Søren Johansen Fred L. Ramsey

A Bang-Bang Representation for 3x3 Embeddable Stochastic Matrix

Preprint
April
1978
7

Institute of Mathematical Statistics
University of Copenhagen

Søren Johansen and Fred L. Ramsey*

A BANG-BANG REPRESENTATIONS
FOR 3x3 EMBEDDABLE STOCHASTIC MATRIX

Preprint 1978 No. 7

INSTITUTE OF MATHEMATICAL STATISTICS
UNIVERSITY OF COPENHAGEN

April 1978

*Dept. of Statistics, Oregon State University, Corvallis, Oregon.

H. C. Ø. - tryk
københavn

Abstract: It is proved that a 3x3 embeddable stochastic matrix has a representation as a product of a finite number of elementary stochastic matrices, with only one off-diagonal element positive. In particular if the determinant is $\geqq \frac{1}{2}$ then only 6 matrices are needed and a necessary and sufficient condition for embeddability in this case is given.

Keywords: Markov chains, embedding problem, Control theory, Kolmogorov's differential equations, Bang-Bang representations.

1. Introduction, summary, and notation.

Consider a finite state Markov chain with transition probabilities $P(s,t)$ satisfying the Chapman-Kolmogorov equations

(1.1) $P(s,t) = P(s,u)\,P(u,t),\ 0 \leqq s \leqq u \leqq t \leqq 1,$

with initial condition

(1.2) $P(s,t) = I \Leftrightarrow s = t$

and regularity condition

(1.3) $P(s,t)$ is continuous in $0 \leqq s \leqq t \leqq 1.$

A stochastic matrix P is called embeddable if there exists a family $\{P(s,t)\}$ satisfying (1.1), (1.2), and (1.3) for which

(1.4) $P(0,1) = P.$

It was proved by Goodman (1970), that by changing the time scale by means of $\phi(t) = -\ln \operatorname{Det} P(0,t)$ one makes the transition probabilities almost surely differentiable and $P(s,t)$ can be formed as the unique solution to the forward or backward Kolmogorov equations

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

(1.6) $\frac{\partial}{\partial s} P(s,t) = -Q(s)P(s,t)$

(1.7) $P(s,s) = I.$

Here the equations are satisfied almost surely with respect to Lebesgue measure and $Q(t)$ is an integrable function on $[0,1]$ with values in the space of intensity matrices

(1.8) $\mathcal{Q} = \{Q = \{q_{ij}\} \mid q_{ij} \geqq 0,\ i \neq j,\ q_{ii} \leqq 0,\ \Sigma_j q_{ij} = 0\}$

Thus the embedding problem can be viewed as a control problem using the intensity matrix $Q(t)$ as a control variable. The embeddable matrices are thus the matrices that can be reached from the identity I using a suitable choice of control variable.

Let us consider some simple cases. If $Q(t) \equiv Q$ then $P(s,t) = \exp((t-s)Q)$ which corresponds to a time homogeneous Markov chain. If $Q(t) \equiv h(t)Q$, where $h(t) \geqq 0$ is integrable on [0,1] then $P(s,t) = \exp(\int_s^t h(u)du)Q$. In particular if Q is an extremal element of $\mathcal{Q}$, i.e. Q has only one off diagonal element positive and equal to 1 say, then the matrix $\exp(aQ)$, $0 \leqq a \leqq \infty$ will be called an <u>elementary</u> (or Poisson) matrix and such matrices will be denoted by K. Notice that since $Q^2 = -Q$ we have $\exp(aQ) = e^{-a}I + (1-e^{-a})(I+Q)$, which means that $\exp(aQ)$ is a convex combination of the two stochastic matrices I and I+Q. In particular $\exp(\int_a^t h(u)du\, Q)$ is contained in the interval from I to I + Q, but does not attain the value I + Q, since this would correspond to $\int_0^1 h(u)du = \infty$.

It is well known, see Lee and Marcus (1968) that any embeddable matrix can be approximated with a finite product of elementary matrices. This is called the chattering principle. It was proved by Johansen (1973) that any matrix in the interior of the set of embeddable matrices has a representation as a finite product of elementary matrices, see also Hazod (1976) for a different approach. Such a representation is called a Bang-Bang representation in control theory because one imagines the target being reached, using a finite number of switches between the extremal controls. A general account of Bang-Bang representations for bilinear systems can be found in the paper by Krener (1974). A related result about the configuration of zeroes in an embeddable matrix can be found in the paper by Kingman and Williams (1973).

With this background we can now state the main result of this paper: For 3 x 3 stochastic matrices we can prove that embeddable matrices have a Bang-Bang representation, and the number of switches is bounded by 6 times the smallest integer larger than or equal to

$$(\ln\tfrac{1}{2})^{-1} \quad \ln \operatorname{Det} P.$$

In particular if Det P $\geqq \frac{1}{2}$ then we need only use 6 factors, where 6 is the dimension of the space of embeddable matrices. The results are formulated in Theorem 9 and Theorem 10.

We also obtain an explicit representation of the set of matrices which are embeddable and have determinant $\geqq \frac{1}{2}$. The set R is defined in (2.1), (2.2), and (2.3), and one can by checking at most 9 simple conditions decide, whether a stochastic matrix with determinant $\geqq \frac{1}{2}$ is embeddable.

For a matrix with positive principal minors and determinant $\geqq \frac{1}{2}$, the conditions are that for some permutation (i,j,k) of (1,2,3) we have either

$$p_{ii}p^{ii} \geqq 1$$

or

$$p_{jj}p_{kk}p^{kj} + p_{kj} \leqq 0.$$

Having explained now the results one can give a very brief outline of the method of proof, leaving out all technicalities.

We first describe the set R, see Fig. 1 and 2 and prove that apart from a regularity condition any matrix in R is a product of at most 6 elementary matrices (Lemma 7). We then prove that, again apart from regularity conditions, if P $\in$ R then also KP $\in$ R. This contractability condition means that if we are ever in R then the extremal controls can not take us out of R. This is then used to prove that since I $\in$ R we get that if Det $K_n \ldots K_1 \geqq \frac{1}{2}$ then also

$K_n \ldots K_1 \in R$ and since R is closed and such products approximate embeddable matrices we find that (Theorem 10) any embeddable matrix with determinant $\geqq \frac{1}{2}$ is in R and hence a product of at most 6 elementary matrices.

