J. Austral. Math. Soc. Ser. B 25 (1984), 463–472

INFINITELY DIVISIBLE RANDOM TRANSITION PROBABILITIES WITH APPLICATION TO DEPENDENT MARKOV CHAINS

PETER L. CHESSON[1]

(Received 7 April 1982; revised 30 November 1982)

Abstract

Random transition probability matrices with stationary independent factors define "white noise" environment processes for Markov chains. Two examples are considered in detail. Such environment processes can be used to construct several Markov chains which are dependent, have the same transition probabilities and are jointly a Markov chain. Transition rates for such processes are evaluated. These results have application to the study of animal movements.

1. Introduction

Markov processes provide useful models for a wide variety of natural phenomena. However, in the biological setting the systems being modelled are often subject to parameter variation through time. Examples in animal movements are given by Chesson [2, 3] and in population ecology by May [9]. When parameter variation is modelled as a stochastic process, one obtains a Markov process in a random environment (*e.g.*, Keiding [7]). It is clear that several processes which are independent in a constant environment will generally be dependent in a common random environment, and this fact must be a frequent cause of dependence between processes in nature. This paper provides a simple means of modelling such dependent processes.

Here we consider continuous time finite Markov chains in a random environment. The environmental variation is described by a family of random transition

[1] Department of Zoology, The Ohio State University, 1735 Neil Avenue, Columbus, Ohio 43210, U.S.A.

probabilities which have the property of infinite divisibility. As a result, the processes which are conditionally Markov processes with these transition probabilities, are Markov processes unconditionally. Moreover, if several processes are conditionally independent and identically distributed with these transition probabilities, then, unconditionally these processes will be dependent. Marginally, each process is a Markov process, and the joint vector process is also a Markov process. Thus, an environment process of this kind leads to very simple results. Such Markov processes have application to the study of animal movements.

2. Infinitely divisible transition probabilities

A transition matrix is a $k \times k$ non-negative matrix with rows summing to 1.

DEFINITION 1. *A random transition matrix P will be called infinitely divisible if, for every n, it can be expressed as a product $P_1 \cdots P_n$ where $P_1, \ldots, P_n$ are i.i.d. random transition matrices.*

By a family of random transition matrices we mean a collection $\{P(s,t), 0 \leqslant s \leqslant t < \infty\}$ such that $P(s,t)$ is a random transition matrix, the Chapman-Kolmogorov equations

$$P(r,s)P(s,t) = P(r,t) \tag{1}$$

are satisfied for each realization of the process, and $\{P(s,t)\}$ satisfies the regularity condition

$$\lim_{t \downarrow s} P(s,t) = I \tag{2}$$

for every s.

DEFINITION 2. *A family of random transition matrices $\{P(s,t)\}$ is said to have stationary independent factors if, for every n and $0 \leqslant t_1 \leqslant t_2 \leqslant \cdots \leqslant t_n$, $P(t_1, t_2)$, $P(t_2, t_3), \ldots, P(t_{n-1}, t_n)$ are mutually independent and the distribution of $P(s,t)$ depends only on $t - s$.*

Clearly, if $\{P(s,t)\}$ has stationary independent factors, then $P(s,t)$ is infinitely divisible for every s and t $(s \leqslant t)$.

For every family of random transition matrices $\{P(s,t)\}$ there is a stochastic process $X = \{X(t), t \geqslant 0\}$ such that, given $\{P(s,t)\}$, X is a nonhomogeneous Markov chain with transition matrix $\{P(s,t)\}$. Transition matrices of a nonhomogeneous Markov chain are commonly constructed as solutions to Kolmogorov's forward equation (Goodman [5]):

$$\partial_t P(s,t) = P(s,t)B(t) \tag{3}$$

https://doi.org/10.1017/S0334270000004215 Published online by Cambridge University Press

