**PREPRINT**

OCT 1972

**10**

Søren Johansen

# The Imbedding Problem for Finite Markov Chains IV

**UNIVERSITY OF COPENHAGEN INSTITUTE OF MATHEMATICAL STATISTICS**

Søren Johansen

# THE BANG-BANG PROBLEM FOR STOCHASTIC MATRICES.

Preprint 1972 No. 10

INSTITUTE OF MATHEMATICAL STATISTICS

UNIVERSITY OF COPENHAGEN

October 1972

## 1. Introduction and Summary

We shall consider the imbedding problem for $n \times n$ stochastic matrices. Such a matrix is called imbeddable if it can occur as transition probability matrix for a time continuous Markov process with n states.

It is proved that any matrix which is in the interior of the set of imbeddable matrices admits a representation as a finite product of exponentials of extreme intensity matrices.

Let $\mathcal{P}$ denote the set of stochastic $n \times n$ matrices and $\mathcal{Q}$ the set of $n \times n$ intensity matrices normalized such that $\mathrm{tr}Q = -1$, $Q \in \mathcal{Q}$.

Let $P \in \mathcal{P}$, then P is imbeddable if there exist a null set N and a measurable function $Q(\circ) : [o,1] \to \mathcal{Q}$ such that the solution $P(\circ)$ of the equation

$$(1.1) \qquad \frac{d}{dt}P(t) = P(t)Q(t), \quad t \notin N$$

$$(1.2) \qquad P(0) = I$$

satisfies the condition

$$(1.3) \qquad P(1) = P.$$

For a proof of this see [1] or [2].

Notice that $\mathcal{Q}$ is a convex compact set and that the extreme points of $\mathcal{Q}$ are characterized by the condition that exactly one off diagonal element is 1.

We shall investigate the role of the extreme intensity-matrices. If $Q(t) = Q$, then the solution to (1.1) and (1.2) is

$$P(t) = e^{tQ}$$

which is called a Poisson matrix when Q is extreme, see [2]. Let $\mathcal{A}$ denote the set of matrices which are finite products of Poisson matrices.

It was proved in [2] that $\mathcal{A}$ is dense in $\mathcal{M}$. In this paper we shall investigate further proporties of $\mathcal{A}$. We can prove that exp $tQ \in \mathcal{A}$, $Q \in \text{int}\,\mathcal{Q}$ (Corollary 2.2), that $\mathcal{A}$ is starshaped (Corollary 2.6), and that int $\mathcal{M} \subset \mathcal{A}$, (Theorem 2.2).

The problem of representing an imbeddable matrix as a finite product of Poisson matrices is known as the Bang-Bang problem in control theory: one can reach P by switching abstrubtly between extreme controllers Q a finite number of times. We therefore say that P admits a Bang-Bang representation.

The problem is also related to the representation of infinitely divisible and infinitely factorizable probability measures on finite semigroups, see [2] and in this sense the Bang-Bang problem is a generalization of the Lévy-Khinchin representation.

## 2. Main Results

The basic tool we shall use is a lemma which follows from the Brouwer fixed-point theorem, see Lee and Markus [1] p.251.

2.1 Lemma. Let f be a continuous map from a compact convex set $\mathcal{B}$ having interior points in $R^N$, into the space $R^N$. Let $P_0$ be a point interior to $\mathcal{B}$ and assume

$$(2.1) \qquad ||f(P) - P|| < ||P-P_0||$$

for all $P \in \partial\mathcal{B}$, then the image $f(\mathcal{B})$ covers $P_0$.

We shall apply this lemma to the compact convex set

$$(2.2) \qquad \mathcal{B}_\varepsilon = \{P \mid ||P-I|| \leq 2\varepsilon\},$$

where

$$||P-I|| = \sup_i \Sigma_j |p_{ij}-\delta_{ij}| = 2\sup_i (1-p_{ii}).$$

Notice that

$$\partial\mathcal{B}_\varepsilon = \mathcal{B}_1 \cup \mathcal{B}_2$$

where

$$(2.3) \qquad \mathcal{B}_1 = \{P \mid \exists (i,j) : p_{ij} = 0\},$$

$$(2.4) \qquad \mathcal{B}_2 = \{P \mid \exists_i : p_{ii} = 1-\varepsilon\}.$$

The function f is defined by

$$(2.5) \qquad f(P) = \prod_{i \neq j} e^{p_{ij} Q_{ij}}$$

where $Q_{ij}$ is the extreme intensity matrix with (i,j)'th element equal to 1. The order in which the multiplication is performed will be fixed throughout.