We shall conclude this section by giving a different formulation of the problem which will be used in the proofs. This formulation is also due to Gerald Goodman (1969), (1974). Let a stochastic matrix P be given. We denote the rows by a, b, and c respectively and represent each row as a point in S the simplex spanned by the unit vectors of R^3. Thus a matrix is represented by a labeled triangle which we shall denote (a,b,c), we shall even write P = (a,b,c). We denote the unit vectors A,B, and C and the identity matrix I is then represented by (A,B,C). Let now P_1 be a stochastic matrix and let $P_0 = P_1 P = (a^0, b^0, c^0)$. It is easily seen that the rows of P_0 are convex combinations of the rows of P, with coefficients determined by P_1, hence $(a^0, b^0, c^0) \subset (a,b,c)$.

Now let P be embeddable by the family P(s,t) satisfying (1.1). If we represent P(t,1) by (a,b,c)(t) then it follows from the equation $P(s,1) = P(s,t)\ P(t,1)$, $0 \leqq s \leqq t \leqq 1$, and the above remarks, that (a,b,c) (s) $\subset$ (a,b,c) (t). Hence the embedding problem can be thought of as embedding a labeled triangle (a,b,c) in a monotone continuous family of labeled triangles (a,b,c) (s) which connects (a,b,c) with (A,B,C).

Let us then consider the effect of left multiplication by an elementary matrix K. It is seen that if P = (a,b,c) then KP is a labeled triangle which has two vertices in common with (a,b,c), a and b say, and where the third vertex is a convex combination of a and c (or b and c), one can say that K moves c towards a (or b).

If we consider a finite product of elementary matrices $K_m K_{m-1} \cdots K_1 = (a_m, b_m, c_m)$ and where K_{m+1} moves the third vector towards the first we shall say that K_{m+1} moves c from c_m towards a_m into c_{m+1}, some times leaving out a_m if it is obvious from the context.

Thus a,b,and c without index will in general denote variable vertices and with an index they are used to denote starting or terminal points.

With this geometric formulation the Bang-Bang problem is that of reaching an embeddable triangle using a finite number of the above described moves.

In order to describe the moves of the vertices, the following notation has proved useful. Let x and y be different points of S. We let [x,y] denote the closed interval between x and y, and [x,y[denote the interval without y. We let xy denote that point in S on the halfline from x through y which is as far away from x as possible. One can think of xy as the projection of x through y onto the boundery of S. When needed we shall denote the coordinates of a point $x \in S$ by x_1,x_2, and x_3.

2. Main results.

For $P = (a,b,c)$ we define the functions

(2.1) $$f(P) = (a_3b_2 - a_2b_3)\ (c_1 + c_2) - c_2(a_3 - b_3)$$

(2.2) $$g(P) = (c_2 - b_2(c_1 + c_2))\ (c_3a_2 - a_3c_2) - (a_2 - c_2)\ b_3c_2.$$

Notice that f is linear in a,b, and c and g is linear in a and b, and quadratic in c.

Let σ denote a permutation matrix, then we define

$$f_\sigma(P) = f(\sigma P\sigma^{-1}),\ g_\sigma(P) = g(\sigma P\sigma^{-1}),$$

and the set of matrices

(2.3) $$R = \mathrm{cl} \bigcup_\sigma \{P \mid (f_\sigma \vee g_\sigma) \wedge \mathrm{Det}(P) > 0\},$$

where cl denotes closure.

The functions f and g are defined for a particular labeling of the states, but the set R is made independent of this labeling, since obviously $P \in R \Rightarrow \sigma P\sigma^{-1} \in R$.

A matrix will be called <u>regular</u> if all principal minors are positive and we can at once prove the result:

<u>Lemma 1</u> If P is imbeddable and Det P $\geqq \frac{1}{2}$ then P is regular.

Proof: It was proved by Goodman (1970) that each of the diagonal elements of an embeddable matrix dominate the determinant, and that this determinant is positive: $p_{ii} \geqq \text{Det } P > 0$, where equality holds only if $P = I$. The matrix I is clearly regular and if $P \neq I$, then $p_{ii} > \frac{1}{2}$ and $p_{ii}p_{jj} - p_{ij}p_{ji} > \frac{1}{2}\frac{1}{2} - \frac{1}{2}\frac{1}{2} = 0$.

We shall now describe the set R and prove various properties of R. The main idea is to fix two rows, a and b say, and consider the section $R_{ab} = \{c \mid (a,b,c) \in R\}$. This set is clearly defined for all a and b but we shall only deal with it in the following cases.

<u>Case 1</u> The rows a and b satisfy Det $(a,b,A) \leqq 0$, Det $(a,b,B) \leqq 0$, Det $(a,b,C) > 0$.

<u>Case 2</u> The rows a and b satisfy Det $(a,b,A) < 0$, Det $(a,b,B) > 0$, Det$(a,b,C) > 0$.

In order to clarify the meaning of these concepts, notice that Det (a,b,c) is proportional to the area of the triangle (a,b,c) with a positive sign if the orientation is the same as that of (A,B,C) otherwise negative. Thus in case 1 the directed line through a and b has C on its positive side and A and B on the negative, and in case 2 both C and B on the positive side and only A on the negative. One can also characterize case 1 by the inequalities $ab \in [C,B]$, $ba \in [A,C]$ and case 2 by $ab \in]A,B[$ and $ba \in]A,C]$, see Figure 1 and 2.

<u>Lemma 2</u> If $P = (a,b,c)$ is regular then either (a,b) is in case 1 or 2 or $\sigma_0 P \sigma_0^{-1}$ has this property, where σ_0 is the permutation matrix that interchanges state 1 and 2.

Proof: Regularity of P implies that $\text{Det}\ (a,b,C) > 0$. With this condition it follows that $\text{Det}\ (a,b,A) \geqq 0 \Rightarrow \text{Det}\ (a,b,B) \leqq 0$ and $\text{Det}\ (a,b,B) \geqq 0 \Rightarrow \text{Det}\ (a,b,A) \leqq 0$.

