Kybernetika

Petr Mandl; Ma.Rosario Romera Ayllón
On controlled Markov processes with average cost criterion

Terms of use:

ON CONTROLLED MARKOV PROCESSES
WITH AVERAGE COST CRITERION

PETR MANDL, Mª ROSARIO ROMERA AYLLON

An asymptotic lower bound for the probability distribution of the cost functional is presented together with an analogous result related to the arcsine law.

1. INTRODUCTION

A controlled Markov process is a mathematical model of the evolution X_t, $t \geq 0$, of a system S such that the transition law governing the state changes of S depends only on the current state and on a control variable. In the present paper we assume that the state space I and the control spaces $J(i)$ are finite, and that the state changes according to the transition rates

$$(1) \qquad q(i, j; z), \quad i, j \in I, \quad z \in J(i).$$

For general information we refer to [2], but a reader having certain familiarity with the subject need not consult any reference.

We denote by Z_t the value of the control parameter at time t. It has to depend only on the trajectory of S up to time t. Given the initial state, the control $Z = \{Z_t, t \geq 0\}$ together with (1) define the probability distribution of $X = \{X_t, t \geq 0\}$. X_t and Z_t are assumed to be piecewise constant right- and left-continuous, respectively. The left-continuous version of X_t will be denoted by X_t^-. Z of the form

$$(2) \qquad Z_t = z(X_t^-), \quad t \geq 0,$$

will be called stationary.

Introduce the counting processes of jumps

$$N_t^{(i,j,z)} = \sum_{s \leq t} \chi\{X_s^- = i, \ X_s = j, \ Z_s = z\}, \quad t \geq 0,$$

abbreviate $(i, j, z) = k$, and set

$$K = \{k = (i, j, z): q(i, j; z) > 0\} .$$

To evaluate the performance of S under a control Z we define

$$C_T = \int_0^T c(X_t, Z_t) \, dt + \sum_{k \in K} d(k) N_T^k , \quad T \geqq 0 .$$

C_T, which will be called cost up to time T, consists of a continuous part and of the evaluation of the jumps. We shall deal with the controls aiming to minimize the average cost

$$\lim_{T \to \infty} T^{-1} C_T .$$

We make the following hypothesis.

Assumption 1. For any stationary control the matrix

$$\|q(i, j; z(i))\|_{i, j \in I}$$

is indecomposable.

Assumption 1 means that under (2) the Markov process X has recurrent states only.

The optimality equation for our problem can be stated as follows: The minimal average cost θ is the unique constant to which auxiliary constants $w(j)$, $j \in I$, can be found so that

$$\min_{z \in J(i)} \{c(i, z) + \sum_{j \neq i} q(i, j; z) (d(i, j; z) + w(j) - w(i)) - \theta\} = 0 , \quad i \in I .$$

Denoting the expression in the braces by $f(i, z)$, we have

$$f(i, z) \geqq 0 , \quad i \in I , \quad z \in J(i) .$$

Let $z(i)$ fulfil

(3) $$f(i, z(i)) = 0 , \quad i \in I .$$

Then under (2) holds

(4) $$\lim_{T \to \infty} T^{-1} C_T = \theta .$$

We omit the adverb almost surely. On the other hand, under arbitrary control Z

(5) $$\liminf_{T \to \infty} T^{-1} C_T \geqq \theta .$$

Assumption 2. $z(i)$ fulfilling (3) is unique.

In the sequel $z(i)$ will always denote the one mentioned in Assumption 2.

Let us observe that (4) and (5) characterize the optimality of (2) in the sense of the law of large numbers. In the present paper we show that there is a meaningful optimality with respect to the central limit theorem and with respect to the arcsine law. Results obtained previously for linear systems with a quadratic cost (see [3]) are extended to finite-state Markov processes.

434

2. STATEMENT OF RESULTS

Let ξ, η be random variables. The validity of

$$(6) \qquad\qquad \mathsf{P}(\xi \le x) \le \mathsf{P}(\eta \le x), \quad x \in (-\infty, \infty),$$

expresses that ξ is stochastically nonsmaller than η in a sense, which is easily grasped from (6). We shall use an analogous relation to compare the asymptotic distributions.