2.2 Lemma. For any $P \in \mathcal{B}_\varepsilon$ we have

$$(2.6) \qquad ||f(P)-P|| \leq 2(n\varepsilon)^2 e^{2n\varepsilon}.$$

Proof. Since

$$P = I + \sum_{i \neq j} p_{ij} Q_{ij}$$

we get

$$||f(P)-P|| = || \prod_{i \neq j} e^{p_{ij}Q_{ij}} - I - \sum_{i \neq j} p_{ij} Q_{ij} ||.$$

Using an elementary inequality for products of exponentials, see [2], we get that

$$||f(P)-P|| \leq \frac{1}{2}\left(\sum_{i \neq j} p_{ij} ||Q_{ij}||\right)^2 \exp\left(\sum_{i \neq j} p_{ij} ||Q_{ij}||\right)$$

$$\leq 2(n\varepsilon)^2 e^{2n\varepsilon},$$

since

$$||Q_{ij}|| = 2, \quad i \neq j$$

and

$$\sum_{i \neq j} p_{ij} = \Sigma_i (1-p_{ii}) \leq n \sup_i (1-p_{ii}) \leq n\varepsilon.$$

The main idea in using Lemma 2.1 is now that from (2.6) it follows that the left-hand side of (2.1) decreases with $\varepsilon^2$, where as by keeping $P_0$ in the interior of $\mathcal{B}_\varepsilon$ and restricting $P$ to $\mathcal{B}_1$ or $\mathcal{B}_2$, the right-hand side of (2.1) will decrease with $\varepsilon$ and hence (2.1) will be satisfied for $\varepsilon$ sufficiently small.

Let us define

$$(2.7)\qquad \mathcal{P}(p) = \{P \mid p_{ij} \geqq p\},\ 0 < p < \frac{1}{n}.$$

<u>2.3 Proposition</u>. Let p be fixed, then for ε sufficiently small, the matrix

$$P_0 = (1-\varepsilon)I + \varepsilon R,\ R\in\mathcal{P}(p)$$

admits a representation of the form

$$\prod_{i\neq j} e^{p_{ij}Q_{ij}}$$

for some stochastic matrix $P = (p_{ij})$ such that $||I-P|| \leq 2\varepsilon$.

<u>Proof</u>. Since

$$||P_0-I|| = \varepsilon||R-I|| \leq 2\varepsilon(1-p) < 2\varepsilon.$$

and since $P_0$ is interior to $\mathcal{P}$ we have that $P_0 \in \mathrm{int}\mathcal{B}_\varepsilon$.

Now let $P\in\mathcal{B}_1$, and let $p_{ij} = 0$, then

$$(2.8)\qquad ||P-P_0|| \geq |p_{ij}-p_{ij}^{(0)}| = \varepsilon r_{ij} \geq \varepsilon p.$$

If $P\in\mathcal{B}_2$ and $p_{ii} = 1 - \varepsilon$, then we evaluate

$$(2.9)\qquad ||P-P_0|| \geq |p_{ii}-p_{ii}^{(0)}|$$

$$= |1-\varepsilon-(1-\varepsilon)-\varepsilon r_{ii}| \geq \varepsilon p.$$

Let now ε be chosen so small that

$$(2.10)\qquad 2(n\varepsilon)^2 e^{2n\varepsilon} < \varepsilon p$$

then we can combine (2.6), (2.8), (2.9), and (2.10) to see that Lemma 2.1 can be applied and that we can represent $P_0$ in the form f(P) for some $P\in\mathcal{B}_\varepsilon$.

2.4 Corollary.

$$P = e^{tQ} \in \mathcal{A}, \quad Q \in \mathrm{int}\,\mathcal{Q}.$$

Proof. Let us define the stochastic matrix R by

$$Q = R - I$$

then

$$P = (P_k)^k = (\exp(k^{-1}t(R-I)))^k$$

and

$$P_k = I + tk^{-1}(R-I) + B$$

where

$$||B|| \leq \frac{1}{2}k^{-2}||tQ||^2 \exp(k^{-1}||tQ||).$$

Rewriting this we obtain

$$P_k = (1-tk^{-1})I + tk^{-1}(R+kt^{-1}B).$$

Now fix $p = \frac{1}{2}\inf_{i,j} r_{ij}$ then $p > 0$ and for k sufficiently large

$$(R+kt^{-1}B) \in \mathcal{P}(p).$$

Now choose k so large that Proposition 2.3 can be applied with $\varepsilon = tk^{-1}$. We then get the existence of a stochastic matrix $P^{(k)}$ such that

$$P_k = \prod_{i \neq j} e^{p_{ij}^{(k)} Q_{ij}}$$

It follows that

$$P = (P_k)^k = \left(\prod_{i \neq j} e^{p_{ij}^{(k)} Q_{ij}}\right)^k$$

which proves that $P \in \mathcal{A}$.

