APPLICATIONES MATHEMATICAE
22,1 (1993), pp. 25–38

Ł. STETTNER (Warszawa)

ERGODIC CONTROL OF PARTIALLY OBSERVED MARKOV PROCESSES WITH EQUIVALENT TRANSITION PROBABILITIES

Abstract. Optimal control with long run average cost functional of a partially observed Markov process is considered. Under the assumption that the transition probabilities are equivalent, the existence of the solution to the Bellman equation is shown, with the use of which optimal strategies are constructed.

1. Introduction. Let $(\Omega, \mathcal{F}, P)$ be a probability space and (x_n) a discrete time controlled Markov process on a compact state space E, endowed with the Borel σ-field $\mathcal{E}$, with transition kernel $P^v(x, dz)$ for $v \in U$, where $(U,\mathcal{U})$ is a compact space of control parameters. Assume the only observations of x_n are $\mathbb{R}^d$-valued random variables $y_1, \ldots, y_n$ such that for $\mathcal{Y}_n = \sigma\{y_1, \ldots, y_n\}$ we have

$$P\{y_{n+1} \in A \mid x_{n+1}, \mathcal{Y}_n\} = P\{y_{n+1} \in A \mid x_{n+1}\} = \int_A r(x_{n+1}, y)\, dy \tag{1}$$

for $n = 0, 1, \ldots$ with $r : E \times \mathbb{R}^d \to \mathbb{R}^+$ a measurable function, and $A \in \mathcal{B}(\mathbb{R}^d)$, the family of Borel subsets of $\mathbb{R}^d$.

The Markov process (x_n) is controlled by a sequence (a_n) of $\mathcal{Y}_n$-measurable U-valued random variables. The best mean square approximation of x_n based on the available observation is given by a filtering process π_n, defined as a measure valued process such that for $A \in \mathcal{E}$,

$$\pi_n(A) = P\{x_n \in A \mid \mathcal{Y}_n\} \quad \text{for } n = 1, 2, \ldots, \tag{2}$$

and

$$\pi_0(A) = \mu(A)$$

where μ is the initial law of (x_n).

1991 *Mathematics Subject Classification*: Primary 93E20; Secondary 93E11.
Key words and phrases: stochastic control, partial observation, long run average cost, Bellman equation.

The following lemma gives the most general formula for π_n. Its proof, unlike those in [5] and [8], which have more restrictive hypotheses, is not based on the reference probability method.

Lemma 1. *Under* (1), *for* $n = 0, 1, \ldots$ *and* $A \in \mathcal{E}$ *we have*

$$\pi_{n+1}(A) = \frac{\int_A r(z_2, y_{n+1}) \int_E P^{a_n}(z_1, dz_2)\, \pi_n(dz_1)}{\int_E r(z_2, y_{n+1}) \int_E P^{a_n}(z_1, dz_2)\, \pi_n(dz_1)}. \tag{3}$$

Proof. Denote the right hand side of (3) by $M^{a_n}(y_{n+1}, \pi_n)(A)$. Let $F : (\mathbb{R}^d)^n \to \mathbb{R}$ be a bounded measurable function, $Y_n = (y_1, \ldots, y_n)$ and $C \in \mathcal{B}(\mathbb{R}^d)$.

By (1), Fubini's theorem and properties of conditional expectations we have

$$\begin{aligned}
&\int_\Omega M^{a_n}(y_{n+1}, \pi_n)(A)\chi_C(y_{n+1})F(Y_n)\, dP \\
&\quad= \int_\Omega E[M^{a_n}(y_{n+1}, \pi_n)(A)\chi_C(y_{n+1}) \mid x_{n+1}, \mathcal{Y}_n]F(Y_n)\, dP \\
&\quad= \int_\Omega \int_C M^{a_n}(y, \pi_n)(A) r(x_{n+1}, y)\, dy\, F(Y_n)\, dP \\
&\quad= \int_\Omega \int_C M^{a_n}(y, \pi_n)(A) E[E[r(x_{n+1}, y) \mid \mathcal{Y}_n, x_n] \mid \mathcal{Y}_n]\, dy\, F(Y_n)\, dP \\
&\quad= \int_\Omega \int_C M^{a_n}(y, \pi_n)(A) E\Big[\int_E r(z, y)\, P^{a_n}(x_n, dz) \Big| \mathcal{Y}_n\Big]\, dy\, F(Y_n)\, dP \\
&\quad= \int_\Omega \int_C M^{a_n}(y, \pi_n)(A) \int_E \int_E r(z, y)\, P^{a_n}(z_1, dz)\, \pi_n(dz_1)\, dy\, F(Y_n)\, dP \\
&\quad= \int_\Omega \int_C \int_A r(z_2, y) \int_E P^{a_n}(z_1, dz_2)\, \pi_n(dz_1)\, dy\, F(Y_n)\, dP \\
&\quad= \int_\Omega \int_E \int_A \int_C r(z_2, y)\, dy\, P^{a_n}(z_1, dz_2)\, \pi_n(dz_1)\, F(Y_n)\, dP \\
&\quad= \int_\Omega E\Big[\int_A \int_C r(z_2, y)\, dy\, P^{a_n}(x_n, dz_2) \Big| \mathcal{Y}_n\Big] F(Y_n)\, dP \\
&\quad= \int_\Omega E\Big[E\Big[\int_C r(x_{n+1}, y)\, dy\, \chi_A(x_{n+1}) \Big| \mathcal{Y}_n, x_n\Big] \Big| \mathcal{Y}_n\Big] F(Y_n)\, dP \\
&\quad= \int_\Omega \int_C r(x_{n+1}, y)\, dy\, \chi_A(x_{n+1}) F(Y_n)\, dP
\end{aligned}$$

