ELSEVIER

Statistics & Probability Letters 38 (1998) 183–186

STATISTICS & PROBABILITY LETTERS

A representation result for finite Markov chains

Peter Spreij *

Department of Econometrics, Vrije Universiteit, De Boelelaan 1105, 1081 HV Amsterdam, Netherlands

Received 1 June 1997; received in revised form 1 October 1997

Abstract

In this short paper we derive a representation result in terms of a solution to a stochastic differential equation that is valid for both continuous and discrete time Markov processes that live on a finite state space. Martingale techniques are used throughout the paper. © 1998 Elsevier Science B.V. All rights reserved

AMS classification: 60G42; 60G44; 60J10; 60J27.

Keywords: Markov chain; Stochastic differential equation; Martingale

1. Introduction

Let $(\Omega, \mathscr{F}, \mathbb{F}, P)$ be a filtered probability space. Assume that the *filtration* $\mathbb{F}$ satisfies the usual conditions in the sense of Dellacherie and Meyer (1980). Let X be a cadlag process on $[0,\infty)$ with a finite state space. Without loss of generality we can assume that the state space is the *standard orthogonal basis* of the Euclidian space $\mathbb{R}^m$. Call this set $B^m = \{b_1, \dots, b_m\}$. (Indeed, if ξ is a stochastic process with values in a set $\{z_1, \dots, z_m\}$, where all the z_i are different, then we can define the process X with values in B^m by $X_t = b_i$ iff $\xi_t = z_i$. Hence the probabilistic structure of ξ determines that of X and vice versa). So, we view X as a mapping $X : \Omega \times [0,\infty) \to B^m$. If we work with a discrete time chain X then we use the familiar way of considering it as a continuous time cadlag chain by the convention $X_t \equiv X_{[t]}$. Similarly, one then usually has for the filtration $\mathscr{F}_t \equiv \mathscr{F}_{[t]}$.

By definition X is called $\mathbb{F}$-Markov if for all $t \geqslant s$ and for all $b \in B^m$ one has $P(X_t = b \mid \mathscr{F}_s) = P(X_t = b \mid \sigma(X_s))$. Denote by $\Phi(t,s)$ the $m \times m$ matrix with elements $\Phi_{ij}(t,s) = P(X_t = b_i \mid X_s = b_j)$. Notice that the *column sums* of $\Phi(t,s)$ are all equal to one.

The advantage of working with the state space B^m can already be illustrated with the following observation:

X is $\mathbb{F}$-Markov if and only if $E[X_t \mid \mathscr{F}_s] = \Phi(t,s)X_s$ for all $t \geqslant s$.

* E-mail: pspreij@econ.vu.nl.

0167-7152/98/$19.00
PII S0167-7152(98)00005-4

Necessity of this equality for the Markov property can be shown as follows. First we note that $P(X_t = b_i \mid \mathscr{F}_s) = \sum_{j=1}^m P(X_t = b_i \mid X_s = b_j)1_{\{X_s=b_j\}} = b_i^{\mathrm{T}}\Phi(t,s)X_s$. Hence $E[X_t \mid \mathscr{F}_s] = \sum_{i=1}^m b_i P(X_t = b_i \mid \mathscr{F}_s) = \sum_{i=1}^m b_i b_i^{\mathrm{T}} \Phi(t,s)X_s = \Phi(t,s)X_s$, since $\sum_{i=1}^m b_i b_i^{\mathrm{T}} = I$.

In order to show sufficiency we use that $X_t = b_i$ *iff* $b_i^{\mathrm{T}}X_t = 1$. So we obtain $P(X_t = b_i \mid \mathscr{F}_s) = E[1_{\{X_t=b_i\}} \mid \mathscr{F}_s] = b_i^{\mathrm{T}}E[X_t \mid \mathscr{F}_s] = b_i^{\mathrm{T}}\Phi(t,s)X_s$.

The representation of a discrete time Markov chain X as the solution of a linear (stochastic) difference equation (Proposition 1.1 below) has been popular in the stochastic systems theory literature for a long time, since this equation looks the same as the one for a Gaussian state process. Especially in references devoted to filtering problems one will often encounter such representations, cf. the book by Elliott et al. (1995) and references therein. Specifically, there one uses the following result, see e.g. Elliott et al. (1995, p. 17) for the time homogeneous case.

Proposition 1.1. *Let $\{X_t\}_{t\in\{0,1,2,\dots\}}$ be a discrete time Markov chain with $A(t)$ the matrix of one step transitions probabilities at time $t: A_{ij}(t) = P(X_{t+1} = e_i \mid X_t = e_j)$ and define for each $t \geqslant 1$*

$$\varepsilon_t = X_t - A(t-1)X_{t-1}. \tag{1.1}$$

The process $\{\varepsilon_t\}$ is then a martingale difference sequence adapted to the filtration generated by X.

