Søren Johansen

Product Integrals and Markov Processes

Preprint
April
1977
3

Institute of Mathematical Statistics
University of Copenhagen

Søren Johansen

PRODUCT INTEGRALS AND MARKOV PROCESSES

Preprint 1977 No. 3

INSTITUTE OF MATHEMATICAL STATISTICS
UNIVERSITY OF COPENHAGEN

April 1977

Product Integrals and Markov Processes

by

Søren Johansen

1. Introduction. A finite state Markov Process is usually described by the Kolmogorov differential equations

$$(1.1) \qquad \frac{\partial}{\partial t} P(s,t) = P(s,t)Q(t),$$

where $P(s,t)$ denotes the transition probabilities and $Q(t)$ the intensity matrix. The solution to (1.1) with initial condition $P(s,s) = I$ is given by the product integral

$$P(s,t) = \prod_s^t (I + Q(u)du)$$

For references to this, see Schlesinger (1931) and Dobrushin (1953). The product integral was used to study the imbedding problem by Johansen (1973).

In the study of the estimation problem for a nonhomogeneous Markov Process (Aalen and Johansen, 1977) it turned out that we needed the product-integral representation of a transition probability which is only piecewise constant. This note contains the definition of the product integral, where the measure $\int_A Q(u)du$ is replaced by an arbitrary matrix valued measure ν on [0,1].

Section 2 contains the product integral and some of its properties and section 3 then applies this integral to the representation of the transition probabilities in terms of its integrated intensities. We thus obtain a different approach to some of the results of Jacobsen (1972).

2. The Product Integral

Let $\mathfrak{B}$ denote the Borel sets of [0,1] and let ν be a σ-additive finite signed measure with values in the set of nxn matrices, i.e. a matrix of n^2 real measure. We let $\nu^{(n)}$ denote product measure on $[0,1]^n$ defined by

$$\nu^{(n)}(B_1 \times \ldots \times B_n) = \nu(B_1) \cdot \ldots \cdot \nu(B_n), B_i \in \mathfrak{B}.$$

We shall use the notation $\| \cdot \|$ to denote matrix norm $\|\nu(B)\| = \sup_i \sum_j |\nu_{ij}(B)|$, and introduce the real positive measure

$$\nu_0 = \sum_i \sum_j |\nu_{ij}| (B)$$

where $|\nu_{ij}| = \nu_{ij}^{+} + \nu_{ij}^{-}$.

One easily checks, that $\|\nu(B)\| \leqslant \nu_0(B)$ and that $\|\nu^{(n)}(B_1 \times \ldots \times B_n)\| \leqslant \prod_{i=1}^{n} \nu_0(B_i)$.

We now give the basic definition:

2.1. Definition. For $B \in \mathfrak{B}$ we define the product integral

$$\prod_B (I+d\nu) = I + \sum_{n=1}^{\infty} \nu^{(n)} (B \times \ldots \times B \cap \{u_1 < \ldots < u_n\}).$$

Notice that from the inequality

$$\|\nu^{(n)} (B \times \ldots \times B \cap \{u_1 < \ldots < u_n\})\| \leqslant$$

$$\nu_0^{(n)} (B \times \ldots \times B \cap \{u_1 < \ldots < u_n\}) \leqslant \frac{1}{n!} \nu_0 (B)$$

follows the convergence of the series.

The following evaluations follow easily from the definition.

2.2. Proposition. For $B \in \mathfrak{B}$ we have

$$\| I + \nu(B) \| \leq e^{\nu_o(B)}$$

$$\| \prod_B (I+d\nu) \| \leq e^{\nu_o(B)}$$

$$\| \prod_B (I+d\nu) - I \| \leq \nu_o(B)\, e^{\nu_o(B)}$$

$$\| \prod_B (I+d\nu) - I - \nu(B) \| \leq \frac{1}{2}\, \nu_o(B)^2 e^{\nu_o(B)}$$

Now we can immediately prove the basic property of multiplicativity:

2.3. Theorem. For any $t \in [0,1]$ we have

$$\prod_B (I+d\nu) = \prod_{B\cap[o,t]} (I+d\nu) \prod_{B\cap]t,1]} (I+d\nu) .$$

Thus $\prod_B (I+d\nu)$ is multiplicative over disjoint intervals, which is the reason for its name.