$$= \int_\Omega E[\chi_C(y_{n+1}) \mid \mathcal{Y}_n, x_{n+1}]\chi_A(x_{n+1})F(Y_n)\,dP$$

$$= \int_\Omega \chi_C(y_{n+1})\chi_A(x_{n+1})F(Y_n)\,dP = \int_\Omega \pi_{n+1}(A)\chi_C(y_{n+1})F(Y_n)\,dP.$$

Therefore, by the definition of conditional expectation, (3) follows. ■

The class of controls $a_n = u(\pi_n)$, where u is a fixed measurable, U-valued function, is of special interest. Namely, we have

LEMMA 2. *Under* (1), *if additionally* $a_n = u(\pi_n)$ *with* u *a fixed measurable function from the space* $\mathcal{P}(E)$ *of probability measures on* E, *endowed with the topology of weak convergence*, *into* $(U,\mathcal{U})$, *then* π_n *is a* $\mathcal{Y}_n$*-Markov process with transition operator*

$$\Pi^{u(\nu)}(\nu, F) = \int_E \int_{\mathbb{R}^d} F(M^{u(\nu)}(y,\nu))r(z,y)\,dy \int_E P^{u(\nu)}(z_1,dz)\,\nu(dz_1) \tag{4}$$

where

$$M^v(y,\nu)(A) = \frac{\int_A r(z,y)\int_E P^v(z_1,dz)\,\nu(dz_1)}{\int_E r(z,y)\int_E P^v(z_1,dz)\,\nu(dz_1)} \tag{5}$$

for $v \in U$, $\nu \in \mathcal{P}(E)$ *and* $F : \mathcal{P}(E) \to \mathbb{R}$ *bounded measurable.*

Proof. By (1) we easily obtain

$$\begin{aligned}
E[F(\pi_{n+1}) \mid \mathcal{Y}_n]&\\
&= E[F(M^{u(\pi_n)}(y_{n+1},\pi_n)) \mid \mathcal{Y}_n]\\
&= E[E[F(M^{u(\pi_n)}(y_{n+1},\pi_n)) \mid \mathcal{Y}_n, x_{n+1}] \mid \mathcal{Y}_n]\\
&= E\Big[\int_{\mathbb{R}^d} F(M^{u(\pi_n)}(y,\pi_n))r(x_{n+1},y)\,dy \,\Big|\, \mathcal{Y}_n\Big]\\
&= E\Big[\int_{\mathbb{R}^d} E[F(M^{u(\pi_n)}(y,\pi_n))r(x_{n+1},y) \mid \mathcal{Y}_n, x_n]\,dy \,\Big|\, \mathcal{Y}_n\Big]\\
&= E\Big[\int_{\mathbb{R}^d}\int_E F(M^{u(\pi_n)}(y,\pi_n))r(z,y)P^{u(\pi_n)}(x_n,dz)\,dy \,\Big|\, \mathcal{Y}_n\Big]\\
&= \int_E\int_{\mathbb{R}^d} F(M^{u(\pi_n)}(y,\pi_n)) \int_E r(z,y)\,P^{u(\pi_n)}(z_1,dz)\,\pi_n(dz_1)\,dy\\
&= \Pi^{u(\pi_n)}(\pi_n,F)\,.
\end{aligned}$$

Thus (π_n) is Markov with transition operator of the form (4). ■

In this paper we are interested in minimizing the following long run average cost functional:

$$J_\mu((a_n)) = \limsup_{n\to\infty} n^{-1} E_\mu\Big\{\sum_{i=0}^{n-1} c(x_i, a_i)\Big\} \tag{6}$$

over all U-valued, $\mathcal{Y}_n$-adapted processes a_n, with $c : E \times U \to \mathbb{R}^+$ a given bounded measurable cost function.

By the very definition of a filtering process we have

$$J_\mu((a_n)) = \limsup_{n\to\infty} n^{-1} E_\mu\Big\{\sum_{i=0}^{n-1} \int_E c(z, a_i)\, \pi_i(dz)\Big\}. \tag{7}$$

The optimal strategies for the cost functional J_μ are constructed with the use of a suitable Bellman equation, the solution of which is found as a limit of $w^\beta(x) = \vartheta^\beta(x) - \inf_{z\in E} \vartheta^\beta(z)$ as $\beta \to 1$, where ϑ^β is the value function of the β-discounted cost functional. Since our limit results are based on compactness arguments, obtained via the Ascoli–Arzelà theorem, in Section 2 we show the continuity of ϑ^β. Then in Section 3 we prove the uniform boundedness of w^β. Using the concavity of w^β, obtained from the concavity of ϑ^β, proved in Section 2, we get equicontinuity of w^β, which allows us to use the Ascoli–Arzelà theorem.