Denote by $\pi(i), i \in I$, the stationary distribution of X under (2). Introduce

$$\Delta = \sum_{i \neq j} \pi(i) \, q(i, j; z(i)) \, (d(i, j; z(i)) + w(j) - w(i))^2,$$

and assume

$$\Delta > 0.$$

Further set

$$\Phi(x) = \frac{1}{\sqrt{(2\pi)}} \int_{-\infty}^{x} \exp\left\{-\frac{y^2}{2}\right\} dy.$$

Proposition 1. Under arbitrary control Z

(i) $\displaystyle \limsup_{T \to \infty} \mathsf{P}((C_T - \theta T)/\sqrt{T} \le x) \le \Phi\left(\frac{x}{\sqrt{\Delta}}\right), \quad x \in (-\infty, \infty).$

If Z is such that

$$(7) \qquad\qquad \lim_{T \to \infty} \frac{1}{\sqrt{T}} \int_0^T \chi\{Z_t \neq z(X_t^-)\} \, dt = 0 \quad \text{in probability},$$

then

(ii) $\displaystyle \lim_{T \to \infty} \mathsf{P}((C_T - \theta T)/\sqrt{T} \le x) = \Phi\left(\frac{x}{\sqrt{\Delta}}\right), \quad x \in (-\infty, \infty).$

A more general version of the second assertion is contained in [4].

Consider next the proportion of time spent by C_t above θt, $t \in [0, T]$, namely

$$B_T = \frac{1}{T} \int_0^T \chi\{C_t > \theta t\} \, dt.$$

B_T should be small for a good control. The asymptotic lower bound for the distribution of B_T is given in the next proposition.

Proposition 2. Under arbitrary control Z

(iii) $\displaystyle \limsup_{T \to \infty} \mathsf{P}(B_T \le x) \le 2/\pi \arcsin \sqrt{x}, \quad x \in [0, 1].$

If Z is such that (7) holds, then

(iv) $\displaystyle \lim_{T \to \infty} \mathsf{P}(B_T \le x) = 2/\pi \arcsin \sqrt{x}, \quad x \in [0, 1].$

3. EXAMPLE

Proposition 1 provides the asymptotic solution to the following problem: Let $0 < p < 1$, find the minimal value $x_p(T)$ such that

$$P(C_T \leq x_p(T)) \geq p \, .$$

We have

$$x_p(T) \approx \theta T + u_p \sqrt{(\Delta T)} \, ,$$

where $\Phi(u_p) = p$, and $x_p(T)$ is valid for the control (2).

Consider next the queueing network depicted in Figure 1. Two types of jobs arrive to the network as Poisson streams with rates λ_i, $i = 1, 2$. At most two jobs
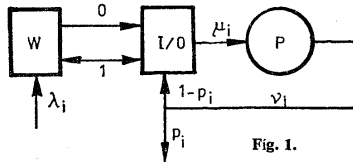

Fig. 1.

can be in the system, one in the waiting line W, and one in the server consisting of an input-output unit, where service rates μ_i apply, and of a processor where the service rates are ν_i. The jobs leaving the processor leave the system with probability p_i. With probability $1 - p_i$ they go back to the I/O unit. The jobs leaving the I/O unit either proceed to the processor (control $z = 0$) or change place with the job in W (control $z = 1$).

The jobs arriving when W is occupied are lost. Denote by N the number of jobs lost during a shift of length T. What bound for N can be guaranteed with 90% reliability?

Set $\lambda = \lambda_1 + \lambda_2$ and let η be the total occupation time of W. To answer the question we have to find the minimal n such that

$$0 \cdot 9 \leq P(N \leq n) = E \sum_{k=0}^{n} \frac{(\lambda \eta)^k}{k!} \, e^{-\lambda \eta} \, .$$

The right-hand side is decreasing with respect to the ordering defined by (6). Thus, Proposition 1 yields the asymptotic solution for $C_T = \eta$.

