New York Journal of Mathematics

New York J. Math. **1** (1995) 120–129.

Characterizations of Embeddable 3×3 Stochastic Matrices with a Negative Eigenvalue

Philippe Carette

Abstract. The problem of identifying a stochastic matrix as a transition matrix between two fixed times, say $t = 0$ and $t = 1$, of a continuous-time and finite-state Markov chain has been shown to have practical importance, especially in the area of stochastic models applied to social phenomena. The embedding problem of finite Markov chains, as it is called, comes down to investigating whether the stochastic matrix can be expressed as the exponential of some matrix with row sums equal to zero and nonnegative off-diagonal elements. The aim of this paper is to answer a question left open by S. Johansen (1974), i.e., to characterize those stochastic matrices of order three with an eigenvalue $\lambda < 0$ of multiplicity 2.

Contents

1. Introduction

Let the stochastic process $(X(t))_{t\geq 0}$ on a probability triple $(\Omega, \mathcal{A}, \mathbb{P})$ be a continuous time Markov chain with a finite number N of states and time-homogeneous transition probabilities

$$P_{ij}(t) = \mathbb{P}[\, X(t+u) = j \mid X(u) = i\,], \qquad i, j = 1, \dots, N,$$

i.e., which are independent of the time $u \geq 0$. Under the assumption

$$\lim_{t \overset{>}{\to} 0} P_{ii}(t) = 1, \qquad i = 1, \dots, N,$$

Received February 20, 1995.

Mathematics Subject Classification. 60J27,15A51.

Key words and phrases. Continuous-time Markov chain, stochastic matrix, matrix exponential, embedding problem.

ISSN 1076-9803/95

the transition probabilities $P_{ij}(t)$ are gathered into stochastic $N \times N$ matrices $\mathbf{P}(t)$, $t \geq 0$, which satisfy

$$\mathbf{P}(0) = \mathbf{I}, \text{ the identity matrix,} \tag{1}$$

$$\mathbf{P}(t+s) = \mathbf{P}(t)\mathbf{P}(s), \qquad t, s \geq 0. \tag{2}$$

It is a well-known result (see e.g., [Fre71], p. 148) that in this case

$$\frac{d}{dt}\mathbf{P}(0) = \mathbf{R} \text{ and } \mathbf{P}(t) = \exp(\mathbf{R}t), \qquad t \geq 0, \tag{3}$$

where $\mathbf{R}$ is a $N \times N$ matrix satisfying

$$\begin{aligned} R_{ij} &\geq 0, && \text{for } i \neq j, \\ \sum_{j=1}^{N} R_{ij} &= 0, && i = 1, \dots, N. \end{aligned} \tag{4}$$

The (i, j)-th element of the matrix $\mathbf{R}$ represents the transition *intensity* from state i to state j. In general, any matrix $\mathbf{R}$ that satisfies (4) will be called an *intensity matrix*. Thus, the Markov chain is completely determined by its intensity matrix.

Now, an interesting problem emerges as follows. Suppose $(Y(t))_{t=0,1,2,\dots}$ is a discrete time Markov chain with N states and time-homogeneous transition probability matrix $\mathbf{P}$, i.e., the elements of $\mathbf{P}$

$$P_{ij} = \mathbb{P}[\,Y(u+1) = j \,|\, Y(u) = i\,], \qquad i, j = 1, \dots, N \tag{5}$$

are independent of the time $u = 0, 1, 2, \dots$. Then, one can ask the question whether or not this process is a discrete manifestation of an underlying time-homogeneous and continuous N-state Markov chain $(X(s))_{s \geq 0}$, or equivalently, under what circumstances $\mathbf{P}$ has the form

$$\mathbf{P} = \exp \mathbf{R} \tag{6}$$

for some intensity matrix $\mathbf{R}$. If this occurs, we say that $\{Y(t) \,|\, t = 0, 1, 2, \dots\}$ is *embeddable* into a continuous and time-homogeneous Markov chain, or briefly that $\mathbf{P}$ is *embeddable* and that $\mathbf{R}$ *generates* $\mathbf{P}$.

This problem has been acknowledged to have practical relevance when mathematical modelling of social phenomena is concerned (see [Bar82, SS77]).

From a purely mathematical point of view however, the question has already been investigated by Kingman ([Kin62]), who published a simple necessary and sufficient condition (due to D. G. Kendall) for the two state case: $\det \mathbf{P} > 0$. He gave also necessary and sufficient conditions for embeddability of any $N \times N$ stochastic matrix, but they are not applicable in practice. At the time, it was already known that for $N \geq 3$ more than one intensity matrix could generate the same $\mathbf{P}$ ([Spe67]).

However, for the three state case, some simplifications do also occur. Papers [Joh74, Cut73] have shown that necessary and sufficient conditions were in a critical sense dependent upon the nature of the eigenvalues of $\mathbf{P}$. In both of them such conditions were given, except for the case when $\mathbf{P}$ has a negative eigenvalue of multiplicity two.