Proof. For $i = 1,\ldots,n-1$ let

$$A(B,i,n) = B\times\ldots\times B \cap \{u_1 <\ldots< u_i \leq t < u_{i+1} <\ldots< u_n\}$$

with the obvious modification for $i = 0$ and n. We let

$$A(B,n) = B\times\ldots\times B \cap \{u_1<\ldots<u_n\} = \bigcup_{i=o}^{n} A(B,i,n) .$$

Now

$$\nu(A(B,n)) = \sum_{i=o}^{n} \nu(A(B,i,n))$$

$$= \sum_{i=o}^{n} \nu^{(i)}(A(B\cap[o,t],i))\, \nu^{(n-i)}(A(B\cap]t,1]),n-i) .$$

Summing over n gives the result.

2.4. Examples. If $d\nu = Qdt$, where Q is a fixed matrix then

$$\nu^{(n)}(0 \leq u_1<\ldots<u_n \leq t) = \frac{t^n Q^n}{n!}$$

and hence

$$\prod_{[o,t]} (I+d\nu) = e^{tQ} .$$

If $d\nu = Q_1 dt$ for $0 \leq t \leq t_1$ and $d\nu = Q_2 dt$ for $t_1 < t \leq 1$, then using Theorem 2.3, and the above example, we get

$$\prod_{[o,t]}(I+d\nu) = \begin{cases} e^{Q_1 t} & , 0 \leq t \leq t_1 , \\ e^{Q_1 t_1} e^{Q_2 (t-t_1)} & , t_1 < t \leq 1 . \end{cases}$$

As a final example we let $\nu = Q\varepsilon_{t_1}$ where Q is fixed and ε_{t_1} is a one point measure at t_1, then

$$\prod_{[o,t]}(I+d\nu) = \begin{cases} I & , 0 \leq t < t_1 , \\ I+Q & , t_1 \leq t \leq 1 . \end{cases}$$

Thus we get a piecewise constant function of t.

The following results give a different and perhaps more intuitive definition of the product integral. The definition we have chosen seems to give the basic results in a very efficient manner, since we can use existing integration and measure theory.

<u>2.5. Theorem</u>. Let $0 = t_o < t_1 < \ldots < t_n = 1$, then

$$\| \prod_{]o,t]}(I+d\nu) - \prod_{i=o}^{n-1}(I+\nu]t_i,t_{i+1}]) \| \leq c \sup_i \nu_o]t_i,t_{i+1}[.$$

<u>Proof</u>. We split the product integral into the corresponding factors and define

$$M_i = \prod_{]t_i,t_{i+1}]}(I+d\nu) , N_i = I + \nu]t_i,t_{i+1}]$$

then by Proposition 2.2, we get

$$\|M_i\| \leq e^{\nu_o]t_i,t_{i+1}]} , \quad \|N_i\| \leq e^{\nu_o]t_i,t_{i+1}]} .$$

We also get

$$\|M_i - N_i\| = \Big\| \prod_{]t_i,t_{i+1}[} (I+d\nu)(I+\nu[t_{i+1}]) - I - \nu]t_i,t_{i+1}[- \nu[t_{i+1}] \Big\|$$

$$\leqslant \Big\| \prod_{]t_i,t_{i+1}[} (I+d\nu) - I - \nu]t_i,t_{i+1}[\Big\| + \Big\| \prod_{]t_i,t_{i+1}[} (I+d\nu) - I \Big\| \, \|\nu[t_{i+1}]\|$$

$$\leqslant \tfrac{1}{2}(\nu_o]t_i,t_{i+1}[)^2 e^{\nu_o]t_i,t_{i+1}[} + \nu_o]t_i,t_{i+1}[\, e^{\nu_o]t_i,t_{i+1}[} \cdot \nu_o[t_{i+1}]$$

$$\leqslant e^{\nu_o]t_i,t_{i+1}[} (\sup_i \nu_o]t_i,t_{i+1}[) \, \nu_o]t_i,t_{i+1}] \ .$$

