KYBERNETIKA — VOLUME 17 (1981), NUMBER 3

ON AN OPTIMUM PRINCIPLE FOR PARTIALLY OBSERVED CONTROLLED JUMP MARKOVIAN PROCESSES

NGUYEN VAN HUU

This paper deals with an optimum principle for partially observed controlled jump Markovian processes in which at each time moment some function of the state of the process is observable. A representation of the cost function for each control is given and a local dynamic programming condition involving transition intensities and aposteriori probabilities of the states of the process is presented.

1. INTRODUCTION

Controlled jump processes have been intensively investigated by many authors, e.g. S. R. Pliska [7], R. Boel, P. Varaiya [1], M. H. A. Davis, R. Elliott [2], C. B. Wan, M. H. A. Davis [9], and R. W. Rishel [8].

Pliska considered completely observed controlled jump Markovian processes. Boel and Varaiya investigated general jump process, gave a local optimality condition for value increasing controls and derived a differential equation to define the value function in the case of complete information. The paper by Davis and Elliott dealt with optimality condition for completely observed controlled jump processes the probability distributions of which is defined by the so-called Lévy system (λ^u, Λ^u) depending on control u and showed that a control is optimal iff it maximizes some Hamiltonian. Wan and Davis derived conditions for existence of an optimal control for the case of complete information. The paper by Rishel dealt with the minimum principle and dynamic programming conditions for two dimensional jump Markovian processes when only one component is observable.

This paper concerns partially observable controlled jump Markovian processes in which at each time moment some function of the state is observable. A representation of the value function and a local dynamic programming condition involving transition intensities and aposteriori probabilities of the states of the process are obtained.

2. STATEMENT OF THE CONTROL PROBLEM

Let Ω be the class of all right continuous piece-wise constant functions $\omega(t)$ defined on $[0, T]$ with values in $I = \{1, 2, \ldots, m\}$.

For each $t \in [0, T]$ let $X_t(\omega)$ be the function defined on Ω by the coordinate mapping $X_t(\omega) = \omega(t)$. Define the σ-algebras

$$\mathscr{F}_t = \sigma(X_s(\omega), s \leqq t), \quad 0 \leqq t \leqq T.$$

Since $X_t(\omega)$ is right continuous piece-wise constant in t, $\mathscr{F} = \{\mathscr{F}_t, 0 \leqq t \leqq T\}$ is a right continuous family, i.e. $\mathscr{F}_t = \bigcap_{s>t} \mathscr{F}_s$.

To state control problem for a partially observed controlled jump Markovian process on finite horizon suppose the following:

(S_1) A subset Z of Euclidian space R^k is given and the σ-algebra $\mathscr{R}$ of its Borel subsets is defined. $(Z, \mathscr{R})$ represents the control space.

(S_2) A family of increasing σ-algebras $\{\mathscr{G}_t, 0 \leqq t \leqq T\}$ is defined so that $\mathscr{G}_t \subset \mathscr{F}_t$. $\mathscr{G}_t$ is regarded as the σ-algebra of all observable information up to time t.

(S_3) Consider the class $\mathscr{U}$ of all functions $u : \Omega \times [0, T] \to Z$ such that u is a $\mathscr{G}$-predictable function, i.e. $u(\omega, t)$ is measurable with respect to the σ-algebra on $\Omega \times [0, T]$ generated by all $\mathscr{G}_t$-adapted left continuous functions $v : \Omega \times [0, T] \to Z$, where $\mathscr{G} = \{\mathscr{G}_t, 0 \leqq t \leqq T\}$. $\mathscr{U}$ is called the class of admissible controls.

(S_4) There is a bounded measurable function

$$q : I \times I \times Z \to R_+$$

such that $q(i, i, z) = 0$ for all $i \in I$, $z \in Z$.

Let $\tau_1, \tau_2, \ldots (\tau_1 < \tau_2 < \ldots \leqq T)$ be the jump times of X_t on $[0, T]$, $\chi(A)$ denote the indicator of the set A, and

$$N_i(t) = \sum_{\tau_k \leqq t} \chi(X_{\tau_k} = i), \quad i \in I. \tag{1}$$

Thus $N_i(t)$ is the counting process of the jumps of the process $\{X_t, 0 \leqq t \leqq T\}$ into state i on time interval $[0, t]$.

It is easy to see that

$$\mathscr{F}_t = \sigma(X_s(\omega), s \leqq t) = \sigma(N_i(s), s \leqq t, i \in I).$$

According to Theorem 3.6 of Jacod [4], to each u and to fixed initial probabilities p_i $(p_i \geqq 0, \sum_{i \in I} p_i = 1)$ there exists a unique probability measure P^u on $(\Omega, \mathscr{F}_T)$ such that

(i) $$P^u\{X_0 = i\} = p_i, \quad i \in I,$$

(ii) $$\left\{N_i(t) - \int_0^t q(X_s, i, u_s)\, \mathrm{d}s, \quad 0 \leqq t \leqq T\right\}$$

is an $(\mathscr{F}, P^u)$-martingale for all $i \in I$.