The discrete time ergodic optimal control problem with partial observation was studied in [1], [2], [3], [6], [8], [9]. In [1] and [8] the observation was corrupted with white noise. In addition, in [1] there was a finite state space and a rich observation structure. In [8] the state space was general but there were some restrictions on controls. The papers [2] and [3] contain a general theory but the fundamental example used is a very simple maintenance-replacement model.

In [6] a model with a finite state space and almost steady state transition probabilities was studied. Finally, finite state space semi-Markov decision processes with a completely observable state were considered in [9]. Our paper generalizes [6] in various directions. Namely, we have a general, compact state space. Although the techniques to show the boundedness and the equicontinuity of w^β follow in some sense the arguments of [6], by a more detailed estimation we obtain the results under the assumptions which are much less restrictive than the corresponding ones in [6], even when E is finite.

2. Discounted control problem. In this section we characterize the value function ϑ^β of the discounted cost functional J_μ^β defined as follows:

$$J_\mu^\beta((a_n)) \stackrel{\text{def}}{=} E_\mu\Big\{\sum_{i=0}^{\infty} \beta^i c(x_i, a_i)\Big\} = E_\mu\Big\{\sum_{i=0}^{\infty} \beta^i \int_E c(z, a_i)\, \pi_i(dz)\Big\} \tag{8}$$

with $\beta \in (0, 1)$.

The theorem below provides a complete solution to the discounted partially observed control problem.

THEOREM 1. *Assume* (1) *and*

(A1) $c : E \times U \to \mathbb{R}^+$ *is continuous*,

(H1) *for* $F \in C(\mathcal{P}(E))$, *the space of continuous functions on* $\mathcal{P}(E)$, *if* $\mu_n \Rightarrow \mu$, *i.e.* μ_n *converges weakly in* $\mathcal{P}(E)$ *to* μ, *we have*

$$\sup_{a\in U} |\Pi^a(\mu_n, F) - \Pi^a(\mu, F)| \to 0 \quad \text{as } n \to \infty, \tag{9}$$

(H2) *for* $F \in C(\mathcal{P}(E))$, *if* $U \ni a_n \to a$ *we have*

$$\Pi^{a_n}(\mu, F) \to \Pi^a(\mu, F). \tag{10}$$

Then

$$\vartheta^\beta(\mu) \overset{\text{def}}{=} \inf_{(a_n)} J^\beta_\mu((a_n)) \tag{11}$$

is a continuous function of $\mu \in \mathcal{P}(E)$ *and is a unique solution to the Bellman equation*

$$\vartheta^\beta(\mu) = \inf_{a\in U} \Big[\int_E c(x, a)\, \mu(dx) + \beta\Pi^a(\mu, \vartheta^\beta)\Big]. \tag{12}$$

There exists a measurable selector $u^\beta : \mathcal{P}(E) \to (U, \mathcal{U})$ *for which the infimum on the right hand side of* (12) *is attained. Moreover*, *we have*

$$\vartheta^\beta(\mu) = J^\beta_\mu((u^\beta(\pi_n)))\,. \tag{13}$$

In addition, ϑ^β *can be uniformly approximated from below by the sequence*

$$\begin{aligned} &\vartheta^\beta_0(\mu) \equiv 0\,,\\ &\vartheta^\beta_{n+1}(\mu) = \inf_{a\in U} \Big[\int_E c(x, a)\, \mu(dx) + \beta\Pi^a(\mu, \vartheta^\beta_n)\Big], \end{aligned} \tag{14}$$

and each ϑ^β_n *is concave*, *i.e. for* $\mu, \nu \in \mathcal{P}(E)$ *and* $\alpha \in [0, 1]$,

$$\vartheta^\beta_n(\alpha\mu + (1-\alpha)\nu) \ge \alpha\vartheta^\beta_n(\mu) + (1-\alpha)\vartheta^\beta_n(\nu). \tag{15}$$

Proof. We only point out the main steps since the proof is more or less standard (for details see [4], Thm. 2.2).

Define, for $\vartheta \in C(\mathcal{P}(E))$,

$$T\vartheta(\mu) = \inf_{a\in U} \Big[\int_E c(x, a)\, \mu(dx) + \beta\Pi^a(\mu, \vartheta)\Big].$$

By (A1) and (H1), T is a contraction on $C(\mathcal{P}(E))$. Thus, by the Banach principle there is a unique fixed point ϑ^β of T, which is a unique solution to the Bellman equation (12). Since by (A1) and (H2) the map

$$U \ni a \to \int_E c(x,a)\,\mu(dx) + \beta \Pi^a(\mu, \vartheta^\beta)$$

is continuous, there exists a measurable selector u^β. The identity (13) is then almost immediate. Since T is monotonic and contractive, ϑ_n^β is increasing and converges to ϑ^β. It remains to show the concavity of ϑ_n^β. We prove this by induction. Clearly, $\vartheta_0^\beta \equiv 0$ is concave. Provided ϑ_n^β is concave, by Jensen's lemma we have for $\alpha \in (0,1)$,

$$\Pi^a(\alpha\mu + (1-\alpha)\nu, \vartheta_n^\beta) \geq \alpha \Pi^a(\mu, \vartheta_n^\beta) + (1-\alpha)\Pi^a(\nu, \vartheta^\beta)$$

and therefore from (14),

$$\vartheta_{n+1}^\beta(\alpha\mu + (1-\alpha)\nu) \geq \alpha\vartheta_{n+1}^\beta(\mu) + (1-\alpha)\vartheta_{n+1}^\beta(\nu)\,,$$

i.e. ϑ_{n+1}^β is concave. By induction, ϑ_n^β is concave for each n. The proof of the theorem is complete. ■

Below we formulate sufficient conditions for (H1) and (H2).