First it is to be decided whether to deroute the type with a larger average service time to W if a job of the other type is present there. (2) means just the better one of the alternatives. Investigating the characteristic function

$$E \, e^{iuN} = E \exp \{\lambda \eta (e^{iu} - 1)\}$$

it is seen that under (2) N is asymptotically normal with mean $\lambda \theta T$ and variance

$\left(\lambda\theta + \lambda^2\Delta\right) T$ as $T \to \infty$. Consequently,

$$n \approx \lambda\theta T + 1\cdot 282 \sqrt{\left(\lambda\theta + \lambda^2\Delta\right)} \sqrt{T}.$$

For numerical illustration let the average time spent in the I/O unit be 2 sec. and the time spent in the processor be in average for type 1 4 sec. with $1\cdot 25$ cycle repeti-

Table 1.

λ_1	λ_2	θ	Δ	n
0·9	0·1	0·038	0·040	26·0
0·8	0·2	0·063	0·074	40·8
0·7	0·3	0·088	0·104	54·3

tions, for type 2 20 sec. with 2 repetitions. Table 1 presents the quantities θ, Δ and n for $T = 480$ min. in dependence on the proportion of incoming types. The service rule gives preference to type 1.

4. PROOFS

Introduce the random process

$$(8) \qquad M_T = C_T - \theta T + w(X_T) - w(X_0) - \int_0^T f(X_t, Z_t)\,\mathrm{d}t\,, \quad T \geqq 0\,,$$

and note that

$$M_T = \sum_{k \in K} a(k)\left(N_T^k - Q_T^k\right),$$

where for $k = (i, j, z)$

$$a(k) = d(i, j, z) + w(j) - w(i)\,,$$

$$Q_T^k = q(i, j; z) \int_0^T \chi\{X_t = i,\, Z_t = z\}\,\mathrm{d}t\,,$$

Q_T^k is the compensator of N_T^k, and therefore we can construct the representation

$$N_T^k = \mathscr{P}^k(Q_T^k)\,,$$

where

$$\mathscr{P}^k(t)\,, \quad t \geqq 0\,, \quad k \in K\,,$$

are mutually independent Poisson processes with unit rate. The processes

$$\mathscr{M}^k(t) = \mathscr{P}^k(t) - t\,, \quad t \geqq 0\,,$$

are martingales, and

$$(9) \qquad M_T = \sum_{k \in K} a(k)\, \mathscr{M}^k(Q_T^k)\,, \quad T \geqq 0\,.$$

Let us treat the process Q_T^k in the same way as C_T. Take constants Δ^k, $v^k(l)$, $l \in I$, and set

$$(10) \qquad R_T^k = Q_T^k - \Delta^k T + v^k(X_T) - v^k(X_0) - \int_0^T g^k(X_t, Z_t)\,\mathrm{d}t\,,$$

437

where for $k = (i, j, z)$

$$g^k(l, u) = q(i, j; z) \chi\{l = i, u = z\} + \sum_{m \neq l} q(l, m; u) (v^k(m) - v^k(l)) - \Delta^k .$$

We want to have

(11) $$\left| g^k(l, u) \right| \leq \text{const.} f(l, u) , \quad l \in I , \quad u \in J(i) , \quad k \in K .$$

Since

$$f(l, u) > 0 \quad \text{for} \quad u \neq z(l) ,$$

we can set

$$\Delta^k = 0 , \quad v^k(l) = 0 , \quad l \in I , \quad \text{for} \quad k = (i, j, z) , \quad z \neq z(i) ,$$

and determine the constants for $k = (i, j, z(i))$ so that

$$g^k(l, z(l)) = 0 , \quad l \in I .$$

Assumption 1 guarantees that such determination is feasible with

$$\Delta^k = \pi(i) q(i, j; z(i)) .$$

A representation of R_T^k similar to (9), namely

$$R_T^k = \sum_{l \in K} a^k(l) \mathcal{M}^l(Q_T^l) ,$$

yields

(12) $$\text{Var } R_T^k = \sum_{l \in K} a^k(l)^2 \, \mathsf{E} Q_T^l \leq \text{const.} \, T , \quad T \geq 0 , \quad k \in K .$$

Finally, introduce

$$\mathcal{M}_T = \sum_{k \in K} a(k) \mathcal{M}^k(\Delta^k T) , \quad T \geq 0 .$$

It holds

$$\text{Var } \mathcal{M}_T = \sum_{k \in K} a(k)^2 \, \Delta^k T = \Delta T .$$

Lemma 1. The probability distribution of

$$\frac{1}{\sqrt{(\Delta T)}} \, \mathcal{M}_{sT} , \quad s \in [0, 1] ,$$

converges weakly as $T \to \infty$ to the Wiener measure.

