ON CONTINUOUS-TIME MARKOV PROCESSES WITH REWARDS, I

BY HIROHISA HATORI

1. In the previous paper [1] we have discussed Markov chains with rewards. In this paper we shall extend our previous work to continuous-time Markov processes with rewards.

2. As the preparation of the following sections, we shall state some well-known properties of Markov processes. Let X_t, $t\geqq 0$ be a continuous-time Markov process with the state space $\boldsymbol{S}=\{1, 2, \cdots, N\}$. The quantity a_{jk} is defined as follows: In a short time interval dt, the process that is now in state $j\in\boldsymbol{S}$ will make a transition to state $k\in\boldsymbol{S}$ with probability $a_{jk}dt+o(dt)$ $(j\neq k)$. The probability of two or more state transitions is $o(dt)$. Then, this Markov process is described by the transition-rate matrix $A=(a_{jk})$ with elements a_{jk} where the diagonal elements of A are defined by $a_{jj}=-\sum_{k\neq j}a_{jk}$ $(j=1, 2, \cdots, N)$. The probability that the system occupies state j at time t after the start of the process is the state probability $\pi_j(t)\overset{\text{def}}{=}P\{X_t=j\}$ and we have

$$\frac{d}{dt}\pi_k(t)=\sum_{j=1}^{N}\pi_j(t)a_{jk} \qquad (k=1, 2, \cdots, N). \tag{1}$$

In vector-form we may write (1) as

$$\frac{d}{dt}\boldsymbol{\pi}(t)=\boldsymbol{\pi}(t)\cdot A, \tag{2}$$

where $\boldsymbol{\pi}(t)\overset{\text{def}}{=}[\pi_1(t), \cdots, \pi_N(t)]$ is the vector with the components $\pi_j(t)$. Let us designate by $\boldsymbol{\Pi}(s)$ the Laplace transform of the state-probability vector $\boldsymbol{\pi}(t)$. If we take the Laplace transform of (2), we obtain

$$s\boldsymbol{\Pi}(s)-\boldsymbol{\pi}(0)=\boldsymbol{\Pi}(s)\cdot A$$

and so

$$\boldsymbol{\Pi}(s)=\boldsymbol{\pi}(0)[sI-A]^{-1}, \tag{3}$$

where I is the identity matrix. Under a certain weak condition, the equation $\det(sI-A)=0$ has a simple root $s=0$ and, $\alpha_1, \cdots, \alpha_k$ being its remaining roots, the real parts $\Re(\alpha_l)$ of α_l $(l=1, 2, \cdots, k)$ are negative. Each element of $[sI-A]^{-1}$ is a function of s with a factorable denominator $s(s-\alpha_1)^{m_1}\cdots(s-\alpha_k)^{m_k}$, where $m_1, \cdots, m_k$, are the multiplicities of $\alpha_1, \cdots, \alpha_k$, respectively. By partial-fraction expansion we can express each element as the sum of the fractions whose forms are const./s and

Received December 16, 1965.

const.$/(s-\alpha_l)^\nu$ $(\nu=1,\cdots,m_l;\ l=1,2,\cdots,k)$. Expressing this fact in matrix-form, we have

$$[sI-A]^{-1}=\frac{1}{s}S+\sum_{l=1}^{k}\sum_{\nu=1}^{m_l}\frac{1}{(s-\alpha_l)^\nu}T_{l\nu},$$

where S and $T_{l\nu}$ $(\nu=1,\cdots,m_l;\ l=1,2,\cdots,k)$ are $N\times N$ matrices independent of s, which implies that $[sI-A]^{-1}$ is the Laplace transform of

$$H(t)\overset{\text{def}}{=}S+\sum_{l=1}^{k}\sum_{\nu=1}^{m_l}\frac{t^{\nu-1}}{(\nu-1)!}e^{\alpha_l t}T_{l\nu},$$

that is,

$$[sI-A]^{-1}=\int_0^\infty H(t)e^{-st}dt, \qquad (s>0).$$

Therefore, we have

$$H(t)\to S \qquad (t\to\infty),$$

which implies with (3)

$$\boldsymbol{\pi}(t)=\boldsymbol{\pi}(0)H(t)\to\boldsymbol{\pi}(0)S \qquad (t\to\infty). \tag{4}$$

S is a stochastic matrix and its j-th row is the limiting-state-probability vector of the process if it starts in the j-th state.