subject to $P(s, s) = I$, where elements $b_{ij}(t)$ of $B(t)$ are the transition rates which satisfy $b_{ii}(t) \leqslant 0$, $b_{ij}(t) \geqslant 0$, $i \neq j$, and $\Sigma_j b_{ij}(t) = 0$. The $b_{ij}(t)$ are assumed integrable in every finite interval. In the case $k = 2$ an explicit solution to (3) can be found and to express this we define

$$b_i(t) = -b_{ii}(t), \qquad \beta(t) = \exp\left\{\int_0^t (b_1(u) + b_2(u))\, du\right\}$$

and

$$p_i(s, t) = 1 - [\beta(t)]^{-1} \int_s^t \beta(u) b_i(u)\, du. \tag{4}$$

We then have

$$P(s, t) = \begin{bmatrix} p_1(s, t) & 1 - p_1(s, t) \\ 1 - p_2(s, t) & p_2(s, t) \end{bmatrix}.$$

A family of random transition matrices can also be constructed using the differential equation (3) starting with a matrix valued stochastic process $\{B(t)\}$ satisfying a.s. the above conditions on transition rates. However no family of random transition matrices with stationary independent factors can be found as a solution of (3). To see this we note from Goodman [5] that a solution of (3) will be continuous in s and t and $0 < \det P(s, t) \leqslant 1$. Using the stationary independent factor property, and equation (1), it is easy to see that $-\log \det P(s, t) = Z(t) - Z(s)$ where $\{Z(t)\}$ is a process with nonnegative stationary independent increments. The process $\{Z(t)\}$ necessarily has jump discontinuities (Breiman [1]) which contradicts the continuity of $P(s, t)$. Thus other methods must be found to define a family with stationary independent factors. However in the examples below we shall see that an equation very much like (3) is satisfied in a certain sense.

EXAMPLE 1. An elementary example of a family with stationary independent factors is given by

$$P(s, t) = \exp\{Q[A(t) - A(s)]\} \tag{5}$$

where A is a right-continuous process with stationary independent non-negative increments and Q is a matrix with non-negative off diagonals and zero row sums, *i.e.*, Q is a matrix of transition rates for a homogeneous Markov chain. A process that is conditionally a Markov chain with transition probabilities (5) is unconditionally a homogeneous Markov chain and is called a subordinated process (Feller [4]).

https://doi.org/10.1017/S0334270000004215 Published online by Cambridge University Press

EXAMPLE 2. A more interesting example of a family with stationary independent factors can be given for $k = 2$ as follows:

Let $\mathbf{q}(t) = (q_1(t), q_2(t))'$, $t \geqslant 0$, be a right-continuous non-negative bivariate process with stationary independent increments. Define $q(t) = \mathbf{l}'\mathbf{q}(t)$, $\beta(t) = e^{q(t)}$ and

$$\beta^*(t) = \lim_{h \downarrow 0} \frac{\beta(t) - \beta(t-h)}{q(t) - q(t-h)}, \tag{6}$$

where the ratio is interpreted as $e^{q(t)}$ if $q(t) = q(t-h)$. For $i = 1, 2$ define

$$p_i(s, t) = 1 - [\beta(t)]^{-1} \int_{(s,t]} \beta^* dq_i. \tag{7}$$

The integral in (7) is Lebesgue-Stieltjes and we define

$$P(s, t) = \begin{bmatrix} p_1(s, t) & 1 - p_1(s, t) \\ 1 - p_2(s, t) & p_2(s, t) \end{bmatrix}. \tag{8}$$

LEMMA. *The family* $\{P(s, t)\}$ *defined by* (8) *is a family of random transition matrices with stationary independent factors.*