Definition 2.1. The family $\{X_t, \mathcal{G}_t, \mathcal{F}_t, P^u, 0 \leqq t \leqq T\}$ defined above is called a *partially observed controlled jump Markovian process.*

For $u_t \equiv z \in Z$ the process $\{X_t, \mathcal{F}_t, P^u, 0 \leqq t \leqq T\}$ is Markovian with transition intensities $q(i, j, z)$ for $i \neq j$, and with the initial distribution $p = (p_1, \ldots, p_m)$, i.e. the probability measure P^z is defined by p and transition probabilities

$$P^z\{X_{t+\Delta t} = j/X_t = i\} = \begin{cases} q(i, j, z)\,\Delta t + o(\Delta t) & \text{for } i \neq j \\ 1 - q(i, z)\,\Delta t + o(\Delta t) & \text{for } i = j \end{cases}$$

as $\Delta t \to 0$, where

$$q(i, z) = \sum_{j \neq i} q(i, j, z)\,.$$

The process maintains the Markovian property for $u_t = u(X_t)$. For this reason the process $(X_t, \mathcal{F}_t, P^u)$, $u \in \mathcal{U}$ is called a Markovian controlled process with controlled transition intensity.

Furthermore, we can fix a measure P on $(\Omega, \mathcal{F}_T)$ such that

(i) $$P(X_0 = i) = p_i\,, \quad i \in I\,,$$

(ii) $$\left\{N_i(t) - \int_0^t [1 - \chi(X_s = i)]\,\mathrm{d}s\right\}$$

is an $(\mathcal{F}, P)$-martingale for all $i \in I$.

Thus $(X_t, \mathcal{F}_t, P, 0 \leqq t \leqq T)$ is Markovian process with transition intensities 1 for $i \neq j$.

It follows from Proposition 4.3 and Theorem 4.5 in [4] that $P^u \ll P$ on $\mathcal{F}_T$ and

(2) $$\mathrm{d}P^u/\mathrm{d}P = \exp\left(L^u_T\right)$$

where

(3) $$L^u_t = \int_0^t [m - 1 - \sum_{i \in I} q(X_s, i, u_s)]\,\mathrm{d}s + \sum_{0 \leqq s \leqq t} \chi(X_{s-} = X_s) \log q(X_{s-}, X_s, u_s)$$
$$= \int_0^t [m - 1 - \sum_{i \in I} q(X_s, i, u_s)]\,\mathrm{d}s + \sum_{i \in I} \int_0^t \log q(X_s, i, u_s)\,\mathrm{d}N_i(s)$$

with the convention that $0 \log 0 = 0$, and $X_{s-} = \lim_{t \downarrow s} X_t$.

Further, let E and E^u denote the expectation with respect to P and P^u, respectively. Putting

$$\varrho^u = \exp\left(L^u_T\right), \quad \varrho^t_s(u) = \exp\left(L^u_t - L^u_s\right), \quad 0 \leqq s \leqq t \leqq T,$$

we have

$$\varrho^t_s(u) = \varrho^r_s(u)\,\varrho^t_r(u)\,, \quad s \leqq r \leqq t$$

and

$$\mathsf{E}\{\varrho^t_s(u)/\mathcal{F}_s\} = 1 \quad \text{for all} \quad 0 \leqq s \leqq T\,.$$

Let $\mathscr{U}_s^t$ be the collection of the restrictions of admissible controls u of $\mathscr{U}$ to $(s, t]$ with $\mathscr{U}_0^T = \mathscr{U}$.

It is easy to see that the class $\mathscr{U}$ has the following property:

Let $u \in \mathscr{U}_0^s$, $v \in \mathscr{U}_s^T$, then $w = uv$ defined by

$$w(t) = \begin{cases} u(t) & \text{for} \quad 0 \leqq t \leqq s \\ v(t) & \text{for} \quad s < t \leqq T \end{cases}$$

also belongs to $\mathscr{U}$.

Let us now consider the cost structure.

Assume given bounded functions

$$c : [0, T] \times I \times Z \to R\,, \quad r : [0, T] \times I \times I \times Z \to R$$

such that $r(t, i, i, z) = 0$ for all $(t, i, z) \in [0, T] \times I \times Z$. We associate with each trajectory the cost of the form

$$\int_0^T c(t, X_t, u_t)\, \mathrm{d}t + \sum_{0 \leqq t \leqq T} r(t, X_{t-}, X_t, u_t)\,.$$

Then the expected cost on $[0, T]$ is given by

$$J(u) = \mathsf{E}^u \left[\int_0^T c(t, X_t, u_t)\, \mathrm{d}t + \sum_{0 \leqq t \leqq T} r(t, X_{t-}, X_t, u_t) \right]. \tag{4}$$

Definition 2.2. Suppose that $u \in \mathscr{U}_0^t$, $v \in \mathscr{U}_t^T$, then the quantity

$$J_t(u, v) = \mathsf{E}^{uv} \left\{ \int_t^T c_s^v\, \mathrm{d}s + \sum_{t \leqq s \leqq T} r_s^v / \mathscr{G}_t \right\}, \tag{5}$$

where for brevity

$$c_s^v = c(s, X_s, v_s)\,, \quad r_s^v = r(s, X_{s-}, X_s, v_s)\,,$$

is called the remaining cost from time t onward.