Proposition 1. *Assume*

(A2) $r \in C(E \times \mathbb{R}^d)$,

(A3) *for fixed* $a \in U$, $P^a(x, \cdot\,)$ *is Feller, i.e. for any* $\varphi \in C(E)$, *if* $x_n \Rightarrow x$ *we have*

$$P^a(x_n, \varphi) \to P^a(x, \varphi), \tag{16}$$

(H3) *if* $U \ni a_n \to a$, *then for each* $\varphi \in C(E)$,

$$\sup_{x\in E} |P^{a_n}(x,\varphi) - P^a(x,\varphi)| \to 0, \tag{17}$$

(A4) *for* $R(z,\psi) \overset{\text{def}}{=} \int_{\mathbb{R}^d} r(z,y)\psi(y)\,dy$ *where* $\psi \in C(\mathbb{R}^d)$, *if* $E \ni z_n \to z$ *we have*

$$R(z_n, \cdot\,) \Rightarrow R(z, \cdot\,). \tag{18}$$

Then (H1) *and* (H2) *are satisfied.*

Proof. Notice first that from (16) and (17), if $U \ni a_n \to a$ and $\mu_n \Rightarrow \mu$, we have

$$P^{a_n}(\mu_n, \varphi) \overset{\text{def}}{=} \int_E P^{a_n}(x,\varphi)\,\mu_n(dx) \to P^a(\mu,\varphi) \quad \text{as } n \to \infty, \text{ for } \varphi \in C(E)\,. \tag{19}$$

Since $U \times \mathcal{P}(E)$ is compact, to prove (H1) and (H2) it is sufficient to show that

$$U \times \mathcal{P}(E) \ni (a,\mu) \to \Pi^a(\mu, F) \text{ is continuous for } F \in C(\mathcal{P}(E)).$$

Therefore we shall show that

$$\Pi^{a_n}(\mu_n, F) \to \Pi^a(\mu, F) \tag{20}$$

for $U \ni a_n \to a$, $\mathcal{P}(E) \ni \mu_n \Rightarrow \mu$ and $F \in C(\mathcal{P}(E))$. We have

$$\begin{aligned} &|\Pi^{a_n}(\mu_n, F) - \Pi^a(\mu, F)| \\ &\quad \le \Big| \int_E \int_{\mathbb{R}^d} (F(M^{a_n}(y, \mu_n)) - F(M^a(y, \mu)))r(z, y)\, dy\, P^{a_n}(\mu_n, dz)\Big| \\ &\qquad + \Big| \int_E \int_{\mathbb{R}^d} F(M^a(y, \mu))r(z, y)\, dy\, (P^{a_n}(\mu_n, dz) - P^a(\mu, dz))\Big| \\ &\quad = I_n + II_n\,. \end{aligned} \tag{21}$$

From (19), $II_n \to 0$, provided

$$E \ni z \to \int_{\mathbb{R}^d} F(M^a(y, \mu))r(z, y)\, dy \in C(E). \tag{22}$$

By (A4), $\mathbb{R}^d \ni y \to M^a(y, \mu) \in \mathcal{P}(E)$ is continuous. Then, again by (A4), the map (22) is continuous, and consequently $II_n \to 0$.

If

$$\sup_{z \in E} \Big| \int_{\mathbb{R}^d} (F(M^{a_n}(y, \mu_n)) - F(M^a(y, \mu)))r(z, y)\, dy\Big| \to 0 \tag{23}$$

then clearly $I_n \to 0$.

By (A4), for each $\varepsilon > 0$ there exists a compact set $K \subset \mathbb{R}^d$ such that for any $z \in E$,

$$R(z, K^c) < \frac{\varepsilon}{2\|F\|}. \tag{24}$$

Therefore

$$\begin{aligned} &\Big| \int_{\mathbb{R}^d} (F(M^{a_n}(y, \mu_n)) - F(M^a(y, \mu)))r(z, y)\, dy\Big| \\ &\qquad\qquad \le \int_K |F(M^{a_n}(y, \mu_n)) - F(M^a(y, \mu))| r(z, y)\, dy + \varepsilon \end{aligned}$$

and to obtain (23) it remains to show that

$$M^{a_n}(y, \mu_n)(\varphi) \to M^a(y, \mu)(\varphi) \tag{25}$$

for any $\varphi \in C(E)$, uniformly in $y \in K$.

Using the Stone–Weierstrass approximation theorem (see [7], Thm. 9.28, cf. also the proof of Lemma A.1.2 of [8]) and (19), we obtain

$$\Big|\int_E r(z,y)\varphi(z)\int_E P^{a_n}(z_1,dz)\,\mu_n(dz_1)-\int_E r(z,y)\varphi(z)\int_E P^a(z_1,dz)\,\mu(dz_1)\Big|$$
$$\le\Big|\int_E r(z,y)\varphi(z)(P^{a_n}(\mu_n,dz)-P^a(\mu,dz))\Big|\to 0$$

uniformly in $y\in K$. Thus, we have uniform convergence of the numerators and denominators in the formula defining M^{a_n}, and consequently convergence of the ratios from which (25) follows.