Using these evaluations we get

$$\Big\| \prod_{i=o}^{n-1} M_i - \prod_{i=o}^{n-1} N_i \Big\| \leqslant \sum_{i=o}^{n-1} \| M_o \cdots M_{i-1} (M_i - N_i) N_{i+1} \cdots N_{n-1} \|$$

$$\leqslant e^{\nu_o[o,1]} (\sum_i \nu_o]t_i,t_{i+1}]) \sup_i \nu_o]t_i,t_{i+1}[$$

as was the result we wanted to prove.

<u>2.6. Corrollary</u>. Let t_{in} satisfy the conditions

a) $o = t_{on} < t_{1n} < \ldots < t_{nn} = 1$

and

b) $\lim\limits_{n\to\infty} \sup\limits_i \nu_o]t_{in}, t_{(i+1)n}[\; = o$,

then the product integral can be computed as

$$\prod_{]o,1]} (I+d\nu) = \lim_{n\to\infty} \prod_{i=o}^{n-1} (I+\nu]t_{in}, t_{(i+1)n}]) \ .$$

Notice that condition b) can always be satisfied, since we can make sure that the atoms of ν (i.e. of ν_o) eventually are among the division points.

The next results are about differentiability of the product integral.

<u>2.7. Theorem</u>. For $B \in \mathfrak{B}$ we have

$$\prod_B (I+d\nu) - I = \int_B \prod_{B\cap[o,u[} (I+d\nu)\,\nu(du)$$

<u>Proof</u>. Using Fubinis theorem on the $(n+1)$'st term we get

$$\nu^{(n+1)}(B\times\ldots\times B,\ 0 \leq u_1<\ldots<u_{n+1} \leq 1)$$

$$= \int_B \nu^{(n+1)}(B\times\ldots\times B,\ 0 \leq u_1<\ldots<u_{n+1} \leq 1 \mid u_{n+1}= u)\,\nu(du)$$

$$= \int_B \nu^{(n)}(B\times\ldots\times B,\ 0 \leq u_1<\ldots<u_n < u)\,\nu(du) \ .$$

Summing over n gives the result.

<u>2.8. Theorem</u>. For $B \in \mathfrak{B}$ we have

$$\prod_B (I+d\nu) - I = \int_B \nu(du) \prod_{]u,1]\cap B} (I+d\nu)$$

<u>Proof</u>. Similar to that of Theorem 2.7. If we define the function F by

$$t \rightarrow \prod_{[o,t]} (I+d\nu)$$

then F is right continuous by Theorem 2.7 and it is of bounded variation. It thus determines a matrix valued measure, which by Theorem 2.7 is absolutely continuous with respect to ν_o. The integral relation can thus be reformulated as

$$\frac{dF}{d\nu_o}(t) = F(t^-)\frac{d\nu}{d\nu_o} \ , \text{ a.s. } [\nu_o] \ .$$

If $H : t \rightarrow \prod_{]t,1]} (I+d\nu)$, then Theorem 2.8 can be reformulated as

$$\frac{dH}{d\nu_o}(t) = -\frac{d\nu}{d\nu_o}(t)H(t) \ , \text{ a.s. } [\nu_o] \ .$$

3. Markov Processes

In constructing a Markov Process one can start with the transition probabilites, satisfying the Chapman-Kolmogorov equations, then construct the process, i.e. the measure on a suitable function space, by the Kolmogorov consistency theorem, see Doob (1953), or via a general theorem of extension of continuous linear functionals, see Nelson (1959), or Goodman and Johansen (1973). In this case the discussion of the differential equations for the transition probabilities becomes a discussion of when a process is determined by its infinitesimal properties.

One can also start out with the waiting time distributions and the jump intensities and then construct the measure very directly, and then prove that certain variables form a Markov Process and define the transition probabilities in terms of these. The differential equations can now be viewed as a convinient and different way of obtaining the transition probabilities, see Jacobsen (1972).