The function

$$V(t, u) = \bigwedge_{v \in \mathscr{U}_t^T} J_t(u, v) \tag{6}$$

is called the value function, where $\bigwedge$ denotes the infimum operator.

Note that the function $V(t, u)$ defined as the infimum of $\{J_t(u, v),\, v \in \mathscr{U}_t^T\}$ in space $L_1(P^u)$ — a complete lattice with customary ordering structure of functions — always exists.

Definition 2.3. A control $u^* \in \mathscr{U}$ is said to be optimal if

$$J(u^*) = \inf_{u \in \mathscr{U}} J(u) = J^*\,.$$

The problem is to find an optimal control, if it exists.

3. ON NECESSARY AND SUFFICIENT CONDITIONS FOR OPTIMALITY

Using the relative completeness of the set $\mathscr{U}$ – a modification of Lemma 3.1 in [3] – we obtain the following Bellman's dynamic programming condition for optimality:

Theorem 3.1. Suppose that $h > 0$ and $0 \leqq s \leqq s + h \leqq T$. Then for each $\mathscr{U}_0^{s+h}$ holds

$$V(s, u) \leqq \mathsf{E}^u \left\{ \int_s^{s+h} c_t^u \, \mathrm{d}t + \sum_{s \leqq t \leqq s+h} r_t^u + V(s + h, u) / \mathscr{G}_s \right\} \quad P^u\text{–a.s.} \tag{7}$$

In particular, u^* is optimal iff the equality in (7) holds for $u = u^*$ and for all $s, h > 0$.

Let us now consider a special form of $\mathscr{G}$ as follows:

Assume given a mapping $\varphi : I \to \hat{I} = \{1, \ldots, n\}$, $n \leqq m$, and let $Y_t = \varphi(X_t)$ be the observable process,

$$\mathscr{G}_t = \mathscr{F}_t^\wedge = \sigma(Y_s, s \leqq t), \quad \mathscr{F}^\wedge = \{\mathscr{F}_t^\wedge, 0 \leqq t \leqq T\},$$

$\hat{N}_j(t)$ be the counting process of jumps of $\{Y_t, 0 \leqq t \leqq T\}$ into the state $j \in \hat{I}$ on $[0, t]$.

Lemma 3.1. $\{\hat{N}_j(t) - \hat{A}_j(t), 0 \leqq t \leqq T\}, j \in \hat{I}$, are $(\mathscr{F}^\wedge, P^u)$–martingales, where

$$\hat{A}_j(t) = \int_0^t \mathsf{E}^u\{\hat{q}(X_s, j, u_s) / \mathscr{F}_s^\wedge\} \, \mathrm{d}s \tag{8}$$

with

$$\hat{q}(i, j, z) = 0 \quad \text{for} \quad \varphi(i) = j, \quad z \in Z,$$
$$\hat{q}(i, j, z) = \sum_{k : \varphi(k) = j} q(i, k, z) \quad \text{for} \quad \varphi(i) \neq j, \quad i \in I. \tag{9}$$

Proof. Put

$$S_j = \{i : \varphi(i) = j\} \subset I.$$

It is clear that

$$\hat{N}_j(t) = \sum_{\tau_n \leqq t} \chi(\varphi(X_{\tau_n -}) \neq j) \, \chi(\varphi(X_{\tau_n}) = j) =$$
$$= \sum_{i \in S_j} \sum_{\tau_n \leqq t} \chi(\varphi(X_{\tau_n -}) \neq j) \, \chi(X_{\tau_n} = i) = \sum_{i \in S_j} \int_0^t \chi(\varphi(X_{s-}) \neq j) \, \mathrm{d}N_i(s),$$

hence

$$\hat{N}_j(t) - \int_0^t \sum_{i \in S_j} \chi(\varphi(X_{s-}) \neq j) \, q(X_{s-}, i, u_s) \, \mathrm{d}s =$$
$$= \sum_{i \in S_j} \int_0^t \chi(\varphi(X_{s-}) \neq j) \, [\mathrm{d}N_i(s) - q(X_s, i, u_s) \, \mathrm{d}s]$$

is an $\mathscr{F}$-martingale.

According to Theorem 18.3, [5] $\{\hat{N}_j(t) - \hat{A}_j(t),\ 0 \leqq t \leqq T\}$ is an $\mathscr{F}^{\wedge}$-martingale. □

The following theorem gives us a sufficient condition for optimality.

