- [13] J. R. BLUM AND J. R. ROSENBLATT, *On the central limit theorem for the sum of a random number of independent random variables*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete, 1, 4 (1962–1963), pp. 389–393.
- [14] W. RICHTER, *Limit theorems for sequences of random variables with sequences of random indices*, Theory Prob. Applications, 10 1 (1965), pp. 74–89.
- [15] J. MOGYORODI, *A remark on the stable sequences of random variables*, Acta Math. Acad. Sci. Hungar., 17, 3 (1966), pp. 401–409.
- [16] H. SREEHARI, *An invariance principle for random partial sums*, Sankhyā–Ser. A, 30, 4 (1968), pp. 432–442.
- [17] R. PRAKASA, *Random central limit theorems for martingales*, Acta. Math. Acad. Sci. Hungar., 20, 1–2 (1969), pp. 217–222.

ON HOMOGENEOUS MARKOV MODELS WITH CONTINUOUS TIME AND FINITE OR COUNTABLE STATE SPACE

A. A. YUSHKEVIC AND E. A. FAINBERG

(Translated by E. Lukacs)

Without assuming the homogeneity in time, processes indicated in the title of this paper were investigated in [13]. In our note, the results of [13] are applied to the homogeneous case, and the concept of canonical strategy is carried over from [16], [17] to models with continuous time. In contrast to papers [1], [3]–[11], it is assumed that the controls depend on the whole history. We note that, aside from [13] and [14], inhomogeneous models with more general controls were studied by Gikhman and Skorokhod under additional conditions of compactness and continuity ([12], Chapter 2, § 5). The physical aspects of descriptions depending on the past of the strategy were first discussed in [2].

1. The homogeneous model Z is given by the following elements: (i) *the state space X* —a finite or denumerable set; (ii) *the space of controls A* —a Borel space (in § 4, it is an arbitrary measurable space); (iii) *the projection j* —a measurable mapping of A to X , where $A(x) = j^{-1}(x)$ is the set of controls admitting the states x ; (iv) *the transition density $q(a, \Gamma)$* —a measure depending measurably on the state a and satisfying the conditions $q(a, j(a)) = 0$, $q(a, X) \leq K < \infty$; (v) *the rate of payoff $r(a)$* —a measurable function on A ; (vi) *the coefficient of discontinuity α* —a non-negative constant.

The estimate of the policy π and the estimate of the model on the time interval $[s, u]$ are given by the formulas

$$(1) \quad w_s^u(x, \pi, R) = \mathbf{E}_x^\pi \left[\int_s^u e^{-\alpha t} r(\pi[x_s^t]) dt + e^{-\alpha u} R(x_u) \right], \quad v_s^u(x, R) = \sup_{\pi} w_s^u(x, \pi, R),$$

where the final payoff R is a function on X (equal to 0 for $u = \infty$), x_t is a random trajectory of the controlled process, x_s^t is the segment of this trajectory in the time interval $[s, t]$, $\mathbf{E}_x^\pi$ is an integral with respect to the probability measure $\mathbf{P}_x^\pi$ in the space of trajectories which corresponds to the initial state x and the policy π . The formal definitions of the space of trajectories, of the arbitrary policy π and the measure $\mathbf{P}_x^\pi$ are given in [13]. For $s = 0$, $u = \infty$, $R = 0$, these arguments are omitted in the notation of the estimates. The model is *summable (from above or from below)* provided the estimate $w(x, \pi)$ is finite for all x and π if in the model with the function r either its positive or its negative part is substituted for r . The Markovian property is given by the formula $\pi[x_s^t] = \varphi(t, x_t)$ by means of the mapping $\varphi(t; x)$ of the space $[0, \infty) \times X$ on A such that $\varphi(t; x) \in A(x)$ and that $\varphi(\cdot; x)$ is Borel measurable. If φ does not depend on t , then the policy is called *stationary* and the mapping

φ itself is called the *selector*. The policy π is *optimal* if $w(\cdot, \pi) = v$ and it is ε -*optimal* if $w(\cdot, \pi) \geq v - \varepsilon$.