In a previous paper (Spreij, 1990) we showed a similar result (but now of *iff*-type) in continuous time. Let $\Phi(t,s)$ be as before and let (the limit is assumed to exist for all $t \geqslant 0$) $A(t) = \lim_{h\downarrow 0}(1/h)[\Phi(t+h,t) - I]$. Then we have the following convenient equivalent representation of an $\mathbb{F}$-Markov process with values in B^m as the solution of a stochastic differential equation.

Proposition 1.2. *A stochastic process $X : \Omega \times [0,\infty) \to B^m$ is $\mathbb{F}$-Markov with generating matrix A iff X satisfies the stochastic differential equation*

$$\mathrm{d}X_t = A(t)X_t\,\mathrm{d}t + \mathrm{d}M_t,\ X_0 \tag{1.2}$$

with $M : \Omega \times [0,\infty) \to \mathbb{R}^n$ a n-dimensional $\mathbb{F}$-martingale.

2. The general representation

The next proposition generalizes Propositions 1.1 and 1.2. As before we consider an $\mathbb{F}$-adapted cadlag stochastic process $X : \Omega \times [0,\infty) \to B^m$ and we denote by $\Phi(t,s)$ the matrix with entries $\Phi_{ij}(t,s) = P(X_t = b_i \mid X_s = b_j)$.

Before stating the main result (Proposition 2.2 below) we need a technical lemma. It involves the Doléans exponent $\mathscr{E}$ (see Jacod (1979, Chapter 6) for a definition). The following holds true.

Lemma 2.1. *If X is $\mathbb{F}$-Markov, then there is a function of bounded variation Q with values in $\mathbb{R}^{m\times m}$ such that $\Phi(t,s) = \mathscr{E}(Q(s+\cdot))_{t-s}$ for all $t \geqslant s$.*

Proof. This follows from Gill and Johansen (1991). Using their Theorem 15, we have $Q(t) = \int_{(0,t]} \mathrm{d}(\Phi - I)$, and $\Phi(t,s) = \prod_{(s,t]}(I + \mathrm{d}Q)$, where the $\prod$ here stands for the product-integral defined as a limit of matrix products, in which the ordering of the product is the *opposite* of the one in Gill and Johansen (1991). As a consequence their formula (40) now takes the form

$$\Phi(t,s) = I + \int_{(s,t]} \mathrm{d}Q(u)\Phi(u-,s) \tag{2.1}$$

from which the assertion follows. □

Remark. The form of the function Q follows from results on product integration. However, in two extreme cases it is easy to define Q without the theory of product integration. Consider first the case in which $\Phi(\cdot,0)$ is differentiable. Then $\Phi(t,0)$ is invertible and $Q(t)$ is simply $\int_0^t \dot{\Phi}(u,0)\Phi(u,0)^{-1}\mathrm{d}u$.

In the other case we assume that X is a Markov chain in discrete time on the integers with $\Phi(t+1,t) = A(t)$. Then $Q(t) = \sum_{k=1}^{t-1}(A(k) - I)$.

Proposition 2.2. *If X is $\mathbb{F}$-Markov, then there is a bounded variation function Q such that M defined by*

$$M_t = X_t - X_0 - \int_{(0,t]} \mathrm{d}Q(s)X_{s-} \tag{2.2}$$

is an $\mathbb{F}$-martingale. Conversely, if there is a martingale M and a bounded variation function Q such that X is the solution of Eq. (2.2), then X is $\mathbb{F}$-Markov with transition probabilities as in Lemma 2.1.

Proof. Using the fact that $E[X_t \mid \mathscr{F}_s] = \Phi(t,s)X_s$, the definition of M and Lemma 2.1 we compute the conditional expectation

$$\begin{aligned} E[M_t \mid \mathscr{F}_s] &= E[X_t \mid \mathscr{F}_s] - X_0 - \int_{(0,s]} \mathrm{d}Q(u)X_{u-} - E\left[\int_{(s,t]} \mathrm{d}Q(u)X_{u-} \mid \mathscr{F}_s\right] \\ &= \Phi(t,s)X_s + M_s - X_s - \int_{(s,t]} \mathrm{d}Q(u)\Phi(u-,s)X_s \\ &= M_s + \left[\Phi(t,s) - I - \int_{(s,t]} \mathrm{d}Q(u)\Phi(u-,s)\right] X_s \end{aligned}$$

and the result follows from Eq. (2.1). For the proof of the converse statement we compute

$$\begin{aligned} E[X_t \mid \mathscr{F}_s] &= X_0 + \int_{(0,s]} \mathrm{d}Q(u)X_{u-} + E\left[\int_{(s,t]} \mathrm{d}Q(u)X_{u-} \mid \mathscr{F}_s\right] + M_s \\ &= X_s + \int_{(s,t]} \mathrm{d}Q(u)E[X_{u-} \mid \mathscr{F}_s]. \end{aligned}$$

So, $E[X_t \mid \mathscr{F}_s]$ satisfies a Volterra in equation t which has a unique solution given by $E[X_t \mid \mathscr{F}_s] = \Phi(t,s)X_s$ in view of the fact that $\Phi(t,s)$ satisfies Eq. (2.1). □

Remark. If we consider again the two cases mentioned in the remark following Lemma 2.1, we see that we obtain Propositions 1.1 and 1.2 as special cases of Proposition 2.2.

