_f(j) = 0, \qquad j \in I_f, \qquad (7)$$

$$\varrho(j) + \sum_{k \neq j} \mu(j, k)\,[r(j, k) + \Theta_f(k) - \Theta_f(j)] - \lambda\Theta_f(j) = 0, \qquad j \notin I_f.$$

Theorem 1

System of equations (7) *has exactly one solution* $\Theta_f(j), j = 1, \ldots, r$.

Proof: For simplicity let us assume $I_f = \{1, \ldots, j - 1\}$, $1 < j \leqq r$. The matrix of system (7) has then the form

$$
\mathbf{M}^* = \left\| \begin{array}{cccc|c} -1 & 0 & \dots & 0 & \text{in any row only one unit,} \\ 0 & -1 & \dots & 0 & \text{the other elements zeros} \\ \dots & \dots & \dots & \dots & \\ 0 & 0 & \dots & -1 & \\ \hline \mu(j,1) & \dots & & \mu(j,j-1) & \mu(j,j)-\lambda \ \dots \ \mu(j,r) \\ \dots & \dots & \dots & \dots & \dots \\ \mu(r,1) & \dots & & \mu(r,j-1) & \mu(r,j) \ \dots \ \mu(r,r)-\lambda \end{array} \right\| =
$$

$$
= \left\| \begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array} \right\| .
$$

For finding the value of det $\mathbf{M}^*$ we add for every $i = 1, \dots, j - 1$ the i-th column to the $f(i)$-th column. We obtain

$$
\det \mathbf{M}^* = \det \left\| \begin{array}{cc} \mathbf{A} & \mathbf{0} \\ \mathbf{C} & \mathbf{D}^* \end{array} \right\|
$$

where

$$
\mathbf{D}^* = \left\| \begin{array}{ccc} d_{jj} - \lambda & \dots & d_{jr} \\ \dots & \dots & \dots \\ d_{rj} & \dots & d_{rr} - \lambda \end{array} \right\| ,
$$

$$
d_{kk} \leqq 0, \ d_{kl} \geqq 0, \ k \neq l, \ k, l = j, j+1, \dots, r; \ \sum_{l=j}^{r} d_{kl} = 0.
$$

As the only nonnegative characteristic number of the quasistochastic matrix (see [4], page 181) is $\lambda = 0$, it holds det $\mathbf{D}^* \neq 0$ for $\lambda > 0$. Thus det $\mathbf{M}^* =$ $=$ det $\mathbf{A}$. det $\mathbf{D}^* \neq 0$ and the matrix $\mathbf{M}^*$ is of full rank.

Let us introduce the *maximal expected discounted reward* (see [3], page 24)

$$
\hat{\Theta}(j) = \max_f \{\Theta_f(j)\}, \qquad j \in I.
$$

The stationary replacement policy $\hat{f}$ is called optimal, if

$$
\hat{\Theta}(j) = \Theta_{\hat{f}}(j), \qquad j \in I.
$$

The maximal reward will be characterized by the following theorem, in whose proof Howard's iteration procedure for finding $\hat{\Theta}(j)$, $j \in I$, and the responsive optimal stationary replacement policy will be described (see [1]).

Theorem 2

The maximal reward $\hat{\Theta}(j)$ is the unique solution of the following equation

$$
\max \{v(j,k) + \hat{\Theta}(k) - \hat{\Theta}(j), k \in D_j;
$$

$$
\varrho(j) + \sum_{k \neq j} \mu(j,k) [r(j,k) + \hat{\Theta}(k) - \hat{\Theta}(j)] - \lambda \hat{\Theta}(j)\} = 0, \qquad j \in I. \tag{8}
$$

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

Proof:

We prove first the existence of the solution of system (8) by Howard's iteration procedure. Chosing an arbitrary stationary replacement policy f_0 we succesively determine the stationary replacement policies $f_1, \ldots, f_n, \ldots$ as follows:

a) we solve the system of equations (to simplify the notation we write $\Theta_{f_n}(j) = \Theta_n(j)$)

$$v(j, f_n(j)) + \Theta_n(f_n(j)) - \Theta_n(j) = 0, \qquad j \in I_{f_n}, \tag{9}$$
$$\varrho(j) + \sum_{k \neq j} \mu(j, k) [r(j, k) + \Theta_n(k) - \Theta_n(j)] - \lambda \Theta_n(j) = 0, \qquad j \notin I_{f_n},$$

by Theorem 1 $\Theta_n(j), j \in I$ are determined by the system uniquely;

b) for all $j \in I$ we succesively determine

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

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

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

we choose

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

in the contrary, if the maximum is obtained by the expression

$$v(j, k) + \Theta_n(k) - \Theta_n(j) \qquad \text{for some } k \in D_j,$$

we choose

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

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

c) If the policy f_{n+1} does not posses 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 such 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 the remaining states we have $f'_{n+1}(j) = f_{n+1}(j)$. We now show the correctness of the procedure in c).

Suppose that $f_n(j) \notin I_{f_n}$ for all $j \in I_{f_n}$ and that the policy f_{n+1} was constructed in the above described way. Let

$$j \in I_{f_{n+1}}, f_{n+1}(j) = k \in I_{f_{n+1}}, f_{n+1}(k) = k'. \tag{10}$$

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

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

and therefore by Assumption 2

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

The equality must hold here, because the expression

$$v(j, k) + \Theta_n(k) - \Theta_n(j)$$

is maximal (replacement $j \to k$ under the policy f_{n+1} in the state j) from all expressions $v(j, i) + \Theta_n(i) - \Theta_n(j)$, $i \in D_j$. We are thus led to the conclusion that k' is equivalent to k for state j, moreover

$$v(k, k') + \Theta_n(k') - \Theta_n(k) = 0. \tag{11}$$

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

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

then it would be also

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

which however contradicts the assumption on the replacement policy f_n. It suffices therefore to change the constructed policy f_{n+1} in the way described in c).

For thus constructed replacement policy then

$$v(j, f_{n+1}(j)) + \Theta_n(f_{n+1}(j)) - \Theta_n(j) \geqq 0, \qquad j \in I_{f_{n+1}}, \tag{12}$$
$$\varrho(j) + \sum_{k \neq j} \mu(j, k) [r(j, k) + \Theta_n(k) - \Theta_n(j)] - \lambda \Theta_n(j) \geqq 0, \qquad j \notin I_{f_{n+1}}.$$

By Theorem 1

$$v(j, f_{n+1}(j)) + \Theta_{n+1}(f_{n+1}(j)) - \Theta_{n+1}(j) = 0, \qquad j \in I_{f_{n+1}}, \tag{13}$$
$$\varrho(j) + \sum_{k \neq j} \mu(j, k) [r(j, k) + \Theta_{n+1}(k) - \Theta_{n+1}(j)] - \lambda \Theta_{n+1}(j) = 0, \qquad j \notin I_{f_{n+1}}.$$

Subtracting (12) from (13) we obtain

$$\Theta_{n+1}(f_{n+1}(j)) - \Theta_n(f_{n+1}(j)) - \Theta_{n+1}(j) + \Theta_n(j) \leqq 0, \qquad j \in I_{f_{n+1}}, \tag{14}$$
$$\sum_{k \neq j} \mu(j, k) [\Theta_{n+1}(k) - \Theta_n(k) - \Theta_{n+1}(j) + \Theta_n(j)] - \lambda(\Theta_{n+1}(j) - \Theta_n(j)) \leqq 0, j \notin I_{f_{n+1}}.$$

For $j \notin I_{f_{n+1}}$ we obtain from (14)

$$[\Theta_n(j) - \Theta_{n+1}(j)] (\lambda + \sum_{k \neq j} \mu(j, k)) \leqq \sum_{k \neq j} \mu(j, k) (\Theta_n(k) - \Theta_{n+1}(k)),$$

whence after some modification

$$\Theta_n(j) - \Theta_{n+1}(j) \leqq \frac{\mu(j)}{\lambda + \mu(j)} \sum_{k \in I} p(j, k) [\Theta_n(k) - \Theta_{n+1}(k)]$$

it means by using the notation

$$c = \max_{j \notin I_{f_{n+1}}} \left\{ \frac{\mu(j)}{\lambda + \mu(j)} \right\}$$

we have for $j \notin I_{f_{n+1}}$

$$\Theta_n(j) - \Theta_{n+1}(j) \leqq c \max_{k \in I} \{\Theta_n(k) - \Theta_{n+1}(k)\}. \tag{15}$$