We shall here start with the intensities or rather the integrated intensities ν, i.e. we assume that

$$\nu_{ii} \leq 0 \ , \ \nu_{ij} \geq 0 \ , \ i \neq j \quad \text{and} \quad \sum_j \nu_{ij} = 0 \ .$$

From this measure we construct the transition probabilities by a product integral and this also gives us the differential equations for P. Thus the approach is highly non-probabilistic as opposed to that of Jacobsen (1972). The solution, however, is the same, as we shall show.

Thus, we let ν be finite signed measure on $[o,1]$ with values in the set of intensity matrices, then the following holds:

<u>3.1. Theorem</u>. If $\nu[t] + I$ is a stochastic matrix, i.e. if $\nu_{ii}[t] \geqslant -1$, then

$$P(B) = \prod_{B}(I+d\nu)$$

is a stochastic matrix.

<u>Proof</u>. Assume first that $\nu_{ii}[t] > -1$, $i = 1,\dots,n$, i.e. no atoms of ν are as large as -1. Then let us choose the partition t_{in} of $[o,1]$ so fine that $I + \nu]t_{in},t_{(i+1)n}]$ is a stochastic matrix. By Corrollary 2.6 $P(B)$ is the limit of stochastic matrices, hence stochastic.

In general there can only be a finite number of points $t_1,\dots t_k$, such that some $\nu_{ii}[t_r] = -1$. By writing

$$P(B) = (I+\nu(o)) \cdot \prod_{i=1}^{k-1} \prod_{B\cap]t_i,t_{i+1}[} (I+d\nu)(I+\nu[t_{i+1}])$$

we see that $P(B)$ is stochastic.

For a given ν we now define

$$P(s,t) = P(]s,t]) = \prod_{]s,t]} (I+d\nu) \ , \quad 0 < s < t \leqslant 1 \ ,$$

$$P(o,t) = P([o,t]) = \prod_{[o,t]} (I+d\nu) \ , \quad 0 \leqslant t \leqslant 1 \ ,$$

then it is seen that $P(s,t)$ is right continuous in s and t (except for $s = o$, $t \downarrow o$), and that

$$P(t^-,t) = I + \nu[t] \ , \quad 0 < t \leqslant 1$$

$$P(t,t^+) = I \ , \quad 0 < t < 1$$

$$P(o,o^+) = I + \nu[o] \ .$$

The multiplicativity of the product integral now immediately gives that $P(s,t)$ satisfies the Chapman-Kolmogorov equations

$$P(s,t) = P(s,u)P(u,t) \qquad 0 \leqslant s < u < t \leqslant 1,$$

and in this formulation, Theorem 2.7 gives the forward differential equation

$$\frac{\partial P(s,t)}{\partial \nu_o(t)} = P(s,t)\frac{d\nu}{d\nu_o}(t) \qquad \text{a.s. } [\nu_o]$$

whereas Theorem 2.8 gives the backward equation

$$\frac{\partial P(s,t)}{\partial \nu_o(s)} = -\frac{d\nu}{d\nu_o}(s)\; P(s,t) \quad , \text{a.s.}[\nu_o]$$

which show that $P(s,t)$ do infact have ν as integrated intensity measure.

Using a result similar to Theorem 2.5 one can prove that for $s < t_{on} < \ldots < t_{nn} = t$ such that $\sup_i \nu_o]t_{in}, t_{(i+1),n}] \to 0$ we have

$$\prod_{]s,t]} (1+d\nu_{kk}) = \lim_{n\to\infty} \prod_{i=0}^{n-1} (1+P_{kk}]t_{in}, t_{(i+1),n}])$$

which is nothing but the waitingtime distribution in state i, i.e.

$$P\{X_u = i,\ s < u \leqslant t \mid X_s = i\} = \prod_{]s,t]} (1+d\nu_{ii}).$$

With the notation

$$G_i[o,t] = 1 - \prod_{[o,t]} (1+d\nu_{ii})$$

and

$$\pi_{ij}(t) = -\frac{d\nu_{ij}}{d\nu_{ii}}(t),\ i \neq j,\ \pi_{ij}(t) = 0$$

we can now prove that the solution provided by Jacobsen, starting with G and π is in fact the same as the solution provided here starting from ν.