Theorem 3.2. Suppose that there is a constant J^* and to each $u \in \mathscr{U}$ exist an $\mathscr{F}^{\wedge}$-adapted function $v(t, u)$ and $\mathscr{F}^{\wedge}$-predictable function $W_j(t, u)$, $t \in [0, T]$, $j \in \hat{I}$, satisfying the following conditions:

(i) $$\mathsf{E}^u \int_0^T |v(t, u)| \, dt < \infty\,, \quad \mathsf{E}^u \int_0^T |W_j(t, u)| \, q(X_t, j, u_t) \, dt < \infty\,,$$

(ii) $$J^* + \int_0^T v(t, u) \, dt + \sum_{j \in \hat{I}} \int_0^T W_j(t, u) \, d\hat{N}_j(t) = 0\,, \quad P^u\text{-a.s.},$$

(iii) $$0 \leqq v(t, u) + \sum_{j \in \hat{I}} W_j(t, u) \, \mathsf{E}^u\{\hat{q}(X_t, j, u_t) / \mathscr{F}_t^{\wedge}\} + \mathsf{E}^u\{c^*(t, X_t, u_t) / \mathscr{F}_t^{\wedge}\} = \\ = G(t, u)\,,$$

say, for almost all $t \in [0, T]$, where

(10) $$c^*(t, i, z) = c(t, i, z) + \sum_{j \in I} q(i, j, z) \, r(t, i, j, z)\,.$$

Then $J^* \leqq J(u)$ for all u. If u^* is a control such that $G(t, u^*) = 0$, then u^* is an optimal control and in this case

(11) $$V(s, u^*) = J^* + \int_0^s v(t, u^*) \, dt + \sum_{j \in \hat{I}} \int_0^s W_j(t, u^*) \, d\hat{N}_j(t)$$

is the value function corresponding to u^*.

Proof. Take $u \in \mathscr{U}$. It follows from (ii) and (iii) that

$$0 = J^* + \mathsf{E}^u \int_0^T [v(t, u) + \sum_{j \in \hat{I}} W_j(t, u) \, \hat{q}(X_t, j, u_t)] \, dt + \\ + \mathsf{E}^u \sum_{j \in \hat{I}} \int_0^T W_j(t, u) \, [d\hat{N}_j(t) - \hat{q}(X_t, j, u_t) \, dt] = \\ = J^* + \mathsf{E}^u \int_0^T [v(t, u) + \sum_{j \in \hat{I}} W_j(t, u) \, \mathsf{E}^u\{\hat{q}(X_t, j, u_t) / \mathscr{F}_t^{\wedge}\}] \, dt \geqq \\ \geqq J^* - \mathsf{E}^u \int_0^T \mathsf{E}^u\{c^*(t, X_t, u_t) / \mathscr{F}_t^{\wedge}\} \, dt = J^* - \mathsf{E}^u \int_0^T c^*(t, X_t, u_t) \, dt = \\ = J^* - \mathsf{E}^u \left[\int_0^T c_s^u \, ds + \sum_{0 \leqq s \leqq T} r_s^u \right] = J^* - J(u)\,.$$

The before last equality follows from the fact that

$$(12)\qquad \mathsf{E}^u\Big[\sum_{0\leqq t\leqq T} r(t, X_{t-}, X_t, u_t)\Big] = \mathsf{E}^u \sum_{i\in I}\int_0^T r(t, X_{t-}, i, u_t)\,\mathrm{d}N_i(t) =$$

$$= \sum_{i\in I} \mathsf{E}^u \int_0^T r(t, X_t, i, u_t)\, q(X_t, i, u_t)\,\mathrm{d}t\,.$$

Thus $J^* \leqq J(u)$.

It is clear that the inequality becomes equality if $G(s, u) = 0$ at $u = u^*$. Hence $J^* = J(u^*)$, i.e. u^* is an optimal control.

Further, for $V(s, u^*)$ defined by the right-hand side of (11) we have

$$V(s, u^*) = \mathsf{E}^{u^*}\{V(s, u^*)/\mathscr{F}^\wedge_s\} =$$

$$= -\mathsf{E}^u\left\{\int_s^T v(t, u^*)\,\mathrm{d}t + \sum_{j\in I}\int_s^T W_j(t, u^*)\,\mathrm{d}\hat{N}_j(t)/\mathscr{F}^\wedge_s\right\} =$$

$$= -\mathsf{E}^u\left\{\int_s^T v(t, u^*)\,\mathrm{d}t + \sum_{j\in I}\int_s^T W_j(t, u)\,\mathsf{E}^{u^*}\{\hat{q}(X_t, j, u_t^*)/\mathscr{F}^\wedge_t\}\,\mathrm{d}t/\mathscr{F}^\wedge_s\right\} =$$

$$= \int_s^T \mathsf{E}^{u^*}\{c^*(t, X_t, u_t^*)/\mathscr{F}^\wedge_s\}\,\mathrm{d}t =$$

$$= \mathsf{E}^{u^*}\left\{\int_s^T c_t^{u^*}\,\mathrm{d}t + \sum_{s\leqq t\leqq T} r_t^{u^*}/\mathscr{F}^\wedge_s\right\},\quad \text{by (12)}\,.$$

On the other hand by (iii)

$$V(s, u^*) \leqq E^{u^*v}\left\{\int_s^T c_t^v\,\mathrm{d}t + \sum_{s\leqq t\leqq T} r_t^v/\mathscr{F}^\wedge_s\right\} = J_s(u^*, v)$$

for all $v \in \mathscr{U}_s^T$. Therefore u^* is an optimal control and

$$V(s, u^*) = \inf_{v\in\mathscr{U}_s T} J_s(u^*, v)$$

is the value function corresponding to the control u^*. □

We shall show below that functions $v(t, u)$, $W_j(t, u)$, $j \in I$, and constant J^* satisfying conditions (i)–(iii) of Theorem 3.2 always exist, and the value function has the form (11) not only for optimal control, but also for arbitrary admissible controls. The result is more general than a similar result in [1] in which the representation holds only for the so-called increasing value controls.