It is then easy to see that if (a,b) is not in case 1 or 2, the only remaining possibility is that $\text{Det}\ (a,b,A) > 0$ and $\text{Det}\ (a,b,B) < 0$ and in this case clearly the permitation σ_0 will have the desired effect.

<u>Lemma 3</u> If (a,b) is in case 1 then $R_{ab} = \{c \mid \text{Det}\ (a,b,c) \geqq 0\} = \text{co}\ \{ab, ba, C\}$ where co denotes the convex hull, see Figur 1.

<u>Figure 1</u>

Proof: Consider the function f. We shall prove that $c \in S \Rightarrow f(a,b,c) \geqq 0$. Indeed $f(a,b,C) = 0$, $f(a,b,A) = -\det(a,b,A) \geqq 0$ and $f(a,b,B) = -\det(a,b,B) \geqq 0$. Now $f(a,b,c)$ is linear in c and it must therefore be nonnegative on S.

Hence $\max\ (f_\sigma \vee g_\sigma)\ (a,b,c) \geqq 0$ and the set R_{ab} is thus only given as the closure of the set where $\text{Det}\ (a,b,c) > 0$ which is $\{c \mid \text{Det}\ (a,b,c) \geqq 0\}$.

In order to describe the set R_{ab} in case 2 we introduce the following notation: σ_1 is the permutation matrix corresponding to the permutation (312) and σ_2 corresponds to (231).

Lemma 4 If (a,b) is in case 2 then

$$R_{ab} = \{c \mid (g \vee g_{\sigma_1} \vee g_{\sigma_2}) \wedge \mathrm{Det}\ (a,b,c) \geqq 0\}$$

$$= \{c \mid g \wedge \mathrm{Det}\ (a,b,c) \geqq 0\} \cup \mathrm{co}\{ab,ba,C,Aa\} \cup \mathrm{co}\{ba,b,Bba,C\}.$$

Proof: The proof of this result, which is a little tedious will be defered to the appendix. We shall however, give a few remarks to explain how R_{ab} is constructed. The nontrivial part of the boundary is contained in the set [Aa,ab] ∪ [Bba,b] ∪ h_{ab}, where h_{ab} is the curve that connects Bba and ab and satisfies $g(a,b,h_{ab}) = 0$. The function g is quadratic in c and the discriminant is

$$D = 4\ b_2 a_2 (b_3 - a_3) - (a_2 - b_2 a_3)^2$$

which is $\leqq 0$ since $b_3 - a_3 = \mathrm{Det}\ (a,b,A) - \mathrm{Det}\ (a,b,B) < 0$. If D = 0 then $b_2 a_2 = a_2 - b_2 a_3 = 0$ but this contradicts $\mathrm{Det}\ (a,b,A) = a_2 b_3 - a_3 b_2 < 0$.

Thus the curve determined by $\{c \mid g(a,b,c) = 0\}$ is a hyperbola. It is not difficult to see that one branch goes through C,a, and A while the other goes through Bba and ab. The function g is positive between the branches.

The hyperbola section h_{ab} is constructed as follows: First take $b^* \in [ab,B]$ then define $h_{ab}(b^*) = [C,b^*] \cap [b^*b, b^*ba]$. It is seen that $g(a,b,h_{ab}(b^*)) = 0$.

The expression for g was of course found from this construction and $g \geqq 0$ only means that c_3 lies "above" the value h_{ab} (Cc).

As a help in the calculations involved here the coordinates of a few of the points will be given if (a,b) is in case 2.

$$ab = (a_3b_1 - a_1b_3, a_3b_2 - a_2b_3, 0)/(a_3 - b_3)$$

$$Cc = (c_1, c_2, 0)/(c_1 + c_2)$$

$$Ccb = (c_2b_3, 0, c_2b_1 - c_1b_2)/(c_2 - b_2(c_1 + c_2))$$

$$h_{ab}(Cc) = (c_1, c_2, \frac{c_2}{a_2} a_3 + (1 - \frac{c_2}{a_2}) \frac{c_2b_3}{c_2 - b_2(c_1 + c_2)})$$

$$Bba = (0, a_2b_1, a_3b_1 - a_1b_3)/(b_1 - a_1(b_1 + b_3))$$

Lemma 5 If (a,b) is in case 1 or 2 then R_{ab} is starshaped around a and b, i.e. $c \in R_{ab} \Rightarrow [a,c] \subset R_{ab}$ and $[b,c] \subset R_{ab}$.

Proof: In case 1 this is rather obvious since the triangle $\text{co}\{ab, ba, C\}$ is convex. In case 2 it is also quite clear if $g_{\sigma_1} \vee g_{\sigma_2}(a,b,c) \geqq 0$, since this function is linear in c and nonnegative at $c = a$ and $c = b$.

Let us therefore take c such that $g \wedge \text{Det}(a,b,c) \geqq 0$ and $g_{\sigma_1} \vee g_{\sigma_2}(a,b,c) < 0$. This implies, see Figure 2, that $g(a,b,bc) < 0$ and $g(a,b,ac) < 0$, since ac and bc lie on the boundary of S determined by $]Bba, B] \cup [B, ab[$ where $g(a,b,\cdot)$ is negative.

Now consider the function $g(a,b,\cdot)$ on the line through a and c. The function is quadratic, it has a zero at a, is nonnegative at c and negative at ac, hence it is positive on $[a,c]$ which means $[a,c] \subset R_{ab}$. Similarly the function on the line through b and c is nonnegative at b and c and negative at bc, which implies that it is nonnegative on $[b,c]$. Hence $[b,c] \subset R_{ab}$.

Corrollary 6 The set R is contractable in the sense that if $P \in R$ is regular and K is elementary then $KP \in R$.

Proof: Let $P = (a^0, b^0, c^0)$. There is no loss in generality in assuming that K moves vertex c, and that (a^0, b^0) is in case 1 or 2. That is $KP = < a^0, b^0, c^1 >$ where $c^1 \in [a^0, c^0]$ or $[b^0, c^0]$. Thus $c^0 \in R_{a^0b^0}$ but by Lemma 5 $R_{a^0b^0}$ is starshaped, and this means $c^1 \in R_{a^0b^0}$ and hence $(a^0, b^0, c^1) \in R$.