In much the same way as in the second part of the proof of Proposition 2.2 we obtain that an $\mathbb{F}$-Markov chain X also satisfies the Strong Markov property. In the book of Breiman (1968, Proposition 15.25) one can find a classical proof.

Proposition 2.3. *Let X be $\mathbb{F}$-Markov and T an a.s. finite $\mathbb{F}$-stopping time. Then*

$$E[X_{T+t} \mid \mathscr{F}_T] = \Phi(T+t,T)X_T \quad \textit{for all } t \geqslant 0. \tag{2.3}$$

Proof. First we assume that T is bounded. Since we know that X satisfies Eq. (2.2) and that $E[M_{T+t} \mid \mathscr{F}_T] = M_T$ by the optional stopping theorem, we find by application of Fubini's theorem (Q has bounded variation over finite intervals) for conditional expectations and the fact that T is $\mathscr{F}_T$-measurable that

$$\begin{aligned} E[X_{T+t} \mid \mathscr{F}_T] &= X_T + E\left[\int_{(T,T+t]} \mathrm{d}Q(u)X_{u-} \mid \mathscr{F}_T\right] \\ &= X_T + E\left[\int \mathrm{d}Q(u)1_{\{T<u\leqslant T+t\}}X_{u-} \mid \mathscr{F}_T\right] \\ &= X_T + \int \mathrm{d}Q(u)E[1_{\{T<u\leqslant T+t\}}X_{u-} \mid \mathscr{F}_T] \\ &= X_T + \int_{(T,T+t]} \mathrm{d}Q(u)E[X_{u-} \mid \mathscr{F}_T]. \end{aligned}$$

So we obtain again a Volterra equation, the solution of which yields the desired expression under the condition that T is bounded.

Next we consider the general case. Let n be a natural number, then $T \wedge n$ is a bounded stopping time. In the string of equalities below we use Eq. (2.3) for this bounded stopping time. Because $\{T\leqslant n\}$ is $\mathscr{F}_{T\wedge n}$-measurable, we can write

$$\begin{aligned} E[X_{T+t} \mid \mathscr{F}_{T\wedge n}] &= E[1_{\{T\leqslant n\}}X_{T+t} \mid \mathscr{F}_{T\wedge n}] + 1_{\{T>n\}}E[X_{T+t} \mid \mathscr{F}_{T\wedge n}] \\ &= E[1_{\{T\leqslant n\}}X_{T\wedge n+t} \mid \mathscr{F}_{T\wedge n}] + 1_{\{T>n\}}E[X_{T+t} \mid \mathscr{F}_{T\wedge n}] \\ &= 1_{\{T\leqslant n\}}E[X_{T\wedge n+t} \mid \mathscr{F}_{T\wedge n}] + 1_{\{T>n\}}E[X_{T+t} \mid \mathscr{F}_{T\wedge n}] \\ &= 1_{\{T\leqslant n\}}\Phi(T\wedge n+t, T\wedge n)X_{T\wedge n} + 1_{\{T>n\}}E[X_{T+t} \mid \mathscr{F}_{T\wedge n}] \\ &= 1_{\{T\leqslant n\}}\Phi(T+t, T)X_T + 1_{\{T>n\}}E[X_{T+t} \mid \mathscr{F}_{T\wedge n}]. \end{aligned}$$

Now take limits for $n \to \infty$ in the extreme members of the above string. Clearly the last term converges to $\Phi(T+t,T)X_T$ and the first one, being a martingale in n, to $E[X_{T+t} \mid \mathscr{F}_T]$ according to the martingale convergence theorem. □

Remark. Notice that we can now also deduce that $E[M_{T+t} \mid \mathscr{F}_T] = M_T$, with M as in Eq. (2.2) for finite stopping times T. This does not immediately follow from the optional sampling theorem, where one requires M to be uniformly integrable or T bounded. As a second remark we mention that Eq. (2.3) is also valid on the set $T<\infty$ for an arbitrary stopping time, as can easily be verified.

References

Breiman, L., 1968. Probability. Addison-Wesley, Reading, MA.

Elliott, R.J., Aggoun, L., Moore, J.B., 1995. Hidden Markov Models. Estimation and Control. Springer, Berlin.

Dellacherie, C., Meyer, P.A., 1980. Probabilités et Potentiel. Hermann.

Gill, R.D., Johansen, S., 1991. A survey of product-integration with a view toward application in survival analysis. Ann. Statist. 18, 1501–1555.

Jacod, J., 1979. Calcul Stochastique et Problèmes de Martingales, Lecture Notes in Mathematics, vol. 714. Springer, Berlin.

Spreij, P.J.C., 1990. Self-exciting counting process systems with finite state space. Stoch. Proc. Appl. 34, 275–295.