The proof of Proposition 1 is complete. ■

Remark 1. (A4) is satisfied when $\sup_{z\in E} r(z,y)$ is integrable.

Define

$$w^\beta(\nu)=\vartheta^\beta(\nu)-\vartheta^\beta(\mu_\beta)\quad\text{and}\quad w^\beta_n(\nu)=\vartheta^\beta_n(\nu)-\vartheta^\beta_n(\mu^n_\beta)\tag{26}$$

where $\mu_\beta=\arg\min\vartheta^\beta$ and $\mu^n_\beta=\arg\min\vartheta^\beta_n$. Clearly, w^β is a solution to the equation

$$w^\beta(\nu)+(1-\beta)\vartheta^\beta(\mu_\beta)=\inf_{a\in U}\Big[\int_E c(x,a)\,\nu(dx)+\beta\Pi^a(\nu,w^\beta)\Big]\tag{27}$$

and $w^\beta_n(\nu)\to w^\beta(\nu)$ uniformly in $\nu\in\mathcal{P}(E)$. We would like to let $\beta\uparrow 1$ in (27) and thus obtain a solution $w(\nu)$ to the long run average Bellman equation

$$w(\nu)+\gamma=\inf_{a\in U}\Big[\int_E c(x,a)\,\nu(dx)+\Pi^a(\nu,w)\Big].\tag{28}$$

Since we wish to apply the Ascoli–Arzelà theorem, we have to show the boundedness and the equicontinuity of w^β for $\beta\in(0,1)$, which are studied successively in the next sections.

3. Boundedness of w^β.

We make the following assumption:

$$\text{(A5)}\quad \inf_{z,z'\in E}\ \inf_{a,a'\in U}\ \inf_{C\in\mathcal{E},\,P^a(z,C)>0}\frac{P^{a'}(z',C)}{P^a(z,C)}\stackrel{\text{def}}{=}\lambda>0.\tag{29}$$

We have

PROPOSITION 2. *Under* (A5) *and the assumptions of Theorem* 1, *the functions $w^\beta(\nu)$ are uniformly bounded for $\beta\in(0,1)$, $\nu\in\mathcal{P}(E)$.*

Proof. We improve the proof of Theorem 2 of [6]. Namely, we show by induction the uniform boundedness of $w^\beta_n(\nu)$ for $\nu\in\mathcal{P}(E)$, $\beta\in(0,1)$, $n=0,1,\ldots$ For $n=0$, $w^\beta_0(\nu)\equiv 0$.

Assume that for any $\beta\in(0,1)$, $\nu\in\mathcal{P}(E)$, $w^\beta_n(\nu)\le L$ where $L\ge\|c\|\lambda^{-2}$.

Let $a, a' \in U$ be such that for fixed $\nu \in \mathcal{P}(E)$,

$$w_{n+1}^\beta(\nu) = \int_E c(x,a)\,\nu(dx) - \int_E c(x,a')\,\mu_\beta^{n+1}(dx) + \beta[\Pi^a(\nu,\vartheta_n^\beta) - \Pi^{a'}(\mu_\beta^{n+1},\vartheta_n^\beta)]\,. \tag{30}$$

For $y \in \mathbb{R}^d$, define

$$m(y)(B) = M^{a'}(y,\mu_\beta^{n+1})(B) - \lambda^2 M^a(y,\nu)(B)$$

for any $B \in \mathcal{E}$.

By (29) we have

$$\begin{aligned}\int_B r(z,y) \int_E P^{a'}(z_1,dz)\,\mu_\beta^{n+1}(dz_1) &\ge \lambda \int_B r(z,y) \int_E P^a(z_1,dz)\,\nu(dz_1)\\ &= \lambda M^a(y,\nu)(B) \int_E r(z,y) \int_E P^a(z_1,dz)\,\nu(dz_1)\\ &\ge \lambda^2 M^a(y,\nu)(B) \int_E r(z,y) \int_E P^{a'}(z_1,dz)\,\mu_\beta^{n+1}(dz_1)\end{aligned}$$

and therefore $m(y)(B) \ge 0$ for $B \in \mathcal{E}$.

If $\lambda = 1$ we have a stationary, noncontrolled Markov chain with $P^a(z,C) = \eta(C)$ for any $a \in U$, $z \in E$ and some fixed $\eta \in \mathcal{P}(E)$, and consequently $w_n^\beta \equiv 0$ for any $n = 0,1,\ldots$ Therefore we restrict ourselves to the case $\lambda < 1$. Then $(1-\lambda^2)^{-1}m(y) \in \mathcal{P}(E)$. Since

$$M^{a'}(y,\mu_\beta^{n+1}) = \lambda^2 M^a(y,\nu) + (1-\lambda^2)[(1-\lambda^2)^{-1}m(y)]\,,$$