REMARK 1. If the matrix A is indecomposable, then the equation $\det(sI-A)=0$ has a simple root $s=0$ and $\alpha_1,\cdots,\alpha_k$ being its remaining roots, the real parts $\Re(\alpha_l)$ of α_l $(l=1,2,\cdots,k)$ are strictly negative. In what follows, we shall prove this fact. For a root α of $\det(sI-A)=0$, there exists a non-zero vector $\begin{bmatrix} z_1\\ \vdots\\ z_N\end{bmatrix}$ such that

$$A\begin{bmatrix} z_1\\ \vdots\\ z_N\end{bmatrix}=\alpha\begin{bmatrix} z_1\\ \vdots\\ z_N\end{bmatrix}. \tag{5}$$

Taking $j_0\in\boldsymbol{S}$ such that $\mathrm{Max}_{j=1,2,\cdots,N}|z_j|=|z_{j_0}|>0$, we may assume without loss of generality $z_{j_0}=1$ and $|z_j|\leqq 1$ $(j=1,2,\cdots,N)$, because (5) holds again by replacing z_j/z_{j_0} for z_j $(j=1,2,\cdots,N)$. Then, we have from (5)

$$\alpha=\alpha z_{j_0}=\sum_{j\neq j_0}a_{j_0j}z_j+a_{j_0j_0}$$

and so

$$\Re(\alpha)=\sum_{j\neq j_0}a_{j_0j}\Re(z_j)+a_{j_0j_0}\leqq\sum_{j\neq j_0}a_{j_0j}+a_{j_0j_0}=0,$$

because $\Re(z_j)\leqq|z_j|\leqq 1$, $a_{j_0j}\geqq 0$ $(j\neq j_0)$ and $a_{j_0j_0}=-\sum_{j\neq j_0}a_{j_0j}$. Therefore we get $\Re(\alpha_l)\leqq 0$ $(l=1,2,\cdots,k)$. Now, the matrix $A+\lambda I$, where $\lambda=\mathrm{Max}_{j=1,\cdots,N}|a_{jj}|$, has $\lambda,\lambda+\alpha_1,\cdots,\lambda+\alpha_k$ as its eigen values. Since $A+\lambda I$ is an indecomposable matrix with non-negative elements, we have by the well-known theorem on matrices with non-negative elements that $s=\lambda$ is a simple root of $\det(sI-A-\lambda I)=0$ and $|\lambda+\alpha_l|\leqq\lambda$

$(l=1, 2, \cdots, k)$. Hence we know that $s=0$ is a simple root of $\det(sI-A)=0$ and $\Re(\alpha_l)<0$, $(l=1, 2, \cdots, k)$.

3. To simplify the explanation of our method in this section, we assume that the equation $\det(sI-A)=0$ has the simple roots $0, \alpha_1, \cdots, \alpha_{N-1}$, that is, $k=N-1$ and $m_1=m_2=\cdots=m_k=1$. Let us suppose that the system earns a reward at the rate of r_{jj} dollars per unit time during all the time that it occupies state j. Suppose further that when the system makes a transition from state j to state k $(j\neq k)$, it receives a reward of r_{jk} dollars. Then, the characteristic function of the distribution of the total reward $R(t)$ that the system will earn in a time t if it starts in state j is

$$\varphi_{jt}(\theta)\overset{\text{def}}{=}E\{e^{i\theta R(t)}|X_0=j\}, \tag{6}$$

where $i=\sqrt{-1}$ and θ is a real variable. Here, dt representing, as before, a very short time interval, we have

$$\varphi_{j,t+dt}(\theta)=(1+a_{jj}dt)e^{i\theta r_{jj}dt}\varphi_{jt}(\theta)+\sum_{k\neq j}a_{jk}dte^{i\theta r_{jk}}\varphi_{kt}(\theta)+o(dt)$$

and so

$$\frac{\partial}{\partial t}\varphi_{jt}(\theta)=(a_{jj}+i\theta r_{jj})\varphi_{jt}(\theta)+\sum_{k\neq j}a_{jk}e^{i\theta r_{jk}}\varphi_{kt}(\theta) \qquad (j=1, 2, \cdots, N). \tag{7}$$