Theorem 3.3. There exists a constant J^*, and to each $u \in \mathscr{U}$ exist $\mathscr{F}^\wedge$-adapted function $v(t, u)$ and $\mathscr{F}^\wedge$-predictable functions $W_j(t, u)$, $j \in I$, $t \in [0, T]$ satisfying conditions (i) – (iii) of Theorem 3.2 such that the value function $V(t, u)$ can be written in the form

$$V(t, u) = J^* + \int_0^t v(s, u)\, \mathrm{d}s + \sum_{j \in I} \int_0^t W_j(t, u)\, \mathrm{d}\hat{N}_j(t) \tag{13}$$

Further, a control u^* is optimal iff $G(t, u^*) = 0$ $(\mathrm{d}t \times \mathrm{d}P^{u^*})$-a.s., where $G(t, u)$ is defined in (iii) of Theorem 3.2.

Proof. Since c and r are bounded, there is constant C_1 such that $|c^*(t, X_t, u_t)| \leqq \leqq C_1$ for all $t \in [0, T]$. Choosing a constant $C_2 (C_2 > C_1)$ and considering the process

$$\bar{V}(s, u) = V(s, u) + C_2 s$$

we obtain by (7) and (12)

$$\bar{V}(s, u) - \mathsf{E}^u\{\bar{V}(s + h, u)/\mathscr{F}_s^\wedge\} \leqq \tag{14}$$

$$\leqq \mathsf{E}^u \left\{ \int_s^{s+h} [c^*(t, X_t, u_t) - C_2]\, \mathrm{d}t/\mathscr{F}_s^\wedge \right\} < 0$$

for each $u \in \mathscr{U}_0^{s+h}$, $0 \leqq s < s + h \leqq T$. Thus $\bar{V}(s, u)$ is a submartingale wrt $(P^u, \mathscr{F}^\wedge)$. In order to obtain the Doob-Meyer decomposition for submartingale $\bar{V}(s, u)$ we have to prove that the mapping $s \to \mathsf{E}^u\, \bar{V}(s, u)$ is continuous, and hence $\bar{V}(t, u)$ has a right continuous modification.

By definition, for each $u \in \mathscr{U}_0^{t+h}$ such that $u = u_0$ on $[0, t]$, $u = u_1$ on $(t, t + h]$,

$$V(t, u) = \bigwedge_{v \in \mathscr{U}_t{}^T} \mathsf{E}^{uv} \left\{ \int_t^T c^*(t, X_s, v_s)\, \mathrm{d}s/\mathscr{F}_t^\wedge \right\} \geqq \tag{15}$$

$$\geqq \bigwedge_{v_1 \in \mathscr{U}_t{}^{t+h}} \mathsf{E}^{u_0 v_1} \left\{ \int_t^{t+h} c^*(s, X_s, v_1)\, \mathrm{d}s/\mathscr{F}_t^\wedge \right\} +$$

$$+ \bigwedge_{v_1 \in \mathscr{U}_t{}^{t+h}} \bigwedge_{v_2 \in \mathscr{U}_{t+h}{}^T} \mathsf{E}^{u_0 v_1 v_2} \left\{ \int_{t+h}^T c^*(s, X_s, v_2)\, \mathrm{d}s/\mathscr{F}_t^\wedge \right\} \geqq$$

$$\geqq -C_1 h + \bigwedge_{v_1, v_2} \mathsf{E}^{u_0 v_1, v_2} \left\{ \int_{t+h}^T c^*(s, X_s, v_2)\, \mathrm{d}s/\mathscr{F}_t^\wedge \right\}.$$

Note that for each $u_0 \in \mathscr{U}_0^t$; $u_1, v_1 \in \mathscr{U}_t^{t+h}$, $v_2 \in \mathscr{U}_{t+h}^T$ we have

$$R = \mathsf{E}^{u_0 v_1 v_2}\left\{\int_{t+h}^{T} c^*(s, X_s, v_2(s))\,\mathrm{d}s/\mathscr{F}_t^\wedge\right\} -$$

$$- \mathsf{E}^{u_0 u_1 v_2}\left\{\int_{t+h}^{T} c^*(s, X_s, v_2(s))\,\mathrm{d}s/\mathscr{F}_t^\wedge\right\} =$$

$$= \frac{\mathsf{E}\left\{\varrho_0^t(u_0)\,\varrho_t^{t+h}(v_1)\,\varrho_{t+h}^T(v_2)\int_{t+h}^{T} c^*(s, X_s, v_2)\,\mathrm{d}s/\mathscr{F}_t^\wedge\right\}}{\mathsf{E}\{\varrho_0^t(u_0)\,\varrho_t^T(v_1 v_2)/\mathscr{F}_t^\wedge\}} -$$