by concavity of ϑ_n^β we obtain

$$\vartheta_n^\beta(M^{a'}(y,\mu_\beta^{n+1})) \ge \lambda^2\vartheta_n^\beta(M^a(y,\nu)) + (1-\lambda^2)\vartheta_n^\beta((1-\lambda^2)^{-1}m(y)) \tag{31}$$

and from (30) we have

$$\begin{aligned}&w_{n+1}^\beta(\nu)\\ &\quad\le \|c\| + \beta \int_E \int_{\mathbb{R}^d} \vartheta_n^\beta(M^a(y,\mu))r(z,y)\,dy\\ &\qquad\times \Big(\int_E P^a(z_1,dz)\,\nu(dz_1) - \lambda^2 \int_E P^{a'}(z_1,dz)\,\mu_\beta^{n+1}(dz_1)\Big)\\ &\qquad - \beta(1-\lambda^2) \int_E \int_{\mathbb{R}^d} \vartheta_n^\beta((1-\lambda^2)^{-1}m(y))r(z,y)\,dy\\ &\qquad\times \int_E P^{a'}(z_1,dz)\mu_\beta^{n+1}(dz_1)\end{aligned} \tag{32}$$

$$
\begin{aligned}
&= \|c\| + \beta \int\limits_E \int\limits_{\mathbb{R}^d} (\vartheta_n^\beta(M^a(y,\mu)) - \vartheta_n^\beta(\mu_\beta^n)) r(z,y)\, dy \\
&\quad \times \Big(\int\limits_E P^a(z_1, dz)\, \nu(dz_1) - \lambda^2 \int\limits_E P^{a'}(z_1, dz)\, \mu_\beta^{n+1}(dz_1) \Big) \\
&\quad - \beta(1-\lambda^2) \int\limits_E \int\limits_{\mathbb{R}^d} (\vartheta_n^\beta((1-\lambda^2)^{-1} m(y)) - \vartheta_n^\beta(\mu_\beta^n)) r(z,y)\, dy \\
&\quad \times \int\limits_E P^{a'}(z_1, dz)\, \mu_\beta^{n+1}(dz_1) \\
&\le \|c\| + \beta L \operatorname{var} \Big(\int\limits_E P^a(z_1, \cdot\,)\, \nu(dz_1) - \lambda^2 \int\limits_E P^{a'}(z_1, \cdot\,)\, \mu_\beta^{n+1}(dz_1) \Big).
\end{aligned}
$$

By (A5) for any $B \in \mathcal{E}$,

$$
\int\limits_E P^a(z_1, B)\, \nu(dz_1) \ge \lambda^2 \int\limits_E P^{a'}(z_1, B)\, \mu_\beta^{n+1}(dz_1)\,. \tag{33}
$$

Thus

$$
w_{n+1}^\beta(\nu) \le \|c\| + \beta L(1-\lambda^2) \le L \tag{34}
$$

and the bound L is independent of $\nu \in \mathcal{P}(E)$, $\beta \in (0,1)$. By induction $w_n^\beta(\nu) \le L$ for any $\nu \in \mathcal{P}(E)$, $n = 0,1,\ldots$, $\beta \in (0,1)$. Since by the very definition $w_n^\beta(\nu) \ge 0$, and for each β, $w_n^\beta(\nu) \to w^\beta(\nu)$ as $n \to \infty$, we finally obtain $w^\beta(\nu) \le L$ for $\nu \in \mathcal{P}(E)$ and $\beta \in (0,1)$. ∎

Remark 2. One can easily see that in the case of a finite state space E, the assumption

$$
\text{(A5}'\text{)} \quad \inf_{z,z' \in E} \inf_{a,a' \in U} \inf_{x \in E,\, P^a(z,x) > 0} \frac{P^{a'}(z',x)}{P^a(z,x)} > 0 \tag{35}
$$

also implies the boundedness of w^β. Thus Proposition 2 significantly improves Theorem 2 of [6]. This was possible because of the choice of μ_β^n in (32) as the argument of minimum of ϑ_n^β.

Remark 3. Assumption (A5) says that the transition probabilities for different controls and initial states are mutually equivalent, with Radon–Nikodym density bounded away from 0. In particular, in the case when $P^a(z,C) = \int_C g^a(z,x)\, \eta(dx)$ the assumption

$$
\inf_{z,z' \in E} \inf_{a,a' \in U} \inf_{x \in E,\, g^a(z,x) > 0} \frac{g^{a'}(z',x)}{g^a(z,x)} > 0 \tag{36}
$$

is sufficient for (A5) to be satisfied.

4. Main theorem. Before we formulate and prove our main result, we show the equicontinuity of w^β for $\beta \in (0,1)$. For this purpose we need an extra assumption:

(A6) If $\mathcal{P}(E) \ni \mu_n \Rightarrow \mu \in \mathcal{P}(E)$ then
$$\sup_{a\in U} \sup_{C\in\mathcal{E}} |P^a(\mu_n, C) - P^a(\mu, C)| \to 0$$

with
$$P^a(\mu, C) \stackrel{\text{def}}{=} \int_E P^a(x, C)\mu(dx)\,.$$

We have

PROPOSITION 3. *Under* (A5), (A6) *and the assumptions of Theorem* 1, *the family of functions* w^β, $\beta \in (0,1)$, *is equicontinuous*, *i.e.*

$$\forall_{\varepsilon>0}\ \exists_{\delta>0}\ \forall_{\mu,\mu'\in\mathcal{P}(E)}\quad \varrho(\mu,\mu') < \delta \Rightarrow \forall_{\beta\in(0,1)}\ |w^\beta(\mu) - w^\beta(\mu')| < \varepsilon \tag{37}$$

with ϱ *standing for a metric compatible with the weak convergence topology of* $\mathcal{P}(E)$.