2. We introduce the operators

$$Qf(a) = \sum_{y \in X} q(a, y)f(y) - q(a, X)f(j(a)), \quad Uf = r + Qf,$$

$$Q^\varphi f(x) = Qf(\varphi(x)), \quad T^\varphi f(x) = Uf(\varphi(x)), \quad Vg(x) = \sup_{A(x)} g(a), \quad Tf = VUf$$

($a \in A$, $x \in X$, f is a function on X , g is a function on A , φ is the selector; the expressions $Qf(a)$ and so on are defined if they do not contain the indeterminacy $\infty - \infty$). The norm of functions is understood to be the supremum of their modulus.

Theorem 1. Let $0 \leq t \leq u < \infty$, $\alpha = 0$, $\|Vr\| + \|R\| < \infty$. We put $w_t(\cdot, \cdot) = w_t^u(\cdot, \cdot, R)$, $v_t(\cdot) = v_t^u(\cdot, R)$.

A. The function $v_t(x)$ is continuously differentiable with respect to t and is the unique solution of Bellmann's equation

$$-\frac{\partial v_t}{\partial t} = Tv_t$$

with the boundary condition $v_u = R$.

B. For any Markovian policy $\varphi(t, \cdot) = \varphi_t(\cdot)$ and initial state x , we have

$$(2) \quad v(x) - w(x, \varphi) = \int_0^u \mathbf{M}_x^\varphi(T - T^{\varphi_t})v_t(x_t) dt,$$

C. A Markovian ε -optimal policy exists for any $\varepsilon > 0$.

PROOF. It is clear that $v_t(x) \leq u\|Vr\| + \|R\|$. For any $\delta > 0$ there exists a selector φ with $r(\varphi) > Vr - \delta \geq -\|Vr\| - \delta$. Therefore,

$$v_t(x) \geq w_t^u(x, \varphi, R) \geq -u(\|Vr\| + \delta) - \|R\|.$$

Hence $\|v_t\| \leq u\|Vr\| + \|R\|$. For uniformly bounded v_t , Corollary 4.1, Lemma 5.1 and Theorem 5.1 of [13] yield B and C and the weaker assertion A': $v_t(x)$ is the unique bounded, absolutely continuous solution of the equation

$$-\frac{\partial v_t}{\partial t} = Tv_t \quad (\text{a.e. in } t)$$

with boundary condition $v_u = R$. It remains to show that Tv_t is continuous in t . It follows from the definition of the operators Q , V and T that

$$\|Vr\| - 2K\|f\| \leq V(r - \|Qf\|) \leq Tf \leq \|Vr\| + 2K\|f\|,$$

so that the uniqueness of the estimate $\|Tv_t\| \leq C < \infty$ holds uniformly in t . Hence, by virtue of A', we have $\|v_t - v_s\| \leq C(t - s)$; therefore, $\|Qv_t - Qv_s\| \leq 2KC|t - s|$, and hence,

$$|Tv_t - Tv_s| \leq 2KC|t - s|.$$

Defining v_t as the supremum of $w_t(\cdot, \varphi)$ with respect to the Markovian policies φ , Miller [3] obtained Bellmann's inequality in the form A' for finite, and Kakumanu [6], Lippman [11], for denumerable jump models, and (in the form A) Pliska for semicontinuous jump models. They also proved that the policy $\varphi(t, x)$ is optimal among the Markovian policies if $T^{\varphi_t}v_t = Tv_t$ for all t .

3. Theorem 2. Let $0 \leq t < \infty$. A. If either $r \leq 0$ or the model of summation from above and $v \neq 0$, then Bellmann's equation $\alpha v = Tv$ is satisfied at all points x for which $v(x)$ is finite. For $\alpha > 0$ and $\|Vr\| < \infty$, the estimate v is the unique bounded solution of $\alpha v = Tv$.

B. If the model is summable and $\|v\| < \infty$, then for any Markovian policy $\varphi(t, \cdot) = \varphi_t(\cdot)$,

$$(3) \quad v(x) - w(x, \varphi) = \int_0^\infty e^{-\alpha t} \mathbf{E}_x^\varphi(T - T^{\varphi_t})v(x_t) dt + \lim_{t \rightarrow \infty} e^{-\alpha t} \mathbf{E}_x^\varphi v(x_t).$$

Moreover, the last limit (different from zero only if $\alpha = 0$) is non-negative.