Introducing the $N\times N$ matrix

$$A(\theta)\overset{\text{def}}{=}\begin{pmatrix} a_{11}+i\theta r_{11} & a_{12}e^{i\theta r_{12}} & \cdots & a_{1N}e^{i\theta r_{1N}} \\ a_{21}e^{i\theta r_{21}} & a_{22}+i\theta r_{22} & \cdots & a_{2N}e^{i\theta r_{2N}} \\ \cdot & \cdot & \cdots & \cdot \\ a_{N1}e^{i\theta r_{N1}} & a_{N2}e^{i\theta r_{N2}} & \cdots & a_{NN}+i\theta r_{NN} \end{pmatrix}$$

and the vector

$$\boldsymbol{\varphi}_t(\theta)\overset{\text{def}}{=}\begin{bmatrix} \varphi_{1t}(\theta) \\ \vdots \\ \varphi_{Nt}(\theta) \end{bmatrix},$$

we may write (7) as

$$\frac{\partial}{\partial t}\boldsymbol{\varphi}_t(\theta)=A(\theta)\boldsymbol{\varphi}_t(\theta). \tag{8}$$

If we take the Laplace transform of (8), we obtain

$$s\boldsymbol{\Phi}(\theta, s)-\boldsymbol{e}=A(\theta)\boldsymbol{\Phi}(\theta, s)$$

and so

$$\boldsymbol{\Phi}(\theta, s)=[sI-A(\theta)]^{-1}\boldsymbol{e} \tag{9}$$

for $s>0$ and θ in a neighborhood of $\theta=0$, where

$$\boldsymbol{\Phi}(\theta, s) \overset{\text{def}}{=} \int_0^\infty \boldsymbol{\varphi}_t(\theta) e^{-st} dt \qquad \text{and} \qquad \boldsymbol{e} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}.$$

The equation $\det(sI - A(\theta)) = 0$ in s has the N roots $\zeta_0(\theta), \zeta_1(\theta), \cdots, \zeta_{N-1}(\theta)$ such that $\Re(\zeta_l(\theta)) \leqq 0$ $(l = 0, 1, \cdots, N-1)$ and

$$\zeta_0(\theta) \to 0, \ \zeta_1(\theta) \to \alpha_1, \ \cdots, \ \zeta_{N-1}(\theta) \to \alpha_{N-1}$$

as $\theta \to 0$. Then, there exist positive constants ε and θ_0 such that

$$-\varepsilon < \Re(\zeta_0(\theta)) \leqq 0 \quad \text{and} \quad \Re(\zeta_l(\theta)) < -2\varepsilon \qquad (l = 1, 2, \cdots, N-1) \tag{10}$$

for $|\theta| < \theta_0$, because $\Re(\alpha_l) < 0$ $(l = 1, \cdots, N-1)$. The consideration similar to the one in the preceeding section give

$$\begin{aligned} \boldsymbol{\Phi}(\theta, s) &= \frac{1}{s - \zeta_0(\theta)} \boldsymbol{\sigma}(\theta) + \sum_{l=1}^{N-1} \frac{1}{s - \zeta_l(\theta)} \boldsymbol{\tau}_l(\theta) \\ &= \int_0^\infty \left\{ e^{\zeta_0(\theta) t} \boldsymbol{\sigma}(\theta) + \sum_{l=1}^{N-1} e^{\zeta_l(\theta) t} \boldsymbol{\tau}_l(\theta) \right\} e^{-st} dt \end{aligned} \tag{11}$$

and so

$$\begin{aligned} \boldsymbol{\varphi}_t(\theta) &= e^{\zeta_0(\theta) t} \boldsymbol{\sigma}(\theta) + \sum_{l=1}^{N-1} e^{\zeta_l(\theta) t} \boldsymbol{\tau}_l(\theta) \\ &= e^{\zeta_0(\theta) t} \left\{ \boldsymbol{\sigma}(\theta) + \sum_{l=1}^{N-1} e^{(\zeta_l(\theta) - \zeta_0(\theta)) t} \boldsymbol{\tau}_l(\theta) \right\}, \end{aligned} \tag{12}$$

where $\boldsymbol{\sigma}(\theta), \boldsymbol{\tau}_1(\theta), \cdots, \boldsymbol{\tau}_{N-1}(\theta)$ are N-dimensional vectors analytic on θ for $|\theta| < \theta_0$. Since $\boldsymbol{\varphi}_t(0) = \boldsymbol{e}$, we have $\boldsymbol{\sigma}(0) = \boldsymbol{e}$. From (12), we have

$$\begin{aligned} \frac{\partial}{\partial \theta} \boldsymbol{\varphi}_t(\theta) &= \zeta_0'(\theta) t e^{\zeta_0(\theta) t} \boldsymbol{\sigma}(\theta) + e^{\zeta_0(\theta) t} \boldsymbol{\sigma}'(\theta) \\ &\quad + \sum_{l=1}^{N-1} \{ \zeta_l'(\theta) t e^{\zeta_l(\theta) t} \boldsymbol{\tau}_l(\theta) + e^{\zeta_l(\theta) t} \boldsymbol{\tau}_l'(\theta) \} \end{aligned} \tag{13}$$