Proof. For $\nu, \mu \in \mathcal{P}(E)$ let

$$\lambda(\nu,\mu) \stackrel{\text{def}}{=} \inf_{a\in U} \inf_{C\in\mathcal{E}, P^a(\mu,C)>0} \frac{P^a(\nu, C)}{P^a(\mu, C)}\,. \tag{38}$$

From (A5) and (A6), if $\nu \Rightarrow \mu$, then

$$\lambda(\nu,\mu) \to 1 \quad \text{and} \quad \lambda(\mu,\nu) \to 1. \tag{39}$$

By (27) for $\nu, \mu \in \mathcal{P}(E)$ we have

$$\begin{aligned} &w^\beta(\nu) - w^\beta(\mu) \\ &\quad\le \sup_{a\in U} \Big| \int_E c(x,a)(\nu(dx) - \mu(dx))\Big| + \beta \sup_{a\in U}(\Pi^a(\nu, w^\beta) - \Pi^a(\mu, w^\beta)). \end{aligned} \tag{40}$$

By analogy with the proof of Proposition 2 define

$$m^a(y,\mu,\nu)(B) = M^a(y,\mu)(B) - \lambda(\mu,\nu)\lambda(\nu,\mu)M^a(y,\nu)(B)$$

for $B \in \mathcal{E}$.

Clearly, $m^a(y,\mu,\nu)(B) \ge 0$ for $B \in \mathcal{E}$, and $\lambda(\mu,\nu)\lambda(\nu,\mu) \le 1$.

If $\lambda(\mu,\nu)\lambda(\nu,\mu) = 1$, then $w^\beta \equiv 0$ for $\beta \in (0,1)$, and consequently the equicontinuity property is satisfied. Therefore assume $\lambda^2 = \lambda(\mu,\nu)\lambda(\nu,\mu) < 1$. Then by the concavity of w^β,

$$w^\beta(M^a(y,\mu)) \ge \lambda^2 w^\beta(M^a(y,\nu)) + (1-\lambda^2)w^\beta((1-\lambda^2)^{-1}m^a(y,\mu,\nu))\,. \tag{41}$$

From (40),

$$w^\beta(\nu) - w^\beta(\mu) \le \sup_{a\in U}\Big|\int_E c(x,a)(\nu(dx)-\mu(dx))\Big| \tag{42}$$
$$+\,\beta\sup_{a\in U}\Big\{\int_E\int_{\mathbb{R}^d} w^\beta(M^a(y,\nu))r(z,y)\,dy\,(P^a(\nu,dz)-\lambda^2P^a(\mu,dz))$$
$$+\int_E\int_{\mathbb{R}^d}(\lambda^2w^\beta(M^a(y,\nu))-w^\beta(M^a(y,\mu)))r(z,y)\,dy\,P^a(\mu,dz)\Big\}$$
$$= I + II + III.$$

Now

$$\begin{aligned} II &\le 2\|w^\beta\|\sup_{a\in U}\sup_{B\in\mathcal{E}}|P^a(\nu,B)-\lambda^2P^a(\mu,B)| \\ &= 2\|w^\beta\|(1-\lambda(\mu,\nu)\lambda(\nu,\mu)) \end{aligned} \tag{43}$$

and using (41) and the nonnegativity of w^β we have

$$\begin{aligned} &III \\ &\le \sup_{a\in U}\int_E\int_{\mathbb{R}^d}(\lambda^2-1)w^\beta((1-\lambda^2)^{-1}m^a(y,\mu,\nu))r(z,y)\,dy\,P^a(\mu,dz)\le 0\,. \end{aligned} \tag{44}$$

Interchanging ν and μ in (40)–(44) we obtain the same estimates and therefore

$$\begin{aligned} |w^\beta(\nu)-w^\beta(\mu)| \le \sup_{a\in U}\Big|\int_E c(x,a)(\nu(dx)-\mu(dx))\Big| \\ +\,2\|w^\beta\|(1-\lambda(\mu,\nu)\lambda(\nu,\mu))\,. \end{aligned} \tag{45}$$

Since by the Stone–Weierstrass theorem (Thm. 9.28 of [7]) $c(x,a)$ can be uniformly approximated on $E\times U$ by continuous functions of the form $\sum_{i=1}^r c_i(x)d_i(a)$, from (39) we obtain

$$\lim_{\nu\Rightarrow\mu}\sup_{\beta\in(0,1)}|w^\beta(\nu)-w^\beta(\mu)| = 0\,. \ \blacksquare$$

Let us comment on the assumption (A6):

Remark 4. (H3) clearly follows from (A6).