C. If $r \geq 0$, then $v(x) = \sup w(x, \varphi)$ where φ varies over the class of Markovian policies.

D. If $r \leq 0$ and $\|v\| < \infty$, then there exists an ε -optimal policy for any $\varepsilon > 0$.

E. If $\alpha > 0$ and $\|Vr\| < \infty$, then there exists a stationary ε -optimal policy for any $\varepsilon > 0$.

F. If A is a metric space, $A(x)$ is compact, $q(a, y)$ is continuous in a , $r(a)$ is semicontinuous from above, $\|Vr\| < \infty$ and either $\alpha > 0$ or $r \leq 0$ and $\|v\| < \infty$, then there exists a stationary optimal policy.

G. If X and A are finite, the model is summable and v is finite, then there exists a stationary optimal policy.

PROOF. According to Theorem 4.2 from [13], the equation of optimality

$$(4) \quad v_t + K \int_t^u v_s ds = v_u + \int_t^u (KI + T_s)v_s ds, \quad t \leq u,$$

is valid under the assumptions A. Here I is the unit operator and $v_t = e^{-\alpha t}v$ due to the homogeneity of the model. The operator T_t is expressed in terms of T by the formula $T_t(e^{-\alpha t}f) = e^{\alpha t}Tf$. Substituting these expressions into (4), we obtain, for finite $v(x)$, the equation $\alpha v(x) = Tv(x)$. It follows easily from (1) that $\|v\| \leq \alpha^{-1}\|Vr\|$ for $\alpha > 0$ (see the proof of Theorem 1). The uniqueness of the bounded solutions is seen from the known representation of Bellmann's equation in the form

$$v = (\alpha + K)^{-1}(KI + T)v,$$

where on the right hand side the contraction operator is obtained for $\alpha > 0$. If $\alpha > 0$, $\|r\| < \infty$ and $v(\cdot) = \sup w(\cdot, \varphi)$ then φ is a Markovian policy. The result A is well-known (cf. [5], [8], [9]).

Under the conditions of B, one has, on account of Theorem 4.1 of [13] and by virtue of the homogeneity of the model, $v'_s(x, e^{-\alpha t}v) = e^{-\alpha s}v(x)$ and according to Lemma 5.1 from [13]

$$\begin{aligned} \int_0^u e^{-\alpha t} \mathbf{E}_x^\varphi(T - T^{\varphi_t})v(x_t) dt &= v^u(x, e^{-\alpha u}v) - w^u(x, \varphi, e^{-\alpha u}v) \\ &= v(x) - w^u(x, \varphi) - \mathbf{E}_x^\varphi e^{-\alpha u}v(x_u). \end{aligned}$$

For $u \rightarrow \infty$, we obtain formula (4). According to Theorem 3.2 from [13], we have

$$w(x, \varphi) = w^u(x, \varphi) + \mathbf{E}_x^\varphi w_u(x_u, \varphi) = w^u(x, \varphi) + e^{-\alpha u} \mathbf{E}_x^\varphi w(x_u, \varphi'),$$

where φ' is the policy obtained from φ by shifting the time to u . Since $w^u(x, \varphi) \rightarrow w(x, \varphi)$ as $u \rightarrow \infty$ we have

$$\lim_{u \rightarrow \infty} e^{-\alpha u} \mathbf{E}_x^\varphi w(x_u, \varphi') = 0.$$

The limit in formula (3) is non-negative, since $v(x) \geq w(x, \varphi')$.

Points C and D are particular cases of Theorems 6.1 and 5.1 from [13].

Under the conditions of E, the model is summable from above and $\|v\| < \infty$; therefore, according to formula (3), the selector φ with $(T - T^\varphi)v \leq \alpha\varepsilon$ generates an ε -optimal stationary policy. For the optimality of the class of Markovian policies for $\|r\| < \infty$ the result E is known (cf. [6]).