Lemma 7 Let $P \in R$ and assume that P is regular, then P can be embedded using at most 6 elementary matrices.

Proof: Since $P = (a^0, b^0, c^0)$ is regular we can without loss of generality assume that (a^0, b^0) is in case 1 or 2.

If (a^0, b^0) is in case 1, then (a^0, b^0, c^0) can be reached as follows: First we take a from A towards C into b^0a^0 then b from B towards C into a^0b^0. Then we take c from C to $b^0a^0c^0$ and then to c^0. Finally we take a from b^0a^0 to a^0 and b from a^0b^0 to b^0.

Let (a^0, b^0) be in case 2. We shall distinguish different cases. Assume first that c^0 is on the hyperbola section $h_{a^0b^0}$. We then embed (a^0, b^0, c^0) as follows: First take b from B to Cc^0 then a from A to Bb^0, then c from C to c^0, and b from Cc^0 to b^0 and a from Bb^0 to a^0. Thus we used only 5 moves. Notice how the points were moved cyclicly $b \to a \to c \to b \to a \to c$, each time as far as possible, compatible with the condition $(a^0, b^0, c^0) \subset (a, b, c)$.

It is easily checked that if $c^0 \in [Aa^0, a^0b^0]$ or $[Bb^0a^0, b^0]$ then (a^0, b^0, c^0) can be embedded in 5 moves which are cyclic as above but starting with $c \to b$ or $a \to c$ respectively.

If $c^0 \in [b^0a^0, C] \cup [C, Bb^0a^0]$ then an easy argument shows that (a^0, b^0, c^0) can be embedded in 5 moves

Finally it remains to notice that if $c^0 \in R_{a^0 b^0}$ then there will be a point c^1 on the boundary of $R_{a^0 b^0}$ which can be reached in 5 moves and such that $c^0 \in [b^0, c^1]$. The starshapedness of $R_{a^0 b^0}$ shows that (a^0, b^0, c^0) can be reached in 6 moves.

Lemma 8 If P is embeddable and Det $P \geqq \frac{1}{2}$ then $P \in R$.

Proof: We clearly have $I \in R$ and that I is regular. Now let $K_1, \ldots, K_n$ be elementary matrices with Det $K_n \ldots K_1 \geqq \frac{1}{2}$. By Corrollary 6 we have $K_1 \in R$, and since also Det $K_1 \geqq \frac{1}{2}$ we have by Lemma 1 that K_1 is regular. Hence $K_2 K_1 \in R$. Continuing like this we prove that $K_n \ldots K_1 \in R$.

Now any embeddable matrix P with Det $P \geqq \frac{1}{2}$ can be approximated by finite products of elementary matrices, with Det $\geqq \frac{1}{2}$, and since R is closed we have $P \in R$.

Theorem 9 If P is embeddable and Det $P \geqq \frac{1}{2}$ then P is the product of at most 6 elementary matrices.

Proof: Follows from Lemma 1, 7, and 8.

Theorem 10 (Bang-Bang representation). Any embeddable matrix P is a product of a finite number of elementary matrices, where the number is bounded by 6 times the smallest integer larger than or equal to

$$(\ln \tfrac{1}{2})^{-1} \ln \text{Det } P.$$

Proof: Let $P \neq I$ and let $k \in \{1, 2, \ldots\}$ be chosen such that $2^{-k} \leqq \text{Det } P < 2^{-(k-1)}$. If the continuous family $P(s,t)$ embeds P, then we can find time points $0 = t_0 < t_1 < \ldots < t_{k-1} < t_k = 1$ such that $P = \prod_{i=0}^{k-1} P(t_i, t_{i+1})$ and such that Det $P(t_i, t_{i+1}) = \frac{1}{2}$,

$i = 0,1,\ldots,k-2$, Det $P(t_{k-1},1) > \frac{1}{2}$. Each of these k factors is a product of at most 6 elementary matrices. Thus the total number is bounded by 6k, but $k \geqq (\ln \frac{1}{2})^{-1} \ln \text{Det } P > (k-1)$ which proves the result.

<u>Corrollary 11</u> A necessary and sufficient condition that a regular 3 x 3 stochastic matrix P with Det $P \geqq \frac{1}{2}$ is embeddable is that for some permutation (i,j,k) of (1,2,3) we have either

$$p_{ii}p^{ii} \geqq 1 \tag{2.4}$$

or

$$p_{jj}p_{kk}p^{kj} + p_{kj} \leqq 0. \tag{2.5}$$

Proof: Let us first prove that a necessary and sufficient condition for a regular matrix P with Det $P \geqq \frac{1}{2}$ to be embeddable is that $P \in \bigcup_\sigma \{f_\sigma \vee g_\sigma \geqq 0\}$. If P has Det $P \geqq \frac{1}{2}$ and is embeddable then by Lemma 8 $P \in R \subset \bigcup_\sigma \{f_\sigma \vee g_\sigma \geqq 0\}$. If on the other hand $P \in \bigcup_\sigma \{f_\sigma \vee g_\sigma \geqq 0\}$, then it follows from the proof of Lemma 4 that $P \in \text{cl} \bigcup_\sigma \{(f_\sigma \vee g_\sigma) \wedge \text{Det} > 0\}$ but for the matrix $P = (a,b,B)$ or permutations of it. This matrix is not regular however, hence P must be embeddable.

Finally we notice that $f(P) \geqq 0$ can be reformulated as follows:

$$\begin{aligned} f(P) &= (a_3b_2 - a_2b_3)\ (c_1 + c_2) - (a_3 - b_3)\ c_2 \\ &= (a_3b_2 - a_2b_3)c_1 - (a_3b_1 - a_1b_3)c_2 \\ &= \text{Det } (a,b,c)\ \{c_1(-\bar{a}_3) - c_2\bar{b}_3\} = \text{Det } (a,b,c)\ \{c_3\bar{c}_3 - 1\} \end{aligned}$$

where the rows of the matrix P^{-1} are called $\bar{a}, \bar{b}$, and $\bar{c}$.

Now $f(P) \geqq 0 \Leftrightarrow c_3\bar{c}_3 \geqq 1$ or $p_{33}p^{33} \geqq 1$.