Let now M be the matrix with entries $n^{-1}$.

2.5 Corollary.

$$(1-\varepsilon)I + \varepsilon M \in \mathcal{A}, \quad 0 \leq \varepsilon < 1.$$

Proof. We define

$$Q = (M-I)(-\ln(1-\varepsilon))$$

then $q_{ij} > 0$ and

$$e^Q = (1-\varepsilon)I + \varepsilon M$$

and we can apply Corollary 2.4.

2.6 Corollary. The set $\mathcal{A}$ is starshaped around M.

Proof. Let $P \in \mathcal{A}$, then since $\mathcal{A}$ is a semigroup we can apply Corollary 2.5 and get that

$$Pe^{(M-I)(-\ln(1-\varepsilon))} = (1-\varepsilon)P + \varepsilon M \in \mathcal{A}, \quad 0 \leq \varepsilon < 1.$$

We shall now prove that the interior of $\mathcal{M}$ has a Bang-Bang representation. The idea here is that the Bang-Bang matrices are dense in $\mathcal{M}$ and that each such matrix P is followed be a small "cone shaped" region $Pf(\mathcal{B}_\varepsilon)$ of Bang-Bang matrices which point towards M.

Hence any matrix $P_0$ in the interior of $\mathcal{M}$ will be surrounded by Bang-Bang matrices and it is proved that one of these will be behind $P_0$ (as viewed from M) in such a way that its "cone" covers $P_0$.

2.7 Theorem. Any imbeddable matrix in the interior of $\mathcal{M}$ has a representation as a finite product of expotentials of extreme intensities or

(2.11) $\quad \text{int}\,\mathcal{M} \subset \mathcal{A}$.

<u>Proof.</u> From Proposition 2.3 it follows that for $0 < p < n^{-1}$, the open set

(2.12) $\quad \mathcal{K} = \{P \mid P = (1-\alpha)I + \alpha R,\ 0 < \alpha \leq \varepsilon,\ r_{ij} > p\}$

is contained in $\mathcal{A}$ for $\varepsilon$ sufficiently small.

Notice for $P \in \mathcal{A}$ we have

$$P\mathcal{K} \subset \mathcal{A}.$$

Now fix $P_0 \in \text{int}\,\mathcal{M}$ and consider the set of matrices in $\mathcal{P}$ from which $P_0$ can be reached by a matrix in $\mathcal{K}$,

(2.12) $\quad \mathcal{O} = \{P \mid P\mathcal{K} \ni P_0\}$.

It is easily seen that

$$\mathcal{O} \subset \{P \mid \text{Det } P > 0\}$$

and since the function $P \to P^{-1}P_0$ is continuous and $\mathcal{K}$ is open it follows that

$$\mathcal{O} = \{P \mid P^{-1}P_0 \in \mathcal{K}\} = P_0\mathcal{K}^{-1}$$

is open.

Now let $\mathcal{N}$ be a neighbourhood of $P_0$ such that

$$\mathcal{N} \subset \text{int}\,\mathcal{M}$$

Let $\phi > 0$ chosen so small that

$$P_{-\phi} = (1+\phi)P_0 - \phi M \in \mathcal{N},$$

then $P_0$ is on the line segment from M to $P_{-\phi}$. Since $\mathcal{K}$ contains the open line segment

$$]I, (1-\varepsilon)I + \varepsilon M[$$

it follows that one can reach $P_0$ from $P_{-\phi}$ using a matrix in $\mathcal{K}$ for $\phi$ sufficiently small. Hence

$$P_{-\phi} \in \mathcal{O}.$$

Now the set $\mathcal{N} \cap \mathcal{O}$ is a non-empty open subset of int $\mathcal{M}$ and hence contains a matrix $P_1$ from $\mathcal{A}$, which was dense in $\mathcal{M}$. But then we can reach $P_1$ using a Bang-Bang controller, since $P_1 \in \mathcal{A}$ and we can reach $P_0$ from $P_1$ using a Bang-Bang matrix from $\mathcal{K}$, since $P_1 \in \mathcal{O}$. Hence $P_0 \in \mathcal{A}$ as was to be proved.

## 3. Acknowledgements.

The author wishes to thank Steffen L. Lauritzen for helpful discussions.

## 3. References.

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

[2] Johansen, S.: A central limit theorem for finite semigroups and its application to the imbedding problem for finite state Markov chains. Preprint, University of Copenhagen and Imperial College (1972).

[3] Lee, E.B. and Markus, L.: Foundations of Optimal Control Theory. John Wiley: New York (1968).