In case F, the function $Uv(a)$ is upper semicontinuous on the compact set $A(x)$ and attains on it its upper bound so that a selector φ with $T^\varphi v = Tv$ exists. According to B, the stationary policy φ is optimal. For finite $A(x)$, $\alpha > 0$ and $\|r\| < \infty$, the existence of stationary optimal (among Markovian) policies was established by Miller [4] and Kakumanu [5];

Pliska [8] and Doshi [9] obtained various forms of sufficient conditions for general semicontinuous jump models.

The point G is proven in the same way as for discrete time: For $\alpha > 0$, a stationary optimal policy exists by virtue of F ; the case $\alpha = 0$ is obtained by going to the limit which is possible in view of the finiteness of the set of selectors (cf. [17], § 6.3).

Some generalizations of Theorem 2 for the case of jump models with Borel spaces of states were announced, without proof, in [14].

3. Let $0 \leq t < \infty$, $\alpha = 0$ and $\|Vr\| < \infty$. We say that the selector φ and the bounded function g and R on X form a *canonical triple* if for all x and t ,

$$(5) \quad w'(\cdot, \varphi, R) = tg + R = v'(\cdot, R).$$

The corresponding stationary policy is also called canonical. Similarly as in the case of discrete time (cf. [17], § 7.1), it follows from (5) that g is an asymptotic estimate of the model and that the canonical policy φ is asymptotically optimal in the sense that for any policy π ,

$$\overline{\lim}_{t \rightarrow \infty} \frac{w'(x, \pi)}{t_i} \leq g(x) = \lim_{t \rightarrow \infty} \frac{w'(x, \varphi)}{t}.$$

Theorem 3. *The selector φ and the bounded functions g and R form a canonical triple if and only if*

$$(6) \quad Q^\varphi g = VQg = 0,$$

$$(7) \quad T^\varphi R = TR = g.$$

PROOF. Since $v'(x, R) = v''_{u-t}(x, R)$ it follows from Theorem 1.A that $v'(x, R)$ is the unique, locally bounded (i.e., bounded on any set $[0, u] \times X$, $u > 0$) solution of the equation

$$(8) \quad \frac{\partial v'(x, R)}{\partial t} = Tv'(x, R), \quad t \geq 0,$$

with the initial condition $v^0(x, R) = R(x)$.

If (5) is satisfied, then (8) takes the form

$$g = T(tg + R) = V[r + tQg + QR].$$

For $t = 0$, we obtain $g = TR$. Dividing by t in the limit as $t \rightarrow \infty$, we have

$$0 = \lim_{t \rightarrow \infty} \sup_{A(x)} \left[Qg(a) + \frac{r(a) + QR(a)}{t} \right].$$

Under the given assumptions, the function $r + QR$ is bounded from above, and hence the right-hand side is $\sup_{A(x)} Qg(a) = VQg$. The right-hand sides of relations (6)–(7) are established. The function $w'(x, \varphi, R)$ is, in view of (5), also locally bounded and, according to Corollary 3.2 from [13], satisfies almost everywhere in t the equation

$$(9) \quad \frac{\partial w'(x, \varphi, R)}{\partial t} = T^\varphi w'(x, \varphi, R).$$

Substituting (5) into (9) we have $g = T^\varphi(tg + R) = r(\varphi) + tQ^\varphi g + Q^\varphi g$ (a.e.). The equalities $g = T^\varphi R$ and $Q^\varphi v = 0$ are obtained by letting $t \searrow 0$ and $t \rightarrow \infty$. The necessity of equations (6)–(7) is proven.

We prove their sufficiency. The function $tg + R$ is locally bounded, and in order to obtain the right-hand equation in (5), it remains to verify that it satisfies (8), i.e.,

$$(10) \quad g(x) = \sup_{A(x)} [r(a) + tQg(a) + QR(a)].$$

On account of (6)–(7),

$$(11) \quad \sup_{A(x)} [r(a) + QR(a)] = g(x), \quad \sup_{A(x)} tQg(a) = t \sup_{A(x)} Qg(a) = 0,$$

Moreover, both suprema are assumed for the same value $a = \varphi(x)$. Therefore, the supremum of the sum equals the sum of the suprema, and (10) follows from (11). In order to obtain the left equation in (5), it is sufficient to apply what has just been said to the model Z_1 , obtained from Z by reducing the set of controls $A(x)$ to one-point sets consisting of points $\varphi(x)$ (in the model Z_1 , the role of $v'(x, R)$, VQ , T and Vr is played by $w'(x, \varphi, R)$, Q^φ , T^φ and $r(q)$, and the boundedness of the function $r(\varphi)$ necessary for the application of Theorem 1.A follows from the boundedness of g and R and from the formula $T^\varphi R = g$). The theorem is proven.