Notice that ν can be recovered from G and π, by the relations

$$\nu_{ii}(A) = - \int_A \frac{G_i(du)}{1-G_i[o,u[}$$

and

$$\nu_{ij}(A) = - \int_A \pi_{ij}(u)\,\nu_{ii}(du).$$

We shall now assume, that ν satisfies the following extra conditions

1) $\nu[0] = 0$

2) $\nu_{ii}[t] > -1.$

i.e. $P(s,t)$ becomes rigth continuous, also at 0, and no atom is as large as -1.

It is then seen that G_i is continuous at zero, and that $G_i[o,t] = 0$ if $\nu_{ii}[o,t] = 0$ and that $G_i[o,t] < 1$, since ν_{ii} $[0,1]$ is finite. In order to see the last result, where condition 2) is needed, we argue as follows: The largest atom of $|\nu_{ii}|$ is $1-\varepsilon$ say.Now choose $0 = t_o < \ldots < t_n = 1$ such that $|\nu_{ii}]t_j,t_{j+1}[| < \frac{\varepsilon}{2}$ then

$1+\nu_{ii}]t_j,t_{j+1}] > \frac{\varepsilon}{2}$

and $\log(1+\nu_{ii}]t_j,t_{j+1}]) \geq \dfrac{\log\frac{\varepsilon}{2}}{-1+\frac{\varepsilon}{2}}\,\nu_{ii}]t_j,t_{j+1}]$.

Summing over j gives

$$\prod_j (1+\nu_{ii}]t_j,t_{j+1}]) \geq c > 0$$

which again implies that $\prod_B (1+d\nu_{ii}) \geq c > 0$.

Thus the functions G and π satisfy the conditions of Jacobsen and his solution $\tilde{P}(s,t)$ is constructed to satisfy the integral equation:

obs!

$$\tilde{p}_{ij}(s,t) = \delta_{ij} \frac{G_i]t,1]}{G_i]s,1]} + \sum_{k\neq i} \int_{]s,t]} \pi_{ik}(u)\tilde{p}_{kj}(u,t) \frac{G_i(du)}{G_i]s,1]} \ .$$

Using the definition of (G,π) in terms of ν this is

$$(3.1) \quad \tilde{p}_{ij}(s,t) = \delta_{ij} \prod_{]s,t]} (1+d\nu_{ii}) - \sum_{k\neq i} \int_{]s,t]} \tilde{p}_{kj}(u,t) \prod_{]s,u[} (1+d\nu_{ii})\nu_{ik}(du)$$

which is known to have a unique solution, see Feller (1940).

The function $P]s,t] = \prod_{]s,t]} (I+d\nu)$ is known to satisfy the equation

$$(3.2) \quad p_{ij}(s,t) - \delta_{ij} = \sum_k \int_{]s,t]} \nu_{ik}(du)p_{kj}(u,t) \ .$$

In this equation we multiply by $\prod_{]a,s[} (1+d\nu_{ii})$ and integrate with respect to $\nu_{ii}(ds)$. If we then use the results:

$$\int_{]a,t]} \prod_{]a,s[} (1+d\nu_{ii})\nu_{ii}(ds) = \prod_{]a,t]} (1+d\nu_{ii}) - 1$$

and

$$\int_{]a,u[} \prod_{]a,s[} (1+d\nu_{ii})\nu_{ii}(ds) = \prod_{]a,u[} (1+d\nu_{ii}) - 1$$

then we get after some reduction that P also satisfies equation (3.1), hence $P = \tilde{P}$.

In fact the equations (3.1) and (3.2) are equivalent. If (3.1) is integrated with respect to $\nu_{ii}(ds)$ on $]a,t]$, we arrive at (3.2).

It should of course be pointed out that we are only dealing with a finite number of states, whereas Jacobsen treats the more general situation of a countable number of states.

4. References

[1] Aalen, O.O. and Johansen, S.: Estimation of the transition probabilities of a finite state Markov Process. In preparation (1977).