PROOF. Let $\mathcal{F}(s, t)$ be the σ-field generated by $\{\mathbf{q}(u) - \mathbf{q}(v),\ u, v \in (s, t]\}$, then $P(s, t)$ is $\mathcal{F}(s, t)$-measurable. To see this note that

$$[\beta(t)]^{-1} \int_{(s,t]} \beta^* dq_i = \lim_{n \to \infty} [\beta(t)]^{-1} \int_{(s,t]} \beta_n^* dq_i, \tag{9}$$

where

$$\beta_n^*(u) = \left(e^{q(u^+)} - e^{q(u^-)}\right) / \left(q(u^+) - q(u^-)\right),$$

with $u^+ = \min(t, [nu + 1]/n)$, $u^- = \max(s, [nu - 1]/n)$ and $[x]$ being the integer part of x. The R.H.S. of (9) is certainly $\mathcal{F}(s, t)$-measurable. The limit exists for each sample path of the $\mathbf{q}$ process and equals the L.H.S. showing that $P(s, t)$ is $\mathcal{F}(s, t)$-measurable. It follows that $P(t_1, t_2), \ldots, P(t_{n-1}, t_n)$ will be mutually independent whenever $0 \leqslant t_1 \leqslant t_2 \leqslant \cdots \leqslant t_n$.

Now

$$\int_{(s,t]} \beta^* dq_1 + \int_{(s,t]} \beta^* dq_2 = \int_{(s,t]} \beta^* dq = \beta(t) - \beta(s), \tag{10}$$

which can be seen from the fact that β^* and q have discontinuities at the same points and the contribution to the integral from the discontinuity at a point u is $\beta(u) - \lim_{h \downarrow 0} \beta(u - h)$. Using (10) we see that $0 \leqslant p_i(s, t) \leqslant 1$ and therefore $P(s, t)$ is a random transition matrix. Equation (10) also implies the Chapman-Kolmogorov equations (1), and the final requirement (2) is clear from the integral representation.

https://doi.org/10.1017/S0334270000004215 Published online by Cambridge University Press

The striking resemblance between (4) and (7) suggests that (8) may be a solution to the stochastic differential equation

$$dP(s, t) = P(s, t)dQ(t) \tag{11}$$

subject to $P(s, s) = I$, with

$$Q(t) = \begin{bmatrix} -q_1(t) & q_1(t) \\ q_2(t) & -q_2(t) \end{bmatrix}.$$

Indeed Example 2 can be interpreted as a solution of (11). To see this we note that

$$|\beta^*(u) - 1| \leqslant e^{q(h)} - 1 \quad \text{for } 0 < u \leqslant h$$

and therefore

$$\beta(h)^{-1}\left|\int_{(0,h]} \beta^* dq_i - \int_{(0,h]} dq_i\right| \leqslant (1 - e^{-q(h)})q(h) = o(q(h)).$$

It follows that

$$P(0, h) - I = Q(h) + o(q(h));$$

and this result quite trivially extends to give

$$P(t, t + h) - I = Q(t + h) - Q(t) + o(q(t + h) - q(t))$$

for $t \geqslant 0$. Combining this with (1) we get

$$P(s, t + h) - P(s, t) = P(s, t)[Q(t + h) - Q(t)] + o(q(t + h) - q(t)) \tag{12}$$

which is a possible interpretation of (11). It is not difficult to see that Example 1 also satisfies (12) with $Q(t) = QA(t)$. In both examples we see that $dQ(t)$ has the interpretation as a matrix of transition rates. With this interpretation it seems likely that Example 2 exhausts the useful examples that might arise in practice for $k = 2$. Example 1, however, is highly specific.