If g is a constant, then (6) is automatically satisfied and only equations (7) remain. In this form, equations (7) were used as sufficient conditions of asymptotic optimality in the class of Markovian policies in the sense of the criterion $\lim_{t \rightarrow \infty} t^{-1} w'(x, \pi)$ by Kakumanu [7] (in the case of denumerable X and A), by Doshi [9] (in case of semicontinuous jump models with general X and A), and by Bather [10] (for denumerable X , unbounded q and r but only in the class of stationary policies).

Theorem 4. *If the spaces X and A are finite, then there exist canonical policies in the model Z .*

The proof is analogous to the case of discrete time (cf. [17] chapter 7). Miller has shown that, in the model considered, the procedure of improving the policy according to Howard, leads finally to a selector φ such that

$$(12) \quad Q^\varphi g(x) = 0, \quad x \in X, \quad Qg(a) \leq 0 \quad a \in A,$$

$$r(\varphi(x)) + Q^\varphi R_1(x) = g(x), \quad x \in X,$$

$$(13) \quad r(a) + QR_1(a) \leq g(x) \quad \text{as } a \in A(x), \quad Qg(a) = 0,$$

where the functions g and R_1 are taken from the decomposition $w_\alpha(x, \varphi) = \alpha^{-1}g(x) + R_1(x) + o(1)$ of the estimate of the policy φ with discount coefficient $\alpha \searrow 0$ ([4], Theorem 8.10 and Lemma 11). As Denardo and Fox [15] have pointed out, it follows from (12) and (13) that

$$(14) \quad r(\varphi(x)) + Q^\varphi R(x) = g(x), \quad r(a) + QR(a) \leq g(x), \quad a \in A(x), \quad x \in X,$$

for the function $R = R_1 + cg$, provided that $c > [Qg(a)]^{-1}[g(j(a)) - r(a) - QR_1(a)]$ for all a for which $Qg(a) < 0$. Relations (12) and (14) are equivalent to the system (5)–(6) so that φ is a canonical policy.

4. It follows from Corollary 6.2 and Theorem 4.1 in [13] that the results of §§ 2–3 remain valid for a more generally defined model in which A is an arbitrary measurable space, provided one introduces the supplementary condition $\|r\| < \infty$ in the assertions 1, 2B and 2E, and the condition $\|r\| + \|v\| < \infty$ in the formulation of 2A. These supplementary conditions are, in fact, unnecessary but in order to get rid of them one must modify the derivation of the equation of optimality in [13] using in it the homogeneity of the model in an essential way.

Moscow Institute for
Transport Engineering

Received by the editors
June 17, 1977

REFERENCES

- [1] R. A. HOWARD, *Dynamic Programming in Markov processes*, Wiley, New York, 1960.
- [2] V. V. RIKOV, *Control of Markov processes with finite state and control space*, Theory Prob. Applications, 11 (1966), pp. 302–313.
- [3] B. L. MILLER, *Finite state continuous time Markov decision processes with a finite planning horizon*, SIAM J. Control, 6 (1968), pp. 266–280.
- [4] B. L. MILLER, *Finite state continuous time Markov decision processes with an infinite planning horizon*, J. Math. Anal. Appl., 22 (1968), pp. 522–569.
- [5] P. KAKUMANU, *Continuously discounted Markov decision model, with countable state and action space*, Ann. Math. Statist., 42 (1971), pp. 919–926.