and so

$$\begin{aligned} \boldsymbol{v}(t) &= \frac{1}{i} \{ \zeta_0'(0) t \boldsymbol{e} + \boldsymbol{\sigma}'(0) \} + \frac{1}{i} \sum_{l=1}^{N-1} \{ \zeta_l'(0) t e^{\alpha_l t} \boldsymbol{\tau}_l(0) + e^{\alpha_l t} \boldsymbol{\tau}_l'(0) \} \\ &\doteqdot -i \zeta_0'(0) t \boldsymbol{e} - i \boldsymbol{\sigma}'(0) \qquad \text{as } t \to \infty, \end{aligned} \tag{14}$$

where $\boldsymbol{v}(t) \overset{\text{def}}{=} \begin{bmatrix} v_1(t) \\ \vdots \\ v_N(t) \end{bmatrix}$ is the vector with the components

$$v_j(t) \overset{\text{def}}{=} E\{R(t) | X_0 = j\} = \frac{1}{i} \left[\frac{\partial}{\partial \theta} \varphi_{jt}(\theta) \right]_{\theta = 0},$$

which implies that $g \overset{\text{def}}{=} -i \zeta_0'(0)$ is a real number. In the similar way, we have

$$E\{R(t)^2|X_0=j\} \doteqdot -\zeta_0'(0)^2t^2-[\zeta_0''(0)+2\zeta_0'(0)\sigma_j']t-\sigma_j'' \qquad \text{as } t\to\infty,$$

where σ_j' and σ_j'' are the j-th component of $\sigma'(0)$ and $\sigma''(0)$, and

$$E\{R(t)^2|X_0=j\}-[E\{R(t)|X_0=j\}]^2 \doteqdot -\zeta_0''(0)t-\sigma_j''+\sigma_j'^2 \qquad \text{as } t\to\infty,$$

which implies $-\zeta_0''(0)$ is positive in general. Now, we shall consider the asymptotic behavior of $R(t)$ as $t\to\infty$. The characteristic function of the distribution of the random variable $[R(t)-gt]/\sqrt{t}$ under the condition $X_0=j$ is

$$\begin{aligned}\psi_{jt}(\theta) &\overset{\text{def}}{=} E\{e^{i\theta[R(t)-gt]/\sqrt{t}}|X_0=j\}\\ &=e^{-ig\sqrt{t}\theta}\varphi_{jt}\left(\frac{\theta}{\sqrt{t}}\right)\\ &=e^{-ig\sqrt{t}\theta}e^{\zeta_0(\theta/\sqrt{t})t}\left\{\sigma_j\left(\frac{\theta}{\sqrt{t}}\right)+\sum_{l=1}^{N-1}e^{(\zeta_l(\theta/\sqrt{t})-\zeta_0(\theta/\sqrt{t}))t}\tau_{lj}\left(\frac{\theta}{\sqrt{t}}\right)\right\},\end{aligned} \tag{15}$$

where $\sigma_j(\theta)$ and $\tau_{lj}(\theta)$ $(l=1,\cdots,N-1)$ are the j-th components of the vectors $\boldsymbol{\sigma}(\theta)$ and $\boldsymbol{\tau}_l(\theta)$ $(l=1,\cdots,N-1)$, respectively. For fixed θ, we have $|\theta/\sqrt{t}|<\theta_0$ for all sufficiently large t so that by (10)

$$\Re\left(\zeta_l\left(\frac{\theta}{\sqrt{t}}\right)-\zeta_0\left(\frac{\theta}{\sqrt{t}}\right)\right)<-\varepsilon<0$$

and

$$e^{(\zeta_l(\theta/\sqrt{t})-\zeta_0(\theta/\sqrt{t}))t}\to 0 \qquad \text{as } t\to\infty. \tag{16}$$