3. The mean of a family with stationary independent factors

An infinitely divisible random transition matrix P has the property that $EP = (P_n)^n$ for some transition matrix P_n, and every n. If EP is also non-singular, it follows that (Kingman [8])

$$EP = \exp(Q^*)$$

where Q^* has non-negative off diagonal elements and row sums 0. Applying the bounded convergence theorem to (2), we have $EP(s, t) \to I$ as $t \downarrow s$, and hence

https://doi.org/10.1017/S0334270000004215 Published online by Cambridge University Press

that $EP(s, t)$ is non-singular for t sufficiently close to s. Thus, if $\{P(s, t)\}$ has stationary independent factors, we have

$$EP(s, t) = \exp(Q^*(t - s)) \tag{13}$$

with

$$Q^* = \lim_{t \downarrow 0} \frac{EP(0, t) - I}{t}.$$

In Example 1, Q^* can be calculated using the fact that the Laplace transform of the process A can be written

$$\log Ee^{-\lambda A(t)} = t \int_{[0,\infty)} \frac{e^{-\lambda x}}{x} d\mu(x) \tag{14}$$

where the measure μ satisfies $\int_{[0,\infty)}(1 + x)^{-1} d\mu(x) < \infty$ (Feller [4]). The matrix Q^* can now be given as

$$Q^* = \int_{[0,\infty)} \frac{e^{Qx} - I}{x} d\mu(x). \tag{15}$$

This result is proved by Nelson [10] as can be seen by noting that e^{Qt}, $t \geqslant 0$, is a contraction semigroup in R^k.

For Example 2 we evaluate Q^* in the case when $\mathbf{q}$ is a generalized Poisson process, *i.e.*,

$$\mathbf{q}(t) = \sum_{j=1}^{N(t)} \mathbf{U}_j, \tag{16}$$

where $N(t)$ is a standard Poisson point process on $[0, \infty)$ with intensity μ, $\mathbf{U}_1, \mathbf{U}_2, \ldots$ is a non-negative i.i.d. bivariate process independent of the process N. Writing

$$Q^* = \begin{bmatrix} -q_1^* & q_1^* \\ q_2^* & -q_2^* \end{bmatrix} \tag{17}$$

and letting $(U_1, U_2)'$ have the same distribution as $\mathbf{U}_t$, we obtain the formula

$$q_i^* = \mu E(1 - e^{-(U_1 + U_2)}) \frac{U_i}{(U_1 + U_2)}. \tag{18}$$

(Here the integrand is interpreted as zero for $U_1 + U_2 = 0$.) To see this we let τ_n be the nth point t in $(0, \infty)$ for which $N(t)$ has a jump and let J_n be the jump in $\beta(t)$ at τ_n. We see that

$$1 - p_i(0, t) = E \sum_{\tau_n \leqslant t} J_n [\beta(t)\mathbf{U}_n' \mathbf{1}]^{-1} U_{in} \tag{19}$$

$$= EJ_1 U_{i1} [\beta(t)\mathbf{U}_1' \mathbf{1}]^{-1} 1_{\{\tau_1 \leqslant t < \tau_2\}} + g(t). \tag{20}$$

https://doi.org/10.1017/S0334270000004215 Published online by Cambridge University Press

Now $|g(t)| \leqslant EN(t)1_{\{\tau_2 \leqslant t\}} = \mu t(1 - e^{-\mu t}) = O(t^2)$. Thus, $\lim_{t\to 0} t^{-1}(1 - p_i(0, t)) = \lim_{t\to 0} t^{-1} EJ_1 U_{i1}[\beta(t)\mathbf{U}_1'\mathbf{1}]^{-1} 1_{\{\tau_1 \leqslant t < \tau_2\}}$ and since $\beta(t) = e^{\mathbf{U}_1'\mathbf{1}}$ on $\{\tau_1 \leqslant t < \tau_2\}$ and $\mathbf{U}_1$ is independent of τ_1 and τ_2, this leads to (18). A useful special case arises when q_1 and q_2 are independent univariate generalized Poisson processes. The bivariate process $\mathbf{q}$ is then generalized Poisson and (18) reduces to

$$q_i^* = \mu_i(1 - Ee^{-U_i}) \tag{21}$$

where μ_i is the intensity of the process q_i and U_i is the random variable giving the jump size in q_i.