- [6] P. KAKUMANU, *Nondiscounted continuous time Markovian decision process with countable state space*, SIAM J. Control, 10 (1972), pp. 210–220.
- [7] P. KAKUMANU, *Continuous time Markovian decision processes. Average return criterion*, J. Math. Anal. Appl., 52 (1975), pp. 173–188.
- [8] S. R. PLISKA, *Controlled jump processes*, Stochastic Processes Appl., 3 (1975), pp. 259–282.
- [9] B. T. DOSHI, *Continuous time control of Markov processes on an arbitrary state space: Average return criterion*, Stochastic Processes Appl., 4 (1976), pp. 55–77.
- [10] J. BATHER, *Optimal stationary policies for denumerable Markov chains in continuous time*, Advances Appl. Probability, 8 (1976), pp. 144–158.
- [11] S. A. LIPMAN, *Countable state continuous time dynamic programming with structure*, Operations Res., 24 (1976), pp. 477–490.
- [12] GIKHMAN AND SKOROKHOD, *Controlled Random Processes*, Kiev Naukova Dumka, 1977.
- [13] A. A. YUSHKEVIC, *Controlled Markovian models with a denumerable set of states and continuous time*, Theory Prob. Applications, 22 (1977), pp. 215–235.
- [14] A. A. YUSHKEVIC, *Controlled Markovian jump processes*, DAN SSSR, 277 (1977), pp. 304–307.
- [15] E. V. DENARDO AND B. L. FOX, *Multichain Markov renewal programme*, SIAM J. Appl. Math., 16 (1968), pp. 469–487.
- [16] A. A. YUSHKEVIC, *On a class of policies in general controlled Markov models*, Theory Prob. Applications, 18 (1973), pp. 775–779.
- [17] E. B. DYNKIN AND A. A. YUSHKEVIC, *Controllable Markov Processes and their applications*, Nauka, Moscow, 1975.

SOME EXAMPLES OF EQUIVALENT AND ORTHOGONAL GAUSSIAN DISTRIBUTIONS

S. M. KRASNITSKII

(Translated by A. B. Aries)

Let E^N be N -dimensional Euclidean space, let $T \subset E^N$, let $(\Omega, \mathfrak{S})$ be a measurable space, and let $\mathbf{P}_1, \mathbf{P}_2$ be probability measures on $(\Omega, \mathfrak{S})$. Assume that the function $\eta: E^N \times \Omega \rightarrow E^1$ is a homogeneous random Gaussian field in E^N with respect to $\mathbf{P}_1$, as well as $\mathbf{P}_2$.

Let $\xi: T \times \Omega \rightarrow E^1$ ($\xi = \xi(\cdot, \cdot)$) be the restriction of η on T . Let us assume also that the correlation functions of the field ξ , determined with respect to $\mathbf{P}_1$ and $\mathbf{P}_2$ (we denote them by R_1 and R_2 , respectively), are distinct. The expectations of the field ξ are taken to be zero. Furthermore, we assume the existence of spectral densities $f_j: E^N \rightarrow [0, \infty)$, $j = 1, 2$, that is,

$$R_j(t, s) = \int_{E^N} e^{-i(\lambda, t-s)} f_j(\lambda) d\lambda, \quad j = 1, 2,$$

$$\lambda = (\lambda_1, \dots, \lambda_N) \in E^N, \quad t = (t_1, \dots, t_N) \in T, \quad s = (s_1, \dots, s_N) \in T,$$

$$(\lambda, \tau) = \lambda_1 \tau_1 + \dots + \lambda_N \tau_N.$$

for $\tau = (\tau_1, \dots, \tau_N)$.

Let $\mathfrak{B} \subset \mathfrak{S}$ be the σ -algebra generated by the family of random variables $\{\xi(t), t \in T\}$, and let ν_1, ν_2 be the restrictions of $\mathbf{P}_1, \mathbf{P}_2$ on $\mathfrak{B}$. As is well known (see [1]–[6]), there exists a definite connection between the behavior of the pair (f_1, f_2) and the equivalence (or orthogonality) of ν_1 and ν_2 . The case $N > 1$ differs essentially from the univariate case in several respects. Some examples illustrating this difference are given in [4], [5]. The examples given below demonstrate other aspects of the difference.