Remark 5. In the case of a finite state space $E=\{1,\ldots,N\}$, (A6) can be written as

$$\sup_{a\in U}\sum_{k=1}^N\Big|\sum_{i=1}^N(s_i^n-s_i)P^a(i,k)\Big|\to 0 \tag{46}$$

for $s^n=(s_1^n,\ldots,s_N^n)\to s=(s_1,\ldots,s_N)$, $0\le s_i^n\le 1$, $0\le s_i\le 1$, $\sum s_i^n=1$, $\sum s_i=1$, and this is satisfied since

$$\sup_{a\in U}\sum_{k=1}^N\Big|\sum_{i=1}^N(s_i^n-s_i)P^a(i,k)\Big|\le\sum_{i=1}^N|s_i^n-s_i|\to 0\quad\text{as } s^n\to s\,.$$

Remark 6. Assume $P^a(z,C)=\int_C g^a(z,x)\eta(dx)$ for $C\in\mathcal{E}$ and that the mapping

$$U\times E\times E\ni(a,z,x)\to g^a(z,x)\text{ is continuous.}\tag{47}$$

Then (A6) is satisfied.

In fact, by the Stone–Weierstrass theorem we can approximate g^a uniformly on $U\times E\times E$ by continuous functions of the form $\sum_{i=1}^k b_i(a)c_i(z)d_i(x)$ and

$$\begin{aligned}&\sup_{a\in U}\sup_{C\in\mathcal{E}}|P^a(\mu_n,C)-P^a(\mu,C)|\\ &\quad\le\sup_{a\in U}\int_E\Big|\int_E g^a(z,x)(\mu_n(dz)-\mu(dz))\Big|\eta(dx)\\ &\quad\le\varepsilon+\sum_{i=1}^k\sup_{a\in U}|b_i(a)|\int_E|d_i(x)|\eta(dx)\Big|\int_E c_i(z)(\mu_n(dz)-\mu(dz))\Big|\to\varepsilon\end{aligned}$$

as $n\to\infty$.

Now we can prove our main result:

THEOREM 2. *Assume* (A1)–(A6). *Then there exist $w\in C(\mathcal{P}(E))$ and a constant γ which are solutions to the Bellman equation*

$$w(\mu)+\gamma=\inf_{a\in U}\Big[\int_E c(x,a)\,\mu(dx)+\Pi^a(\mu,w)\Big].\tag{48}$$

Moreover, there exists $u:\mathcal{P}(E)\to U$ for which the infimum on the right hand side of (48) *is attained. The strategy $a_n=u(\pi_n)$ is optimal for J_μ and*

$$J_\mu((u(\pi_n)))=\gamma\,.\tag{49}$$

Proof. By Theorem 1, each ϑ_n^β is concave. Therefore w_n^β is concave and w^β as limit of w_n^β is also concave. Since by Proposition 2, the w^β are uniformly bounded, and by Proposition 3 equicontinuous, from the Ascoli–Arzelà theorem the family w^β, $\beta\in(0,1)$, is relatively compact in $C(\mathcal{P}(E))$. Moreover, $|(1-\beta)\vartheta^\beta(\mu_\beta)|\le\|c\|$. Therefore one can choose a subsequence $\beta_k\to 1$ such that

$$(1-\beta_k)\vartheta_k^\beta(\mu_{\beta_k})\to\gamma$$

and

$$w^{\beta_k}\to w\quad\text{ in } C(\mathcal{P}(E))\text{ as } k\to\infty.$$

Letting $\beta_k \to 1$ in (27) we obtain (48).

The remaining assertion of the theorem follows easily from Theorem 3.2.2 of [4]. ■

References

[1] G. B. Di Masi and Ł. Stettner, *On adaptive control of a partially observed Markov chain*, Applicationes Math., to appear.

[2] E. Fernandez-Gaucherand, A. Arapostatis and S. J. Marcus, *Adaptive control of a partially observed controlled Markov chain*, in: Stochastic Theory and Adaptive Control, T. E. Duncan and B. Pasik-Duncan (eds.), Lecture Notes in Control and Inform. Sci. 184, Springer, 1992, 161–171.

[3] —, —, —, *On partially observable Markov decision processes with an average cost criterion*, Proc. 28 CDC, Tampa, Florida, 1989, 1267–1272.

[4] O. Hernandez-Lerma, *Adaptive Markov Control Processes*, Springer, New York, 1989.

[5] H. Korezlioglu and G. Mazziotto, *Estimation recursive en transmission numerique*, Proc. Neuvième Colloque sur le Traitement du Signal et ses Applications, Nice, 1983.

[6] M. Kurano, *On the existence of an optimal stationary J-policy in non-discounted Markovian decision processes with incomplete state information*, Bull. Math. Statist. 17 (1977), 75–81.

[7] H. L. Royden, *Real Analysis*, Macmillan, New York, 1968.

[8] W. J. Runggaldier and Ł. Stettner, *Nearly optimal controls for stochastic ergodic problems with partial observation*, SIAM J. Control Optim. 31 (1993), 180–218.

[9] K. Wakuta, *Semi-Markov decision processes with incomplete state observation—average cost criterion*, J. Oper. Res. Soc. Japan 24 (1981), 95–108.

ŁUKASZ STETTNER
INSTITUTE OF MATHEMATICS
POLISH ACADEMY OF SCIENCES
P.O. BOX 137
00-950 WARSZAWA, POLAND
E-mail: STETTNER@IMPAN.IMPAN.GOV.PL

Received on 22.6.1992