$$- \frac{\mathsf{E}\left\{\varrho_0^t(u_0)\,\varrho_t^{t+h}(u_1)\,\varrho_{t+h}^T(v_2)\int_{t+h}^{T} c^*(s, X_s, v_2)\,\mathrm{d}s/\mathscr{F}_t^\wedge\right\}}{\mathsf{E}\{\varrho_0^t(u_0)\,\varrho_t^T(u_1 v_2)/\mathscr{F}_t^\wedge\}} =$$

$$= \frac{\mathsf{E}\left\{\varrho_0^t(u_0)\,\varrho_{t+h}^T(v_2)\,[\varrho_t^{t+h}(v_1) - \varrho_t^{t+h}(u_1)]\int_{t+h}^{T} c^*(s, X_s, v_2)\,ds/\mathscr{F}_t^\wedge\right\}}{\mathsf{E}\{\varrho_0^t(u_0)/\mathscr{F}_t^\wedge\}} =$$

$$= \frac{S}{\mathsf{E}\{\varrho_0^t(u_0)/\mathscr{F}_t^\wedge\}}$$

where S denotes the numerator of the above fraction. Further,

$$(17)\qquad |S| \leqq C\,\mathsf{E}\{\varrho_0^t(u_0)\,|\varrho_t^{t+h}(v_1) - \varrho_t^{t+h}(u_1)|\,\mathsf{E}[\varrho_{t+h}^T(v_2)/\mathscr{F}_{t+h}]/\mathscr{F}_t^\wedge\} =$$

$$= C\,\mathsf{E}\{\varrho_0^t(u_0)\,|\varrho_t^{t+h}(v_1) - \varrho_t^{t+h}(u_1)|/\mathscr{F}_t^\wedge\} \leqq$$

$$\leqq C\,\mathsf{E}\{\varrho_0^t(u_0)\,\mathsf{E}[|\varrho_t^{t+h}(v_1) - 1| + |\varrho_t^{t+h}(u_1) - 1|/\mathscr{F}_t]/\mathscr{F}_t^\wedge\}$$

where C is some constant.

Let us now estimate the conditional expectation

$$\mathsf{E}\{|\varrho_t^{t+h}(v) - 1|/\mathscr{F}_t\} \quad \text{for} \quad v \in \mathscr{U}_t^{t+h}.$$

According to (2), (3) we have

$$(18)\qquad \mathsf{E}\{|\varrho_t^{t+h}(v) - 1|/\mathscr{F}_t\} = \mathsf{E}\left\{\left|\exp\left(\int_t^{t+h}[m - 1 - \sum_{i\in I} q(X_s, i, v_s)]\,\mathrm{d}s + \right.\right.\right.$$

$$\left.\left.\left. + \sum_{i\in I}\int_t^{t+h}\log q(X_{s-}, i, v_s)\,\mathrm{d}N_i(s)\right) - 1\right|/\mathscr{F}_t\right\} =$$

$$= \mathsf{E}|\exp\left(C_3\theta_1 h + C_3\theta_2 \sum_{i\in I}[N_i(t + h) - N_i(t)]\right) - 1|/\mathscr{F}_t\}$$

where θ_1, θ_2 are $\mathscr{F}_{t+h}$-measurable random variables satisfying $|\theta_1|, |\theta_2| \leqq 1$, whereas

C_3 is some constant dominating $|m - 1 - \sum_{j \in I} q(i, j, z)|$, $|\log q(i, j, z)|$ for all $i \in I$, $i \neq j \in I$ and $z \in Z$.

It is easy to verify that for any real numbers $h_1 \geqq 0$, $h_2 \geqq 0$ holds

$$|\exp(\theta_1 h_1 + \theta_2 h_2) - 1| \leqq |\exp(h_1 + h_2) - 1| .$$

Applying this inequality to (18) we obtain

(19)
$$\mathsf{E}\{|\varrho_t^{t+h}(v) - 1| / \mathscr{F}_t\} \leqq \mathsf{E}\{|\exp(C_3 h + C_3 \sum_{i \in I} [N_i(t + h) - N_i(t)] - 1| / \mathscr{F}_t\} .$$

Further, note that

$$\sum_{i \in I} \left[N_i(t) - \int_0^t (1 - \chi(X_s = i)) \, \mathrm{d}s \right] = \sum_{i \in I} N_i(t) - (m - 1) t$$

is an $(\mathscr{F}, P)$-martingale. Thus $\sum_{i \in I} N_i(t)$ is a Poisson process with intensity $m - 1$, hence $\sum_{i \in I} [N_i(t + h) - N_i(t)]$ has under P the Poisson distribution with parameter $(m - 1)h$ and is independent of $\mathscr{F}_t$. Consequently, from (19) follows

(20)
$$\mathsf{E}\{|\varrho_t^{t+h}(v) - 1| / \mathscr{F}_t\} \leqq \sum_{j=0}^{\infty} [\exp(C_3(h + j)) - 1] [(m - 1) h]^j \frac{\mathrm{e}^{-h(m-1)}}{j!} =$$

$$= \exp((C_3 + (m - 1)(\exp(C_3) - 1)) h) - 1 = \alpha(h) = O(h) .$$

It results from (17), (18) that $|S| \leqq 2C \alpha(h) \mathsf{E}\{\varrho_0^t(u) / \mathscr{F}_t^\wedge\}$. Hence $|R| \leqq C_4 h$, uniformly in v.