Relation (15) is valid also for $j \in I_{f_{n+1}}$ since for these j by the first row in (14)

$$\Theta_n(j) - \Theta_{n+1}(j) \leqq \Theta_n(f_{n+1}(j)) - \Theta_{n+1}(f_{n+1}(j)),$$

and Assumption 1 yields $f_{n+1}(j) \notin I_{f_{n+1}}$.

Thus, from (15) we have

$$\max_{j \in I} \{\Theta_n(j) - \Theta_{n+1}(j)\} \leqq c \max_{k \in I} \{\Theta_n(k) - \Theta_{n+1}(k)\}.$$

The last inequality may be satisfied by $0 < c < 1$ if and only if

$$\Theta_n(j) - \Theta_{n+1}(j) \leqq 0, \qquad j \in I,$$

i.e. if

$$\Theta_n(j) \leqq \Theta_{n+1}(j), \qquad j \in I.$$

The sequence $\Theta_n(j)$ is nondecreasing if n is increasing. As the set of the stationary replacement policies is finite, there exists m such that

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

Using (9) and constructing the policy f_{m+1} in the above way we obtain for $j \in I_{f_{m+}}$

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

For $j \notin I_{f_{m+1}}$ we have

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

We can see that $\hat{\Theta}(j) = \Theta_m(j)$, $j \in I$, is a solution of equation (8). We verify now that (8) determines $\hat{\Theta}(j)$ uniquely.

Let $\bar{\Theta}(j)$, $j \in I$, be another solution of equation (8), i.e. let

$$\max \{v(j,k) + \bar{\Theta}(k) - \bar{\Theta}(j), \quad k \in D_j; \tag{16}$$
$$\varrho(j) + \sum_{k \neq j} \mu(j,k)[r(j,k) + \bar{\Theta}(k) - \bar{\Theta}(j)] - \lambda\bar{\Theta}(j)\} = 0, \quad j \in I.$$

Let $\hat{f}$ be the replacement policy defined by Theorem 2. Then

$$v(j, \hat{f}(j)) + \hat{\Theta}(\hat{f}(j)) - \hat{\Theta}(j) = 0, \quad j \in I_{\hat{f}}, \tag{17}$$
$$\varrho(j) + \sum_{k \neq j} \mu(j,k)[r(j,k) + \hat{\Theta}(k) - \hat{\Theta}(j)] - \lambda\hat{\Theta}(j) = 0, \quad j \notin I_{\hat{f}}.$$

According to (16)

$$v(j, \hat{f}(j)) + \bar{\Theta}(\hat{f}(j)) - \bar{\Theta}(j) \leqq 0, \quad j \in I_{\hat{f}}, \tag{18}$$
$$\varrho(j) + \sum_{k \neq j} \mu(j,k)[r(j,k) + \bar{\Theta}(k) - \bar{\Theta}(j)] - \lambda\bar{\Theta}(j) \leqq 0, \quad j \notin I_{\hat{f}}.$$

Subtracting (17) from (18) we obtain

$$\bar{\Theta}(\hat{f}(j)) - \hat{\Theta}(\hat{f}(j)) - \bar{\Theta}(j) + \hat{\Theta}(j) \leqq 0, \quad j \in I_{\hat{f}}, \tag{19}$$
$$\sum_{k \neq j} \mu(j,k)[\bar{\Theta}(k) - \hat{\Theta}(k) - \bar{\Theta}(j) + \hat{\Theta}(j)] - \lambda(\bar{\Theta}(j) - \hat{\Theta}(j)) \leqq 0, \quad j \notin I_{\hat{f}}.$$