4. Dependent Markov chains

Let $\{P(s, t)\}$ have stationary independent factors. Each realization of the process $\{P(s, t)\}$ constitutes a family of regular transition probabilities for a nonhomogeneous Markov chain X on $\{1, 2, \ldots, k\}$. It is not difficult to see that there are processes $X_1, \ldots, X_n$ such that, given $\{P(s, t)\}$, $X_1, \ldots, X_n$ are conditionally independent and each X_i is conditionally a nonhomogeneous Markov chain with transition matrix $P(s, t)$. The initial distributions of these processes can be chosen arbitrarily.

THEOREM. 1. *Unconditionally, each X_i is a homogeneous Markov chain with transition rates Q^*, $i = 1, \ldots, n$.*

2. *The vector process $\mathbf{X} = (X_1, \ldots, X_n)'$ is a homogeneous Markov chain.*

3. *If Q^* is irreducible, then $X_1, \ldots, X_n$ are dependent processes unless $P(s, t) = \exp(Q^*(t - s))$ a.s.*

PROOF. Let $P(s, t) = (p_{ij}(s, t))$, $EP(s, t) = (\bar{p}_{ij}(s, t))$.

Now

$$\begin{aligned} P(X_i(t) = j \mid X_i(u), u \leqslant s) &= E\left[P(X_i(t) = j \mid X_i(u), u \leqslant s, P(\sigma, \tau), 0 \leqslant \sigma \leqslant \tau < \infty) \mid X_i(u), u \leqslant s\right] \\ &= E\left[p_{X_i(s)j}(s, t) \mid X_i(u), u \leqslant s\right] \\ &= \bar{p}_{X_i(s)j}(s, t) \end{aligned}$$

because $P(s, t)$ and $\{X_i(u), u \leqslant s\}$ are independent. This proves 1.

Given $\{P(s, t)\}$ it is clear that $\mathbf{X}$ is a nonhomogeneous Markov chain with transition matrix

$$\Pi(s, t) = P(s, t) \otimes \cdots \otimes P(s, t) = P(s, t)^{[n]}. \tag{22}$$

https://doi.org/10.1017/S0334270000004215 Published online by Cambridge University Press

It follows that $\{\Pi(s, t)\}$ has stationary independent factors. Hence, part 2 of the theorem is a special case of part 1.

To prove 3 we note that

$$P\big(X_l(t) = X_{l'}(t) = j \mid X_l(s) = X_{l'}(s) = i\big) = Ep_{ij}^2(s, t) \tag{23}$$

and since this cannot equal $(\bar{p}_{ij}(s, t))^2$ for all ij, unless $P(s, t)$ is non-random, X_l and $X_{l'}$ are dependent. (The irreducibility condition ensures that X_l and $X_{l'}$ are not restricted to disjoint subsets of the state space.)

For the examples of $P(s, t)$ discussed previously, the transition rates for the Markov chain $\mathbf{X}$ are quite simply calculated.

Consider Example 1 and for simplicity assume that Q is diagonalizable ($Q = G\Lambda G^{-1}$, $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_k)$). Then

$$P(0, t)^{[n]} = G^{[n]}\big(e^{\Lambda A(t)}\big)^{[n]}\big(G^{-1}\big)^{[n]}$$

and it follows that the transition rates for the Markov chain are given by the matrix

$$G^{[n]}\Gamma_n\big(G^{-1}\big)^{[n]}$$

where Γ_n is the diagonal matrix with diagonal elements of the form

$$\int_{[0,\infty)} \frac{\exp\big((\lambda_{i_1} + \lambda_{i_2} + \cdots + \lambda_{i_n})x\big) - 1}{x} d\mu(x). \tag{24}$$