Similarly

$$g(P) = b_2c_3(a_1c_2 - a_2c_1) - c_2 \text{ Det } (a,b,c)$$

$$= \text{Det } (a,b,c) \{b_2c_3(-\bar{c}_2) - c_2\} \geqq 0 \Leftrightarrow p_{22}p_{33}p^{32} + p_{32} \leqq 0.$$

Hence it is seen that $\max_{\sigma} f_{\sigma}(P) \vee g_{\sigma}(P) \geqq 0$ is equivalent to (2.4) and (2.5).

3. Discussion.

Although one can argue that this result for 3×3 matrices is not a very general result, it is our belief that the fact that a Bang-Bang representation holds, where one even gets an upper bound for the number of factors needed, is an indication that a similar result holds for any finite state Markov chain, and perhaps for more general control systems and such a result would give a good insight into the structure of the semigroup of embeddable matrices. The methods of the present paper seems hard to generalize.

We believe to have an example of a 3×3 stochastic matrix which is embeddable in 7 but not 6 moves. If we take $a = (1/4 - \varepsilon, 1/4 + \varepsilon, \frac{1}{2})$, $b = (1/4 + \varepsilon, 1/4 - \varepsilon, \frac{1}{2})$ and $c = (1/4 + \delta, 1/4 + \delta, \frac{1}{2} - 2\delta)$ and choose ε and δ suitably. The reason is that in this case Det $(a,b,C) < 0$ which means that (a,b) has been twisted around, such that a is closest to B and b to A. The set R_{ab} is no longer starshaped and this means that the 7th move can take one outside the region reachable in 6 moves. This conjecture will however, not be substantiated further at this point.

We would like to indicate a different proof that seemingly avoids some of the calculations in the appendix at the expense of some vague symmetry and continuity arguments.

We can start by defining for (a,b) in case 1, $S_{ab} = \{c \mid \mathrm{Det}\,(a,b,c) \geqq 0\}$. For (a,b) in case 2, $S_{ab} = \{c \mid (g \vee g_{\sigma_1} \vee g_{\sigma_2}) \wedge \mathrm{Det}\,(a,b,c) \geqq 0\}$. Now let S be the set of matrices P for which there exists a permutation σ such that if $\sigma P \sigma^{-1} = (a,b,c)$ then (a,b) is in case 1 or 2 and $c \in S_{ab}$.

We want to prove that for P regular, $P \in S$ we have $KP \in S$, for K elementary. There is no loss in generality in assuming $P = (a^0,b^0,c^0)$, where (a^0,b^0) is in case 1 or 2 and where c^0 $c^0 \in S_{a^0b^0}$. Now if c^0 is being moved towards a^0 or b^0, the starshapedness shows that also $KP \in S$. If, however, a^0 or b^0 is being moved by K, the proof goes as follows: Assume $KP = (a^1,b^0,c^0)$, $a^1 \in [a^0,c^0]$. If $c^0 \in R_{a^1b^0}$ then clearly $KP \in S$. If not there exists by continuity a point $\tilde{a} \in\,]a^0,a^1[$ such that $c^0 \in r_{\tilde{a}b^0}$, where r_{ab} is the nontrivial part of the boundary of R_{ab}, given by $r_{ab} = [Aa,ab] \cup [Bba,b] \cup h_{ab}$.

It is easy to see from the construction of h_{ab}, that $c \in h_{ab} \Leftrightarrow a \in [Ccb,c] \Leftrightarrow b \in [Cc,ca]$. This can also be stated as follows:

$$c \in r_{ab} \Leftrightarrow a \in r_{bc} \Leftrightarrow b \in r_{ca}.$$

Thus if we find that $c^0 \in r_{\tilde{a}b^0}$ then we also have $\tilde{a} \in r_{b^0c^0} \subset R_{b^0c^0}$ which by starshapedness again implies that $a^1 \in R_{b^0c^0}$ and that $KP \in S$.

In trying to make this argument precise the proof of the paper appeared. An earlier version of a proof appeared in Johansen and Ramsey (1973).

4. Appendix.

We shall prove here that R_{ab} is in fact given by the set stated in Lemma 4 and sketched in Figure 2.

For a function k(a,b,c) we introduce the notation $\{c|k \geqq 0\}$ for the set $\{c|k(a,b,c) \geqq 0\}$ and similarly for $\{a|k \geqq 0\}$ and $\{b|k \geqq 0\}$.

The set R_{ab} is defined as

$$R_{ab} = \text{cl} \bigcup_{\sigma} \{c|(g_{\sigma} \vee f_{\sigma}) \wedge \text{Det} > 0\}$$

and we shall have to investigate all the sets involved.

First of all the set $\{c|f \geqq 0\}$. Since f is linear in c we shall find where it is zero and where it is positive. For c = C and c = ab we find f(a,b,C) = f(a,b,ab) = 0 and for c = A, f(a,b,A) > 0. Hence we find that $\{c|f \wedge \text{Det} \geqq 0\} = \text{co}\{ab,ba,C\}$.

Now take the set $\{a|g \geqq 0\}$. Since g is linear in a this is a halfspace through the point a = c and the point with coordinates

$$(c_2b_3, 0, b_1c_2 - c_1b_2)/(c_2 - b_2(c_1 + c_2))$$

which is just the point Ccb, provided $b_1c_2 - c_1b_2 = \text{Det}(C,b,c) > 0$. The halfspace contains the point A, since $g(A,b,c) = c_2^2b_3 \geqq 0$.

From

$$c_2f(a,b,c) - g(a,b,c) = a_2c_3(b_2c_1 - c_2b_1)$$

it follows that under the same condition as above we have $\{a|f \geqq 0\} \subset \{a|g \geqq 0\}$.

Now let σ_1 be the permutation (312) that sends c into a, a into b, and b into c. From the above we find that

$$\{c|f_{\sigma_1} \geqq 0\} \subset \{c|g_{\sigma_1} \geqq 0\}$$

and that this set is a halfspace through b and Bba, containing C, provided

$$a_3b_1 - a_1b_3 = \text{Det }(a,b,B) > 0,$$

but this is satisfied in case 2.

Now $\{b|g \geqq 0\}$ is also a halfspace. But this time through the points Cc and ca, containing B, provided $c_3a_2 - a_3c_2 \leqq 0$ and $c_1a_2 - a_1c_2 \leqq 0$. The space $\{b|f \geqq 0\}$ is a halfspace through a and Cc containing B and is thus contained in $\{b|g \geqq 0\}$ under the above conditions.