For simplicity we write $\bar{\Theta}(j) - \hat{\Theta}(j) = w(j)$, $j \in I$, and obtain for $j \notin I_{\hat{f}}$ from the second equation of (19)

$$w(j) \geqq \frac{\mu(j)}{\lambda + \mu(j)} \sum_{k \neq j} p(j,k)\, w(k) \geqq d \min_{k \in I} \{w(k)\},$$

where

$$d = \min_{j \notin I_{\hat{f}}} \left\{ \frac{\mu(j)}{\lambda + \mu(j)} \right\}.$$

The relation

$$w(j) \geqq d \min_{k \in I} \{w(k)\}$$

is valid for all $j \in I$ with respect to (19) and to Assumption 1. This yields

$$\min_{j \in I} \{w(j)\} \geqq d \min_{k \in I} \{w(k)\}.$$

Since $0 < d < 1$, this inequality may hold only if $\min_{j \in I} \{w(j)\} \geqq 0$, i.e. if

$$w(j) = \bar{\Theta}(j) - \hat{\Theta}(j) \geqq 0, \quad j \in I,$$

it is

$$\bar{\Theta}(j) \geqq \hat{\Theta}(j), \quad j \in I.$$

Analogous may be proved that $\bar{\Theta}(j) \leqq \hat{\Theta}(j)$, $j \in I$, therefrom

$$\bar{\Theta}(j) = \hat{\Theta}(j), \quad j \in I.$$

It still remains to verify that the policy $\hat{f}$ is an optimal stationary one.

Theorem 1 tells us that the system

$$\nu(j, \hat{f}(j)) + \Theta(\hat{f}(j)) - \Theta(j) = 0, \qquad j \in I_{\hat{f}}, \tag{20}$$
$$\varrho(j) + \sum_{k \neq j} \mu(j, k) [r(j, k) + \Theta(k) - \Theta(j)] - \lambda\Theta(j) = 0, \qquad j \notin I_{\hat{f}},$$

determines $\Theta_{\hat{f}}(j)$, $j \in I$, uniquely. Comparing (20) and (19) we obtain $\Theta_{\hat{f}}(j) = \hat{\Theta}(j), j \in I$.

REFERENCES

[1] Howard, R. A.: *Dynamic programming and Markov Processes,* M. I. T. Press and John Wiley, New York – London (1960).

[2] Mandl, P.: *An identity for Markovian replacement processes,* J. Appl. Prob. 6, No. 2, 348 – 354 (1969).

[3] Mandl, P.: *Řízené Markovovy řetězce,* příloha časopisu Kybernetika, roč. 5, Academia Praha (1969).

[4] Sarymsakov, T. A.: *Osnovy teo-ii processov Markova,* Moskva (1954).

Author's address:
RNDr. Pavla Kunderová, CSc.
katedra matematické analýzy a numerické matematiky
přírodovědecké fakulty University Palackého
Leninova 26
771 46 Olomouc, ČSSR

OČEKÁVANÝ DISKONTOVANÝ VÝNOS Z MARKOVOVA PROCESU S VÝNOSY A OBNOVAMI

Souhrn

Uvažuje se Markovův proces s výnosy a obnovami popsaný v článku [2]. Je odvozena soustava rovnic pro určování očekávaného diskontovaného výnosu z procesu (viz [3]) při užití stacionární strategie obnovy. Maximální očekávaný diskontovaný výnos je charakterizován větou 2, v jejímž důkaze je popsána Howardova iterační metoda (viz [1]) nacházení maximálního výnosu a metoda určování odpovídající optimální stacionární strategie.

ОЖИДАЕМЫЙ ДОХОД С ПЕРЕОЦЕНКОЙ ИЗ МАРКОВСКОГО ПРОЦЕССА С ДОХОДАМИ И ВОССТАНОВЛЕНИЯМИ

Резюме

В работе рассмотрен процесс Маркова с восстановлениями и доходами определенный в [2]. Найдена система уравнений для определения ожидаемого дохода с переоценкой (смотри [3]) при использовании стационарной стратегии восстановления. Максимальный ожидаемый доход с переоценкой характеризуется теоремой 2, в доказательстве которой описан итерационный метод Ховарда для нахождения максимального дохода и нахождения отвечающей оптимальной стационарной стратегии.