Since then, the problem has mainly been dealt with in the more general context of time-inhomogeneous Markov chains. (See [FS79, JR79, Fry80, Fry83].)

The purpose of this paper is to fill the gap for 3×3 stochastic matrices with a negative eigenvalue of multiplicity 2, by providing a method to check whether or not they are embeddable and giving an explicit form of the possible intensity matrices that generate them (Theorem 3.3). Another interesting characterization involving a lower bound for the negative eigenvalues is contained in Theorem 3.7.

2. Notations and preliminaries

The set of positive integers will be denoted by $\mathbb{N}$, the set of real numbers by $\mathbb{R}$ and the set of complex numbers by $\mathbb{C}$. The sets of real (resp. complex) $m \times n$ matrices is denoted by $\mathbb{R}^{m\times n}$ ($\mathbb{C}^{m\times n}$) and matrices by bold face capital letters. We shall use the notation $\mathrm{diag}(p, q, r)$ for the diagonal matrix $\begin{pmatrix} p & 0 & 0 \\ 0 & q & 0 \\ 0 & 0 & r \end{pmatrix}$, $p, q, r \in \mathbb{C}$. Finally, the (i, j)-th element of the n-th power $\mathbf{M}^n$ of a square matrix $\mathbf{M}$ shall be written as $M_{ij}^{(n)}$, $n = 2, 3, \dots$

It was already remarked in [Joh74] that a negative eigenvalue of an embeddable stochastic matrix has to be of even algebraic multiplicity. Indeed, in our case, if $\mathbf{P}$ is a 3×3 stochastic matrix with an eigenvalue $\lambda < 0$, then $\mathbf{R}$ has an eigenvalue $\theta \in \mathbb{C}$ with $\mathrm{e}^\theta = \lambda$. But then $\overline{\theta}$ must also be an eigenvalue of $\mathbf{R}$ ($\mathbf{R}$ is real), hence the three distinct eigenvalues of $\mathbf{R}$ are 0, θ and $\overline{\theta}$ forcing $\mathbf{R}$ to be diagonalizable:

$$\mathbf{R} = \mathbf{U}\mathrm{diag}(0, \theta, \overline{\theta})\mathbf{U}^{-1}, \tag{7}$$

where $\mathbf{U}$ is a nonsingular 3×3-matrix. We have then

$$\begin{aligned} \mathbf{P} = \exp \mathbf{R} &= \mathbf{U}\mathrm{diag}(\mathrm{e}^0, \mathrm{e}^\theta, \mathrm{e}^{\overline{\theta}})\mathbf{U}^{-1} \\ &= \mathbf{U}\mathrm{diag}(1, \lambda, \lambda)\mathbf{U}^{-1}. \end{aligned}$$

Consequently, λ has algebraic multiplicity 2 and $-\frac{1}{2} < \lambda < 0$, because $1 + 2\lambda = \mathrm{trace}\, \mathbf{P} \geq 0$. Now $\mathbf{P}^\infty = \lim_{n\to+\infty} \mathbf{P}^n$ certainly exists (as is the case for every embeddable stochastic matrix $\mathbf{P}$, see [Kin62]) and

$$\mathbf{P}^\infty = \mathbf{U}\mathrm{diag}(1, 0, 0)\mathbf{U}^{-1}, \quad \text{since } |\lambda| < 1. \tag{8}$$

This leads to the following useful relation between $\mathbf{P}$ and $\mathbf{P}^\infty$ (see also [Joh74])

$$\mathbf{P} = \mathbf{P}^\infty + \lambda(\mathbf{I} - \mathbf{P}^\infty).$$

More can be said about the nature of $\mathbf{P}^\infty$. Indeed, since $\mathbf{P}^\infty$ is a diagonalizable stochastic matrix with eigenvalues 1 and 0 (the latter of algebraic multiplicity 2), $\mathbf{P}^\infty$ must consist of identical rows, equal to say $\begin{pmatrix} p_1 & p_2 & p_3 \end{pmatrix}$. Furthermore, none of the p_i can be equal to 0: Otherwise, if $p_i = 0$ for some i, then $P_{ii} = \lambda < 0$.

Thus, in seeking characterizations for embeddability of $\mathbf{P}$, we only need to focus our attention on these matrices that have the form $\mathbf{P} = \mathbf{P}(\lambda, \mathbf{P}^\infty)$, where

$$\mathbf{P}(\lambda, \mathbf{P}^\infty) = \mathbf{P}^\infty + \lambda(\mathbf{I} - \mathbf{P}^\infty), \tag{9}$$

with

$$\mathbf{P}^\infty = \begin{pmatrix}1\\1\\1\end{pmatrix}\begin{pmatrix}p_1 & p_2 & p_3\end{pmatrix},$$

$$p_1+p_2+p_3=1, \quad p_i>0, \quad i=1,2,3,$$

$$\text{and } \lambda<0, \quad \lambda \geq \kappa := -\min_i \frac{p_i}{1-p_i}, \tag{10}$$

the last condition being a regularity one ensuring that $\mathbf{P}(\lambda,\mathbf{P}^\infty)$ has nonnegative entries.