The lemma is proved using well known properties of the weak convergence of probability measures.

Lemma 2. For $y > 0 , \quad S > 0 , \quad y/S \leq 2$,

(13) $$\mathsf{P}(\sup_{t \in [0, S]} \left| \mathcal{M}^k(t) \right| \geq y) \leq 2 \exp\left\{ -\frac{y^2}{4S} \right\} .$$

Proof. Introduce the first-passage time

$$\tau = \inf\{t: \mathcal{M}^k(t) \gtrless y\} ,$$

and note that

$$\mathsf{E} \exp\{h \, \mathcal{M}^k(t)\} = \exp\{(e^h - 1 - h) t\} .$$

438

For $h > 0$ Wald's fundamental identity holds

$$E\big[\exp\{h\,\mathcal{M}^k(\tau)\}\exp\{-(e^h - 1 - h)\tau\}\big] = 1\,.$$

Hence,

$$E\big[e^{hy}\exp\{-(e^h - 1 - h)\,\tau\}\big] \leqq 1\,,$$

and consequently,

$$(14) \qquad P(\tau \leqq S) \leqq e^{-hy}\exp\{(e^h - 1 - h)\,S\}\,.$$

Similarly for

$$\sigma = \inf\{t\colon \mathcal{M}^k(t) \leqq -y\}\,,$$

it holds

$$E\big[\exp\{-h\,\mathcal{M}^k(\sigma)\}\exp\{-(e^{-h} - 1 + h)\,\sigma\}\big] = 1\,,$$

and hence,

$$(15) \qquad P(\sigma \leqq S) \leqq e^{-hy}\exp\{(e^{-h} - 1 + h)\,S\}\,.$$

From (14), (15) follows

$$P(\tau \leqq S) + P(\sigma \leqq S) \leqq 2\exp\{-hy + h^2 S\}\,, \quad 0 \leqq h \leqq 1\,.$$

Setting $h = y/(2S)$ we obtain

$$P(\tau \leqq S) + P(\sigma \leqq S) \leqq 2\exp\{-y^2/(4S)\}\,,$$

which implies (13). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Proof of (i). Denote

$$A_T = \int_0^T f(X_t, Z_t)\,\mathrm{d}t\,.$$

In virtue of (11)

$$\Big|\int_0^T g^k(X_t, Z_t)\,\mathrm{d}t\Big| \leqq \text{const.}\, A_T\,, \quad T \geqq 0\,, \quad k \in K\,.$$

Let $\varepsilon > 0$ be arbitrarily small. Because of (10) and (12) we can achieve by taking L large enough that

$$(16) \qquad P(Q_T^k - \varDelta^k T \leqq L(A_T + \sqrt{T}),\, k \in K) \geqq 1 - \varepsilon\,.$$

From (8) for T sufficiently large

$$(17) \qquad P(C_T - \theta T \leqq x\sqrt{T}) \leqq P(M_T + A_T \leqq (x + \varepsilon)\sqrt{T}) \leqq$$

$$\leqq P(\mathcal{M}_T \leqq (x + 2\varepsilon)\sqrt{T}) + P(|M_T - \mathcal{M}_T| > A_T + \varepsilon\sqrt{T})\,.$$

With regard to (16) we have

$$P(|M_T - \mathcal{M}_T| > A_T + \varepsilon\sqrt{T}) \leqq$$

$$\leqq \sum_{k \in K} P\big[\# K|a(k)|\,|\mathcal{M}^k(Q_T^k) - \mathcal{M}^k(\varDelta^k T)| \geqq A_T + \varepsilon\sqrt{T}\big] \leqq$$

$$\leqq \varepsilon + \sum_{k \in K} P\big[\# K|a(k)|\sup_{|t - \varDelta^k T| \leqq L(A_T + \sqrt{T})}|\mathcal{M}^k(t) - \mathcal{M}^k(\varDelta^k T)| \geqq A_T + \varepsilon\sqrt{T}\big]\,,$$

$\# K$ denoting the number of elements in K.