On the other hand, we have

$$\begin{aligned}&e^{-ig\sqrt{t}\theta}e^{\zeta_0(\theta/\sqrt{t})t}\\ &=e^{-ig\sqrt{t}\theta+\{\zeta_0'(0)\sqrt{t}\theta+(1/2)\zeta_0''(0)\theta^2+O(1/\sqrt{t})\}}\to e^{(1/2)\zeta_0''(0)\theta^2} \qquad \text{as } t\to\infty,\end{aligned} \tag{17}$$

because $\zeta_0(0)=0$ and $g=-i\zeta_0'(0)$. From (15), (16) and (17), we get

$$\psi_{jt}(\theta)\to e^{(1/2)\zeta_0''(0)\theta^2} \qquad \text{as } t\to\infty$$

and so

$$\begin{aligned}\psi_t(\theta) &\overset{\text{def}}{=} E\{e^{i\theta[R(t)-gt]/\sqrt{t}}\}\\ &=\sum_{j=1}^{N}\pi_j(0)\psi_{jt}(\theta)\to e^{(1/2)\zeta_0''(0)\theta^2} \qquad \text{as } t\to\infty,\end{aligned}$$

which implies that $[R(t)-gt]/\sqrt{t}$ converges in distribution to the normal distribution $N(0, -\zeta_0''(0))$.

REMARK 2. It follows from (8) that

$$\frac{\partial^2}{\partial t\partial\theta}\boldsymbol{\varphi}_t(\theta)=A'(\theta)\boldsymbol{\varphi}_t(\theta)+A(\theta)\cdot\frac{\partial}{\partial\theta}\boldsymbol{\varphi}_t(\theta). \tag{18}$$

Since $\left[\frac{\partial}{\partial\theta}\boldsymbol{\varphi}_t(\theta)\right]_{\theta=0}=i\boldsymbol{v}(t)$, $A(0)=A$ and

$$\frac{1}{i}A'(0)\overset{\text{def}}{=}Q_1=\begin{pmatrix} r_{11} & r_{12}a_{12} & \cdots & r_{1N}a_{1N} \\ r_{21}a_{21} & r_{22} & \cdots & r_{2N}a_{2N} \\ \cdot & \cdot & \cdots & \cdot \\ r_{N_1}a_{N_1} & r_{N_2}a_{N_2} & \cdots & r_{NN} \end{pmatrix},$$

we have, by setting $\theta=0$ in (18),

$$\frac{d}{dt}\boldsymbol{v}(t)=Q_1\boldsymbol{e}+A\boldsymbol{v}(t), \tag{19}$$

which has been shown by Howard [2]. He has given from (19) the asymptotic form of $\boldsymbol{v}(t)$ as $t\to\infty$, which is essentially equivalent to (14). By differentiating the both sides of (18) $w.r.t.\,\theta$ and setting $\theta=0$, we have

$$\frac{d}{dt}\boldsymbol{w}(t)=Q_2\boldsymbol{e}+2Q_1\boldsymbol{v}(t)+A\boldsymbol{w}(t), \tag{20}$$

where $\boldsymbol{w}(t)\overset{\text{def}}{=}\begin{bmatrix} w_1(t) \\ \vdots \\ w_N(t)\end{bmatrix}$ is the vector with components $w_j(t)\overset{\text{def}}{=}E\{R(t)^2|X_0=j\}$ and

$$Q_2\overset{\text{def}}{=}\begin{pmatrix} 0 & r_{12}^2a_{12} & \cdots & r_{1N}^2a_{1N} \\ r_{21}^2a_{21} & 0 & \cdots & r_{2N}^2a_{2N} \\ \cdot & \cdot & \cdots & \cdot \\ r_{N1}^2a_{N_1} & r_{N2}^2a_{N_2} & \cdots & 0 \end{pmatrix}.$$

By taking Laplace transform of (20) and using the method similar to the one in [1], we can find the asymptotic forms that $w_j(t)$ and $\mathrm{Var}\,(R(t))$ assume for large t.

REMARK 3. Let f be any real valued function defined on $\boldsymbol{S}$. In the case where $r_{jj}=f(j)$ and $r_{jk}=0$ $(j\neq k)$, we have that

$$R(t)=\int_0^t f(X_\tau)d\tau$$

and the random variable $[\int_0^t f(X_\tau)d\tau-gt]/\sqrt{t}$ converges in distribution to a normal distribution as $t\to\infty$. Therefore we have the central limit theorem for continuous-time Markov processes.