[2] Dobrušin, R.L.: Generalization of Kolmogorov's equations for Markov Processes with a finite number of possible states. Mat. Sb. N.S. 33, 567-596 (1953).

[3] Doob, J.L.: Stochastic Processes. New York: Wiley 1953.

[4] Feller, W.: On the Integro-Differential Equations of Purely Discontinuous Markoff Processes. Trans. Amer. math. Soc. 48, 488-515 (1940). Errata. Ibid. 58, 474 (1945).

[5] Goodman, G.S. and Johansen, S.: Kolmogorov's differential equations for non-stationary countable state Markov processes with uniformly continuous transition probabilities. Proc. Camb. Phil. Soc. 73, 119-138 (1973).

[6] Jacobsen, M.: A Characterization of Minimal Markov Jump Processes. Z. Wahrscheinlichkeitstheorie verw. Geb. 23, 32-46 (1972).

[7] Johansen, S.: A central Limit Theorem for Finite Semigroups and its Application to the Imbedding Problem for Finite State Markovs Chains. Z. Wahrscheinlichkeitstheorie verw. Geb. 26, 171-190 (1973).

[8] Nelson, E.: Regular probability measures on function space. Ann. of Math. 69, 630-643 (1959).

[9] Schlesinger, L.: Neue Grundlagen für einen Infinitesimalkalkül der Matrizen. Math. Z. 33, 33-61 (1931).

* * *
*

University of Copenhagen

PREPRINTS 1976

COPIES OF PREPRINTS ARE OBTAINABLE FROM THE AUTHOR OR FROM THE INSTITUTE OF MATHEMATICAL STATISTICS, UNIVERSITETSPARKEN 5, 2100 COPENHAGEN Ø, DENMARK.

No. 1 Becker, Niels: Estimation for an Epidemic Model.

No. 2 Kaplan, Norman: A Generalization of a Result of Erdös and Renyi and a Related Problem.

No. 3 Lauritzen, Steffen & Remmer, Ole: An Investigation of the Asymptotic Properties of Geodetic Least Squares Estimates in Certain Types of Network.

No. 4 Asmussen, Søren: Almost Sure Behavior of Linear Functionals of Supercritical Branching Processes.

No. 5 Hald, Anders & Møller, Uffe: On the SPRT of the Mean of a Poisson Process.

No. 6 Lauritzen, Steffen: Projective Statistical Fields.

No. 7 Keiding, Niels: Inference and Tests for Fit in the Birth-Death-Immigration Process.

No. 8 Møller, Uffe: OC and ASN og the SPRT for the Poisson Process.

No. 9 Keiding, Niels: Population Growth and Branching Processes in Random Environments.

No. 10 Andersen, Per K., Andersen, Søren & Lauritzen, Steffen: The Average Noise from a Poisson Stream of Vehicles.

No. 11 Hald, Anders & Møller, Uffe: Multiple Sampling Plans of Given Strength for the Poisson and Binomial Distributions.

No. 12 Johansen, Søren: Two Notes on Conditioning and Ancillarity.

No. 13 Asmussen, Søren: Some Martingale Methods in the Limit Theory of Supercritical Branching Processes.

No. 14 Johansen, Søren: Homomorphisms and general exponential families.

No. 15 Asmussen, Søren & Hering, Heinrich: Some Modified Branching Diffusion Models.

No. 16 Hering, Heinrich: Minimal Moment Conditions in the Limet theory for General Markov Branching Processes.

PREPRINTS 1977

COPIES OF PREPRINTS ARE OBTAINABLE FROM THE AUTHOR OR FROM THE INSTITUTE OF MATHEMATICAL STATISTICS, UNIVERSITETSPARKEN 5, 2100 COPENHAGEN Ø, DENMARK.

No. 1 Asmussen, Søren & Keiding, Niels: Martingale Central Limit Theorems and Asymptotic Estimation Theory for Multitype Branching Processes.

No. 2 Jacobsen, Martin: Stochastic Processes with Stationary Increments in Time and Space.

No. 3 Johansen, Søren: Product Integrals and Markov Processes.