Now let σ_2 be the permutation (231) that takes c into b, b into a, and a into c then $\{c|f_{\sigma_2} \geqq 0\} \subset \{c|g_{\sigma_2} \geqq 0\}$ and this set is given as the halfspace through Aa and ab, containing C, provided $a_1b_3 - a_3b_1 = -\text{Det }(a,b,B) < 0$ and $a_2b_3 - a_3b_2 = \text{Det }(a,b,A) \leqq 0$. Again these conditions are satisfied in case 2.

Notice that also $\{c|f \geqq 0\} \subset \{c|g_{\sigma_2} \geqq 0\}$ which means that we have found all the sets that contribute to R_{ab}. We only have to check that the remaining permutations which interchange a and b do not add anything to R_{ab}.

Let σ_0 be the permutation (213) that interchanges a and b.

Notice that $f_{\sigma_0} = f$, and that $c_1 f_{\sigma_0}(a,b,c) - g_{\sigma_0}(a,b,c) = b_1 c_3(a_1 c_2 - c_1 a_2)$. One can quite easily see that $f(a,b,c) \leqq 0 \Rightarrow a_1 c_2 - a_2 c_1 > 0$ and hence that $g_{\sigma_0}(a,b,c) \leqq 0$. Thus $\{c | g_{\sigma_0} > 0\} \subset \{c | f > 0\}$ and $cl\{c | g_{\sigma_0} > 0\} \subset \{c | f \geqq 0\}$ which has already been accounted for. Notice that $g_{\sigma_0} \geqq 0$ would imply $c_1 f_{\sigma_0} \geqq 0$ which means that either $c_1 = 0$ or $f_{\sigma_0} \geqq 0$. If also $g_{\sigma_0} \geqq 0$ then $c_2 = 0$ or 1 corresponding to C or B. In order to avoid the point B we find $cl\ \{c | g_{\sigma_0} > 0\}$ which does not contain B. This problem only occurs when we consider the quadratic function $g_{\sigma_0}(a,b,\cdot)$ and it is for this reason that R is defined as the closure of an open set.

Figure 3

Now consider the permutation $\sigma_0\sigma_1$. We find

$$a_2 f_{\sigma_0\sigma_1}(a,b,c) - g_{\sigma_0\sigma_1}(a,b,c) = c_2 a_1(b_2 a_3 - a_2 b_3) \geqq 0$$

which means that $\{c | g_{\sigma_0\sigma_1} \geqq 0\} \subset \{c | f_{\sigma_0\sigma_1} \geqq 0\} = \{c | f_{\sigma_1} \geqq 0\} \subset \{c | g_{\sigma_1} \geqq 0\}$. Hence $\sigma_0\sigma_1$ does not give any new contribution.

Finally $\{c | g_{\sigma_0\sigma_2} \geqq 0\}$ is the halfspace through ab and Bb, containing B, and when intersecting with $\{c | \mathrm{Det} \geqq 0\}$ we get the empty set.

This completes the proof of Lemma 4.

5. Acknowledgement.

The solution of this problem would not have been possible without the continuous interest and encouragement of Gerald Goodman, who also critisized an early version of the paper, thereby preventing the results from being published prematurely.

6. References.

[1] Goodman, G. (1969): Private Communication.

[2] Goodman, G. (1970): An intrinsic time for non-stationary finite Markov chains. Z. Wahrscheinlichkeitstheorie verw. Gebiete 16 165-180.

[3] Goodman, G. (1974): Geometric and Control-theoretic approaches to the embedding problem for stochastic matrices (14 pp).

[4] Hazod, W. (1976): Bemerkungen zum Bang-Bang-Problem für stochastische Matrizen. Z. Wahrscheinlichkeitstheorie verw. Gebiete 35, 39-43.

[5] Johansen, S. (1973): A central limit theorem for finite semigroups and its application to the embedding problem for finite state Markov chains. Z. Wahrscheinlichkeitstheorie verw. Gebiete **26**, 171-190.

[6] Johansen, S. (1973): The Bang-Bang problem for stochastic Matrices. Z. Wahrscheinlichkeitstheorie verw. Gebiete **26**, 191-195.