Setting

$$\gamma = \# K|a(k)|\,,$$

and distinguishing the cases $j \sqrt{(T)} \leq A_T < (j + 1) \sqrt{(T)}$ it is concluded that the last probability is majorized by

$$\sum_{j=0}^{\infty} \mathsf{P}\big[\gamma \sup_{|t-\varDelta^k T| \leq L(j+2)\sqrt{T}} |\mathscr{M}^k(t) - \mathscr{M}^k(\varDelta^k T)| \geq (j + \varepsilon) \sqrt{T}\big] \leq$$

$$\sum_{j=0}^{\infty} \mathsf{P}\big[\gamma \sup_{0 \leq s \leq 2L(j+2)\sqrt{T}} |\mathscr{M}^k(s)| \geq \tfrac{1}{2}(j + \varepsilon) \sqrt{T}\big] .$$

Applying Lemma 2 one obtains from here the upper bound

$$2 \sum_{j=0}^{\infty} \exp \left\{ -\frac{(j + \varepsilon)^2 \sqrt{T}}{32\gamma^2 L(j+2)} \right\},$$

which tends to 0 as $T \to \infty$. With regard to Lemma 1 it follows then from (17) that

$$\limsup_{T \to \infty} \mathsf{P}\big(C_T - \theta T \leq x \sqrt{T}\big) \leq \varepsilon + \varPhi \left(\frac{x + 2\varepsilon}{\sqrt{\varDelta}}\right).$$

Letting $\varepsilon \to 0$ we get (i). $\qquad\square$

Proof of (ii). Let (7) hold. Then

(18) $$\lim_{T \to \infty} A_T/\sqrt{(T)} = 0 \quad \text{in probability} .$$

Hence, for T large holds

(19) $$\mathsf{P}\big(A_T \leq \varepsilon \sqrt{T}\big) \geq 1 - \varepsilon .$$

Further,

(20) $$\mathsf{P}\big(C_T - \theta T > x \sqrt{T}\big) \leq \mathsf{P}\big(M_T + A_T > (x - \varepsilon) \sqrt{T}\big) \leq$$

$$\leq \mathsf{P}\big(\mathscr{M}_T > (x - 3\varepsilon) \sqrt{T}\big) + \mathsf{P}\big(|M_T - \mathscr{M}_T| > 2\varepsilon \sqrt{(T)} - A_T\big) .$$

As in the proof of (i), taking (16), (19) into account we obtain

$$\mathsf{P}\big(|M_T - \mathscr{M}_T| > 2\varepsilon \sqrt{(T)} - A_T\big) \leq 2\varepsilon +$$

$$+ \sum_{k \in K} \big[\mathsf{P} \neq K\big] a(k)\big| \sup_{|t-\varDelta^k T| \leq L(\varepsilon+1)\sqrt{T}} |\mathscr{M}^k(t) - \mathscr{M}^k(\varDelta^k T)| \geq \varepsilon \sqrt{T}\big] .$$

For the last probability we have the estimate

$$2 \exp \left\{ -\frac{\varepsilon^2 \sqrt{T}}{4\gamma^2 L(\varepsilon + 1)} \right\}.$$

Letting $T \to \infty$, $\varepsilon \to 0$ we thus get from (20)

$$\limsup_{T \to \infty} \mathsf{P}\big(C_T - \theta T > x \sqrt{T}\big) \leq 1 - \varPhi \left(\frac{x}{\sqrt{\varDelta}}\right),$$

which together with (i) yields (ii). $\qquad\square$

Proof of (iii). Define the function

$$h_b(y) = 0 \quad \text{for} \quad y \leq b , \quad h_b(y) = 1 \quad \text{for} \quad y > b .$$

440

Then

$$B_T = \frac{1}{T} \int_0^T h_0 (C_t - \theta t) \, dt = \int_0^1 h_\varepsilon \left(\frac{1}{\sqrt{T}} \mathcal{M}_{uT} \right) du - \int_0^1 \left(h_\varepsilon \left(\frac{1}{\sqrt{T}} \mathcal{M}_{uT} \right) - \right.$$

$$\left. - h_0 \left(\frac{1}{\sqrt{T}} \left(M_{uT} - w(X_{uT}) + w(X_0) + A_{uT} \right) \right) \right) du \, .$$