From (15), (16) and the above inequality we obtain

$$V(t, u) \geqq -(C_1 + C_4) h + \bigwedge_{v_2 \in \mathscr{U}_{t+h}{}^T} \mathsf{E}^{u_0 u_1} \left\{ \int_{t+h}^T c^*(s, X_s, v_2(s)) \, \mathrm{d}s / \mathscr{F}_t^\wedge \right\} =$$

$$= -(C_1 + C_4) h + \mathsf{E}^u\{V(t + h, u) / \mathscr{F}_t^\wedge\}$$

with notice that $u = u_0 u_1 \in \mathscr{U}_0^{t+h}$, or

(21) $$0 \geqq \bar{V}(t, u) - \mathsf{E}^u\{\bar{V}(t + h, u) / \mathscr{F}_t^\wedge\} \geqq -(C_1 + C_4 + C_2) h = -Ch .$$

Consequently,

$$|\mathsf{E}^u \, \bar{V}(t, u) - \mathsf{E}^u \, \bar{V}(t + h, u)| \to 0 \quad \text{as} \quad h \to 0 .$$

This completes the proof of continuity of $\mathsf{E}^u \, \bar{V}(t, u)$.

According to Theorem T. 4. (iii) in [6], submartingale $\bar{V}(t, u)$ has a right continuous modification and for simplicity we also use the notation $\bar{V}(t, u)$ for the modification.

Further, Theorem T. 31 in [6] allows us to represent $\bar{V}(t, u)$ in the form

$$\bar{V}(t, u) = J^* + R_t^u + M_t^u \tag{22}$$

where M_t^u is an $\mathscr{F}^\wedge$-martingale, R_t^u is an $\mathscr{F}^\wedge$-adapted natural increasing process, $J^* = \bar{V}(0, u) = V(0, u)$. Further,

$$R_t^u = w\text{-}\lim_{h \to 0} \int_0^t h^{-1} \mathsf{E}^u\{\bar{V}(s + h, u) - \bar{V}(s, u)/\mathscr{F}_s^\wedge\} \, \mathrm{d}s \tag{23}$$

where w-lim is the limit meant in the sense of the $\sigma(L_1, L_\infty)$-topology, i.e. $\xi_h(\omega) \xrightarrow{w} \xi(\omega)$ in $L_1(P^u)$ iff for any bounded random variable $g(\omega)$ defined on Ω holds

$$\int_\Omega \xi_h(\omega) \, g(\omega) \, \mathrm{d}P^u(\omega) \to \int_\Omega \xi(\omega) \, g(\omega) \, \mathrm{d}P^u , \quad \text{as} \quad h \to 0 .$$

Let $\xi_h(\omega, s)$ be the integrand in (23). By (21), $|\xi_h(\omega, s)| \leqq C$. Thus for each $s \in [0, T]$ $\{\xi_h(\omega, s), \ 0 \leqq h \leqq T\}$ is a set of uniformly integrable random variables, hence, by the same method as in the proof of Lemma 4.1 in [3] we obtain

$$R_t^u = \int_0^t \beta(s, u) \, \mathrm{d}s \tag{24}$$

where $\beta(s, u) = w\text{-}\lim_{h \downarrow 0} \xi_h(\omega, s)$.

Further, according to Theorem 5.4 of Jacod [4], $\mathscr{F}^\wedge$-martingale M_t^u can be represented in the form

$$M_t^u = \sum_{j \in I} \int_0^t W_j(s, u) \, [\mathrm{d}\hat{N}_j(s) - \mathsf{E}^u\{\hat{q}(X_s, j, u_s)/\mathscr{F}_s^\wedge\} \, \mathrm{d}s] \tag{25}$$

where $W_j(s, u)$ is $\mathscr{F}^\wedge$-predictable and

$$\mathsf{E}^u \int_0^t |W_j(s, u)| \, \hat{q}(X_s, j, u_s) \, ds < \infty .$$

It follows from (22), (24), (25) that

$$\bar{V}(t, u) = J^* + \int_0^t \beta(s, u) \, \mathrm{d}s + \sum_{j \in I} \int_0^t W_j(s, u) \, [\mathrm{d}\hat{N}_j(s) - \mathsf{E}^u\{\hat{q}(X_s, j, u_s)/\mathscr{F}_s^\wedge\} \, \mathrm{d}s] ,$$

hence

$$V(t, u) = \bar{V}(t, u) - \int_0^t C_2 \, \mathrm{d}s = J^* + \int_0^t [v(s, u) \, \mathrm{d}s + \sum_{j \in I} W_j(s, u) \, \mathrm{d}\hat{N}_j(s)]$$

with

$$v(s, u) = \beta(s, u) - \sum_{j \in I} W_j(s, u)\, \mathsf{E}^u\{\hat{q}(X_s, j, u_s)/\mathscr{F}_s^\wedge\} - C_2 .$$