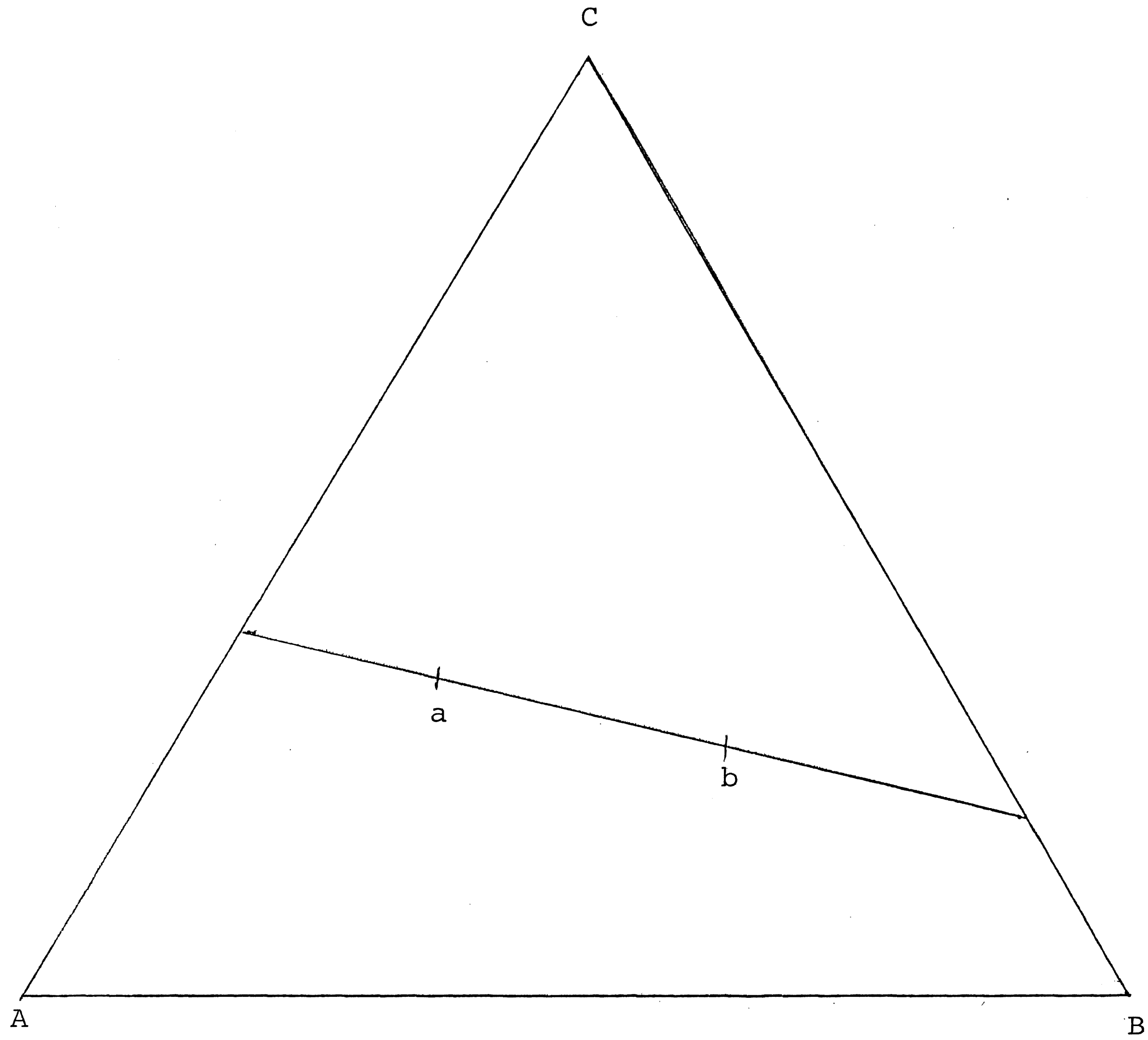
[7] Johansen, S. and Ramsey, F.L. (1973): A representation theorem for imbeddable 3 x 3 stochastic matrices. Tech. Rep. University of Copenhagen 16 pp.

[8] Kingman, J.F.C. and Williams, D. (1973): The combinatorial structure of non-homogeneous Markov chains. Z. Wahrscheinlichkeitstheorie verw. Gebiete **26**, 77-86 (1973).

[9] Krener, A.J. (1974): A generalization of Chow's theorem and the Bang-Bang theorem to nonlinear Control problems. Siam. J. Control **12**, 43-52.

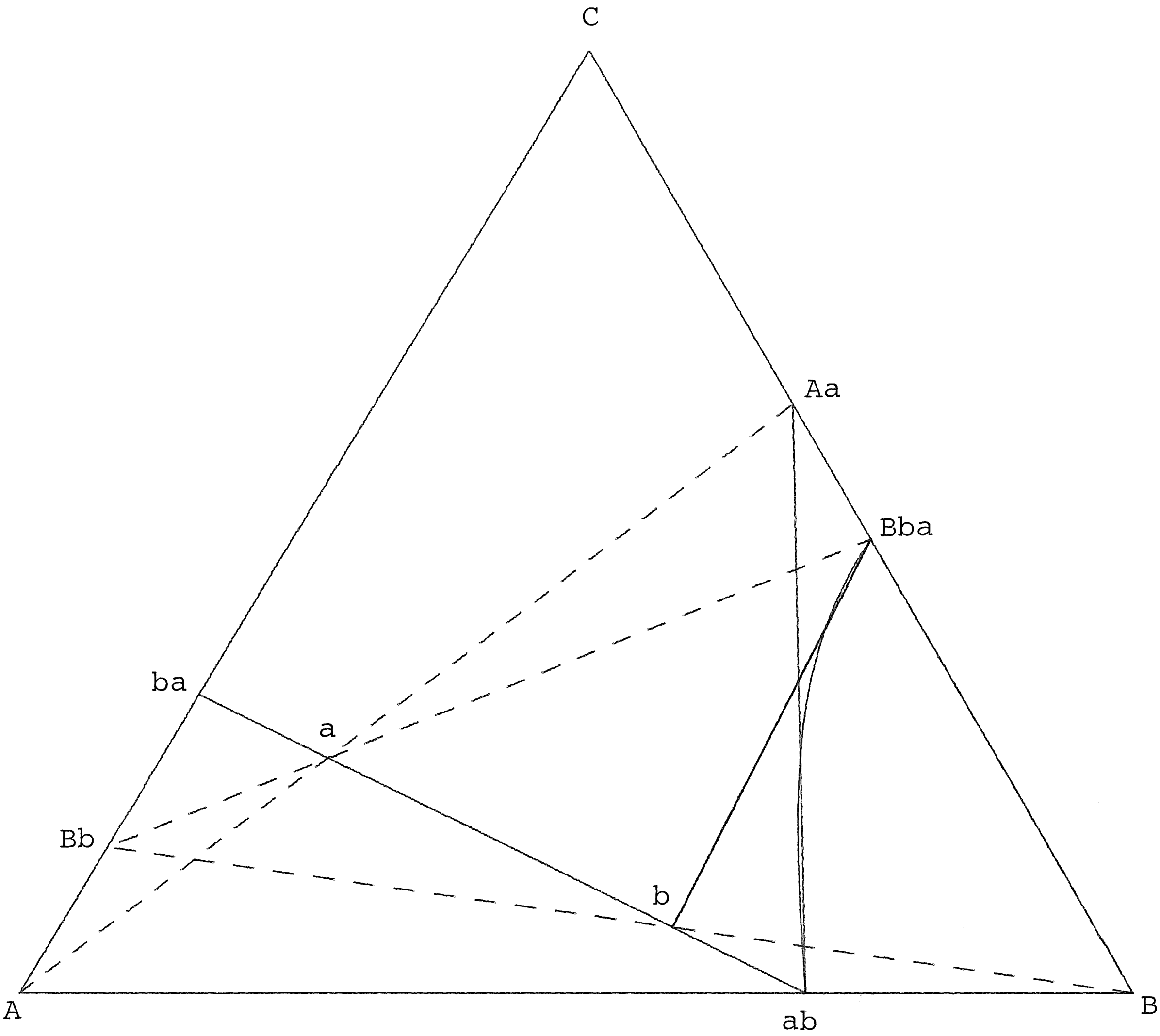
[10] Lee, E.B. and Marcus, L. (1968): Foundations of optimal control, New York: John Wiley.