For the second example, the considerations in Section 2 show that

$$\begin{aligned}
&\lim_{t\to 0} t^{-1} E \prod_{j=1}^{n} \big[1 - p_{i_j}(0, t)\big] \\
&\qquad = \lim t^{-1} E J_1^n \big[\beta(t)\mathbf{U}'\mathbf{1}\big]^{-n} \prod_{j=1}^{n} U_{i_j} 1_{\{\tau_1 \le t < \tau_2\}} \\
&\qquad = \mu E\big\{(1 - \exp(-U_1 - U_2))/(U_1 + U_2)\big\}^n \prod_{j=1}^{n} U_{i_j},
\end{aligned} \tag{25}$$

i.e., the transition rate for $(X_j = i_j) \to (X_j \neq i_j)$, $j = 1, \ldots, n$, is given by expression (25). All other transition rates can be deduced easily from (25).

5. Applications

Consider n animals moving through k discrete areas and let $X_l(t)$ be the area occupied by individual l at time t. Given the environment, the animals' movements are ideally mutually independent (assuming there are no interactions

https://doi.org/10.1017/S0334270000004215 Published online by Cambridge University Press

between individuals) and are governed by the same probability laws because the animals respond in the same way to environmental stimuli. Random fluctuations in the environment cause random fluctuations in $P(s, t)$. Examples 1 and 2 have natural interpretations in this setting.

In Example 1, $dA(t)$ is essentially a rate of movement of animals among all areas; it represents uniform changes in the rate of movement for all areas and may be interpreted as coming from variation in the environment of the animals that affects all areas equally. In Example 2, $dq_i(t)$ represents variation in the rate of leaving area i. Since $q_1(t)$ and $q_2(t)$ can be arbitrarily related, the environmental variation represented by $dq_1(t)$ and $dq_2(t)$ can affect different areas differently. For example it may promote congregation of animals in one area, during one period of time, and in the other area during another period of time. Nothing like this is possible in Example 1.

In a more general setting, the processes $X_1, \ldots, X_n$ simply describe the movements of n particles among k compartments which are subjected to environmental randomness (Saunders [11, 12]). Heuristically, the environment described by $P(s, t)$ can be said to be "independent at every instant," and it is therefore a "white noise" environment process. This is not to be confused with Gaussian white noise (*e.g.* Hida [6]) that generates the Wiener process model of Brownian motion. However both white noise processes have the property of "independence at every instant."

Acknowledgements

With pleasure I acknowledge support from a CSIRO postdoctoral studentship and NSF Grant #DEB77-14470. An anonymous reviewer gave helpful comments which have significantly improved the paper.

References

[1] L. Breiman, *Probability* (Addison-Wesley, Reading, Mass., 1968).
[2] P. L. Chesson, "Models for animal movements", Ph.D. Thesis, University of Adelaide, 1976.
[3] P. L. Chesson, "A theory of animal movements" (1982) (in preparation).
[4] W. Feller, *An introduction to probability theory and its applications*, Vol. II (Wiley, New York, 1971).
[5] G. S. Goodman, "An intrinsic time for non-stationary finite Markov chains", *Z. Wahrsch. Verw. Gebiete* 16 (1970), 165–180.
[6] T. Hida, *Stationary stochastic processes* (Princeton Univ. Press, Princeton, N. J., 1970).
[7] N. Keiding, "Extinction and exponential growth in random environments", *Theoret. Population Biol.* 8 (1975), 49–63.

[8] J. F. C. Kingman, "The imbedding problem for finite Markov chains", *Z. Wahrsch. Verw. Gebiete* 1 (1962), 14–24.
[9] R. M. May, *Stability and complexity in model ecosystems* (Princeton Univ. Press, Princeton, N. J., 2nd edition, 1974).
[10] E. Nelson, "A functional calculus using singular Laplace integrals", *Trans. Amer. Math. Soc.* 88 (1958), 400–413.
[11] R. Saunders, "Conservative processes with stochastic rates", *J. Appl. Probab.* 12 (1975), 447–456.
[12] R. Saunders, "On joint exchangeability and conservative processes with stochastic rates", *J. Appl. Probab.* 13 (1976), 584–590.