Thus $V(t, u)$ has always the form (13) for all $u \in \mathscr{U}$. To prove condition (iii) and the last assertion we employ the representation (13) and Theorem 3.1, according to which we obtain

$$V(t, u) - \mathsf{E}^u\{V(t + h, u)/\mathscr{F}_t^\wedge\} =$$

$$= -\mathsf{E}^u\left\{\int_t^{t+h} v(s, u)\, \mathrm{d}s + \sum_{j \in I} \int_t^{t+h} W_j(t, u)\, \mathrm{d}\hat{N}(s)/\mathscr{F}_t^\wedge\right\} =$$

$$= -\mathsf{E}^u\left\{\int_t^{t+h} v(s, u)\, \mathrm{d}s + \sum_{j \in I} \int_t^{t+h} W_j(s, u)\, \mathsf{E}^u\{\hat{q}(X_s, j, u_s)/\mathscr{F}_s^\wedge\}\, \mathrm{d}s/\mathscr{F}_t^\wedge\right\} \leqq$$

$$\leqq -\mathsf{E}^u\left\{\int_t^{t+h} c(s, X_s, u_s)\, \mathrm{d}s + \sum_{i \in I} \int_t^{t+h} r(s, X_s, u_s)\, q(X_s, i, u_s)\, \mathrm{d}s/\mathscr{F}_t^\wedge\right\}$$

or equivalently

$$h^{-1}\, \mathsf{E}^u\left\{\int_t^{t+h} G(s, u)\, \mathrm{d}s/\mathscr{F}_t^\wedge\right\} \geqq 0 . \tag{26}$$

Let $g(\omega)$ be arbitrary bounded random variable. It is easy to see that

$$\mathsf{E}^u\left[g h^{-1}\, \mathsf{E}^u\left\{\int_t^{t+h} G(s, u)\, \mathrm{d}s/\mathscr{F}_t^\wedge\right\}\right] = h^{-1} \int_t^{t+h} \mathsf{E}^u[g\, \mathsf{E}^u\{G(s, u)/\mathscr{F}_t^\wedge\}]\mathrm{d}s \to$$

$$\to \mathsf{E}^u[g\, G(t, u)] \quad \text{as} \quad h \downarrow 0 \quad \text{for almost all} \quad t \in [0, T] .$$

Consequently, by (26), $G(t, u) \geqq 0$ $(\mathrm{d}t \times \mathrm{d}P^u)$-a.s..

Further, it is clear that u^* is an optimal control iff

$$\mathsf{E}^{u^*}\left\{\int_t^{t+h} G(s, u^*)\, \mathrm{d}s/\mathscr{F}_t^\wedge\right\} = 0$$

and hence, by the same argument as above, $G(s, u^*) = 0$ $(\mathrm{d}t \times \mathrm{d}P^{u^*})$-a.s.. This completes the proof of Theorem 3.3. □

Remark. Put $p_i^u(t) = P^u\{X_t = i/\mathscr{F}_t^\wedge\}$, $i \in I$. p_i^u is the aposteriori probability of state i. Then the function $G(t, u)$ defined in Theorem 3.2 can be represented in the form

$$G(t, u) = v(t, u) + \sum_{i \in I} \left[\sum_{j \in I} W_j(t, u)\, \hat{q}(i, j, u_t) + c(t, i, u_t)\right] p_i^u(t) .$$

ACKNOWLEDGEMENT

The author would like to express his thank to Dr. P. Mandl for very much helpful discussion.

(Received July 28, 1980.)

REFERENCES

[1] R. Boel, P. Varaiya: Optimal control of jump processes. SIAM J. Control and Optimization *15* (1977), 1, 92—119.

[2] M. Davis, R. Elliott: Optimal control of a jump process. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete *40* (1977), 183—202.

[3] M. H. A. Davis, P. Varaiya: Dynamic programming conditions for partially observable stochastic systems. SIAM J. Control *11* (1973), 2, 226—261.

[4] J. Jacod: Multivariate point processes: Predictable projection, Randon-Nikodym derivatives representation of martingales. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete *31* (1975), 2, 235—253.

[5] R. S. Liptser, A. N. Shiryayev: Statistics of Random Processes I, II. Springer-Verlag, New York—Heidelberg—Berlin 1977, 1978.

[6] P. A. Meyer: Probability and Potentials. Blaisdell, Waltham, Mass. 1966.

[7] S. R. Pliska: Controlled jump processes. Stochastic Processes and Appl. *3*, (1975), 3, 259—282.

[8] R. W. Rishel: The minimum principle, separation principle, and dynamic programming for partially observed jump process. IEEE Trans. Autom. Control *AC-23* (1978), 6, 1009 to 1014.

[9] C. B. Wan, M. H. A. Davis: Existence of optimal controls for stochastic jump processes. SIAM J. Control and Optimization *17* (1979), 4, 511—524.

Nguyen van Huu, CSc., matematicko-fyzikální fakulta University Karlovy (Faculty of Mathematics and Physics — Charles University), Sokolovská 83, 186 00 Praha 8. Czechoslovakia. Permanent address: *Khoa Toan, Dại Hoc Tong Hop (Faculty of Mathematics — Hanoi University), Hanoi. Vietnam.*