Abbreviating the denotation of the last integrands we have for $\varepsilon > 0$ arbitrary, $x \in [0, 1]$,

$$P(B_T \leqq x) \leqq P(\textstyle\int_0^1 h_\varepsilon \, du \leqq x + \varepsilon) + P(\textstyle\int_0^1 (h_\varepsilon - h_0) \, du > \varepsilon) \, .$$

In virtue of Lemma 1 and of the arcsine law for the Wiener process,

$$\lim_{\substack{T \to \infty \\ \varepsilon \to 0}} P \left(\textstyle\int_0^1 h_\varepsilon \, du \leqq x + \varepsilon \right) = 2/\pi \arcsin \sqrt{x} \, .$$

Consequently, to prove (iii) it remains to establish

$$(21) \qquad \lim_{T \to \infty} P(\textstyle\int_0^1 (h_\varepsilon - h_0) \, du > \varepsilon) = 0 \, .$$

For T sufficiently large

$$(22) \qquad h_\varepsilon - h_0 \leqq \chi\{ |M_{uT} - \mathcal{M}_{uT}| > A_{uT} + \tfrac{1}{2}\varepsilon \sqrt{T} \} \, .$$

Let R_u^T denote the right-hand side of (22). It holds

$$E R_u^T = P(|M_{uT} - \mathcal{M}_{uT}| > A_{uT} + \tfrac{1}{2}\varepsilon \sqrt{T}) \, .$$

Applying the second part of the proof of (i) it is inferred that

$$\lim_{T \to \infty} E R_u^T = 0 \, ,$$

and hence,

$$\lim_{T \to \infty} E \textstyle\int_0^1 R_u^T \, du = 0 \, .$$

This implies

$$\lim_{T \to \infty} P(\textstyle\int_0^1 R_u^T \, du > \varepsilon) = 0 \, .$$

From here and from (22) follows (21). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Proof of (iv). Let (7) hold. Consider the inequality

$$(23) \qquad P(B_T \leqq x) \geqq P(\textstyle\int_0^1 h_{-\varepsilon} \, du \leqq x - \varepsilon) - P(\textstyle\int_0^1 (h_0 - h_{-\varepsilon}) \, du > \varepsilon) \, .$$

For T sufficiently large

$$h_0 - h_{-\varepsilon} \leqq \chi\{ |M_u - \mathcal{M}_{uT}| > \tfrac{1}{2}\varepsilon \sqrt{T} - A_T \} \, .$$

Denote the last term by S_u^T. From the proof of (ii) follows that (18) implies

$$\lim_{T \to \infty} E \textstyle\int_0^1 S_u^T \, du = 0 \, .$$

Hence,

$$\lim_{T \to \infty} P(\textstyle\int_0^1 (h_0 - h_{-\varepsilon}) \, du > \varepsilon) = 0 \, .$$

Letting $T \to \infty$, $\varepsilon \to 0$ in (23) one obtains

$$\liminf_{T \to \infty} \mathsf{P}(B_T \leqq x) \geqq 2/\pi \arcsin \sqrt{x} \,.$$

This together with (iii) proves (iv). □

REFERENCES

[1] P. Billingsley: Convergence of Probability Measures. J. Wiley, New York 1968.
[2] P. Mandl: Martingale methods in discrete state random processes. Kybernetika *18* (1982), supplement.
[3] P. Mandl: Limit theorems of probability theory and optimality in linear controlled systems with quadratic cost. In: Proceedings of 5th IFIP Working Conference on Stochastic Differential Systems. (Lecture Notes in Control and Information Sciences 96.) Springer-Verlag, Berlin 1987, 316—329.
[4] M. R. Romera Ayllón: Control adaptivo de procesos de Markov con espacio de estados numerable. Thesis. Universidad Complutense, Madrid 1984.
[5] A. Wald: Sequential Analysis. J. Wiley, New York 1947.

Dr. Petr Mandl, DrSc., matematicko-fyzikální fakulta Univerzity Karlovy (Faculty of Mathematics and Physics, Charles University), Sokolovská 83, 186 00 Praha 8. Czechoslovakia. Prof. Dr. Mª Rosario Romera Ayllón, Facultad de Informática, Universidad Politécnica de Madrid (Faculty of Informatics, Polytechnical University of Madrid), Km. 7 Carretera de Valencia, 280 31 Madrid. Spain.