Figure 1



The set Rab in case 1.

Figure 2

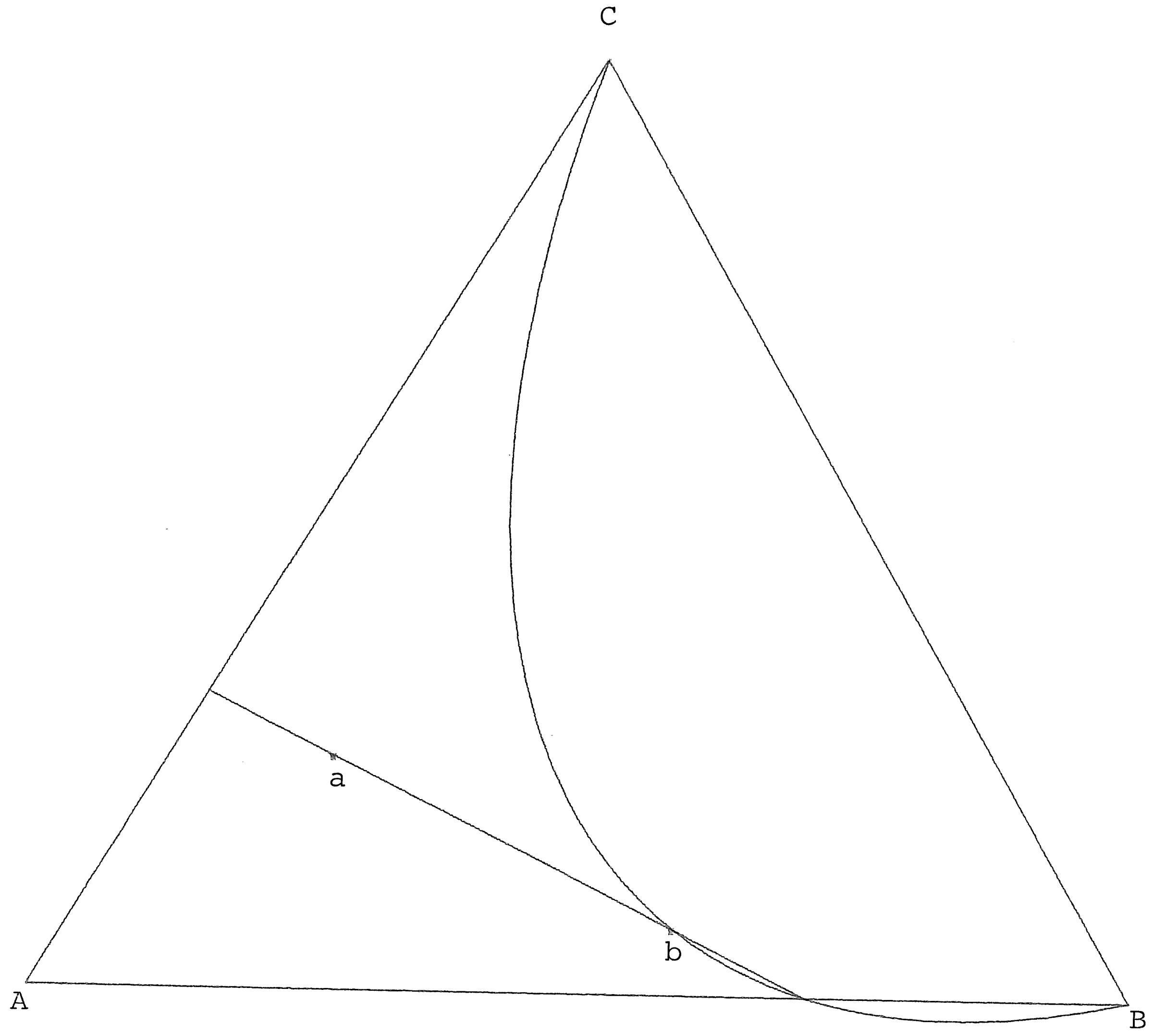


The set Rab in case 2.

a = (0.6 , 0.15, 0.25)

b = (0.38, 0.55, 0.07)

Figure 3



The set $\{c \mid g_{\sigma_1}(a,b,c) \geqq 0,\ \mathrm{Det}\,(a,b,c) \geqq 0\}$ in case 2. Notice that the point c = B belongs to the set. a = (0.6, 0.15, 0.25), b = (0.38, 0.55, 0.07).

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.

No. 4 Keiding, Niels & Lauritzen, Steffen L. : Maximum likelihood estimation of the offspring mean in a simple branching process.

No. 5 Hering, Heinrich: Multitype Branching Diffusions.

No. 6 Aalen, Odd & Johansen, Søren: An Empirical Transition Matrix for Non-Homogeneous Markov Chains Based on Censored Observations.

No. 7 Johansen, Søren: The Product Limit Estimator as Maximum Likelihood Estimator.

No. 8 Aalen, Odd & Keiding, Niels & Thormann, Jens: Interaction Between Life History Events.

No. 9 Asmussen, Søren & Kurtz, Thomas G.: Necessary and Sufficient Conditions for Complete Convergence in the Law of Large Numbers.

No. 10 Dion, Jean-Pierre & Keiding, Niels: Statistical Inference in Branching Processes.

PREPRINTS 1978

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

No. 1 Tjur, Tue: Statistical Inference under the Likelihood Principle.

No. 2 Hering, Heinrich: The Non-Degenerate Limit for Supercritical Branching Diffusions.

No. 3 Henningsen, Inge: Estimation in M/G/1-Queues.

No. 4 Braun, Henry: Stochastic Stable Population Theory in Continuous Time.

No. 5 Asmussen, Søren: On some two-sex population models.

No. 6 Andersen, Per Kragh: Filtered Renewal Processes with a Two-Sided Impact Function.

No. 7 Johansen, Søren & Ramsey, Fred L.: A Bang-Bang Representation for 3x3 Embeddable Stochastic Matrix.