REMARK 4. Although $\Re(\zeta_0(\theta))\leqq 0$ is derived from the analyticity of $\boldsymbol{\Phi}(\theta,s)$, we shall give a proof similar to the one in Remark 1. Since there exist a non-zero vector $\begin{bmatrix} z_1 \\ \vdots \\ z_N\end{bmatrix}$ and a state j_0 such that $A(\theta)\begin{bmatrix} z_1 \\ \vdots \\ z_N\end{bmatrix}=\zeta_0(\theta)\begin{bmatrix} z_1 \\ \vdots \\ z_N\end{bmatrix}$, $|z_j|\leqq 1$ $(j=1,2,\cdots,N)$ and $z_{j_0}=1$, we have

$$\zeta_0(\theta)=\zeta_0(\theta)z_{j_0}=\sum_{j\neq j_0}a_{j_0j}e^{i\theta r_{j_0j}}z_j+(a_{j_0j_0}+i\theta r_{j_0j_0})$$

and so

$$\Re(\zeta_0(\theta)) = \sum_{j \neq j_0} \Re(a_{j_0 j} e^{i\theta r_{j_0 j}} z_j) + a_{j_0 j_0}$$
$$\leqq \sum_{j \neq j_0} a_{j_0 j} |e^{i\theta r_{j_0 j}} z_j| + a_{j_0 j_0} \leqq \sum_{j \neq j_0} a_{j_0 j} + a_{j_0 j_0} = 0.$$

Therefore, we have $\Re(\zeta_0(\theta)) \leqq 0$.

4. In this section, we shall outline the case with the discounting. Let us define a discount rate $0 < \alpha < \infty$ in such a way that a unit quantity of money received after a very short time interval dt is now worth $1 - \alpha dt$. Then, for the characteristic function $\varphi_{jt}(\theta)$ of the present value $R(t)$ of the total reward of the system in time t under the condition $X_0 = j$, we have

$$\varphi_{j,t+dt}(\theta) = (1 + a_{jj} dt) e^{i\theta(1-\alpha dt) r_{jj} dt} \varphi_{jt}((1-\alpha dt)\theta) + \sum_{k \neq j} a_{jk} dt\, e^{i\theta(1-\alpha dt) r_{jk}} \varphi_{kt}((1-\alpha dt)\theta)$$

and so

$$\frac{\partial}{\partial t} \varphi_{jt}(\theta) = (a_{jj} + i\theta r_{jj}) \varphi_{jt}(\theta) - \alpha\theta \cdot \frac{\partial}{\partial \theta} \varphi_{jt}(\theta) + \sum_{k \neq j} a_{jk} e^{i\theta r_{jk}} \varphi_{kt}(\theta),$$

which is expressed in the vector-form

$$\frac{\partial}{\partial t} \boldsymbol{\varphi}_t(\theta) + \alpha\theta \frac{\partial}{\partial \theta} \boldsymbol{\varphi}_t(\theta) = A(\theta) \boldsymbol{\varphi}_t(\theta). \tag{21}$$

By differentiating the both sides of (21) *w.r.t.* θ and setting $\theta = 0$, we have

$$\frac{d}{dt} \boldsymbol{v}(t) + \alpha \boldsymbol{v}(t) = Q_1 \boldsymbol{e} + A \boldsymbol{v}(t) \tag{22}$$

which has been shown by Howard [2]. He has shown from (22) $\boldsymbol{v} = \lim_{t\to\infty} \boldsymbol{v}(t) = [\alpha I - A]^{-1} Q_1 \boldsymbol{e}$. By differentiating twice the both sides of (21) *w.r.t.* θ and setting $\theta = 0$, we have

$$\frac{d}{dt} \boldsymbol{w}(t) + 2\alpha \boldsymbol{w}(t) = Q_2 \boldsymbol{e} + 2 Q_1 \boldsymbol{v}(t) + A \boldsymbol{w}(t), \tag{23}$$

from which we can find without difficulty $\boldsymbol{w} = \lim_{t\to\infty} \boldsymbol{w}(t)$ in terms of α, A, Q_1 and Q_2, where $\boldsymbol{w}(t) \overset{\text{def}}{=} \begin{bmatrix} w_1(t) \\ \vdots \\ w_N(t) \end{bmatrix}$ is the vector with the components

$$w_j(t) \overset{\text{def}}{=} E\{R(t)^2 | X_0 = j\} = -\left[\frac{\partial^2}{\partial \theta^2} \varphi_{jt}(\theta) \right]_{\theta=0}.$$

The author expresses his sincerest thanks to Prof. Y. Kawahara who has given valuable advices.

References

[1] HATORI, H., On Markov chains with rewards. Kōdai Math. Sem. Rep. **18** (1966), 184–192.

[2] HOWARD, R. A., Dynamic Programming and Markov Processes. M. I. T. Press (1960).

TOKYO COLLEGE OF SCIENCE.