3. Characterizations of embeddability

A characterization involving the generating intensity matrix. The idea behind this approach is to seek an explicit form of a general matrix $\mathbf{R}$ with $\exp\mathbf{R} = \mathbf{P}$ and then to submit the elements R_{ij} to the constraints (4). We begin with a few properties of matrix square roots of $\mathbf{P}^\infty - \mathbf{I}$, which we shall need in the proof of our main result (Theorem 3.3).

Lemma 3.1. *Let* $\mathbf{X}$ *be a real square matrix with* $\mathbf{X}^2 = \mathbf{P}^\infty - \mathbf{I}$. *Then*

(a) $\mathbf{X}\mathbf{P}^\infty = \mathbf{P}^\infty\mathbf{X} = \mathbf{0}$

(b) $\exp\Big((2k+1)\pi\mathbf{X}\Big) = 2\mathbf{P}^\infty - \mathbf{I} \quad (k\in\mathbb{N})$

(c) $\sum_j X_{ij} = 0$ *for all* i.

Proof. (a) Using (8), we have

$$\mathbf{P}^\infty - \mathbf{I} = \mathbf{U}\mathrm{diag}(0,-1,-1)\mathbf{U}^{-1},$$

and so (see [Gan60])

$$\mathbf{X} = \mathbf{U}\mathbf{V}\mathrm{diag}(0,\mathrm{i},-\mathrm{i})\mathbf{V}^{-1}\mathbf{U}^{-1},$$

where $\mathbf{V}$ is a real nonsingular matrix that commutes with $\mathrm{diag}(0,-1,-1)$. But then $\mathbf{V}$ commutes also with $\mathrm{diag}(1,0,0)$, and

$$\begin{aligned}\mathbf{P}^\infty\mathbf{X} &= \mathbf{U}\mathbf{V}\mathrm{diag}(1,0,0)\mathbf{V}^{-1}\mathbf{U}^{-1}\mathbf{U}\mathbf{V}\mathrm{diag}(0,\mathrm{i},-\mathrm{i})\mathbf{V}^{-1}\mathbf{U}^{-1}\\ &= \mathbf{U}\mathbf{V}\,\mathbf{0}\,\mathbf{V}^{-1}\mathbf{U}^{-1}\\ &= \mathbf{0}.\end{aligned}$$

Analogously $\mathbf{X}\mathbf{P}^\infty = \mathbf{0}$.

(b) By $\mathbf{X} = \mathbf{U}\mathbf{V}\mathrm{diag}(0,\mathrm{i},-\mathrm{i})\mathbf{V}^{-1}\mathbf{U}^{-1}$, we get

$$\begin{aligned}\exp\Big((2k+1)\pi\mathbf{X}\Big) &= \mathbf{U}\mathbf{V}\mathrm{diag}(1,\mathrm{e}^{(2k+1)\pi\mathrm{i}},\mathrm{e}^{-(2k+1)\pi\mathrm{i}})\mathbf{V}^{-1}\mathbf{U}^{-1}\\ &= \mathbf{U}\mathbf{V}\mathrm{diag}(1,-1,-1)\mathbf{V}^{-1}\mathbf{U}^{-1}\\ &= 2\,\mathbf{U}\mathbf{V}\mathrm{diag}(1,0,0)\mathbf{V}^{-1}\mathbf{U}^{-1} - \mathbf{U}\mathbf{V}\mathrm{diag}(1,1,1)\mathbf{V}^{-1}\mathbf{U}^{-1}\\ &= 2\mathbf{P}^\infty - \mathbf{I}.\end{aligned}$$

(c) By (a), we have $\mathbf{X}\mathbf{P}^{\infty}\begin{pmatrix}1\\1\\1\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix}$. Hence

$$\mathbf{X}\begin{pmatrix}1\\1\\1\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix}.$$

This concludes the proof. □

Corollary 3.2. *If* $\mathbf{R} = \ln|\lambda|(\mathbf{I} - \mathbf{P}^{\infty}) + (2k+1)\pi\mathbf{X}$ *with* $\mathbf{X}^2 = \mathbf{P}^{\infty} - \mathbf{I}$ *and* $k \in \mathbb{N}$*, then*

$$\exp\mathbf{R} = \mathbf{P}^{\infty} + \lambda(\mathbf{I} - \mathbf{P}^{\infty}).$$

Proof. Statement (a) of Lemma 3.1 implies that $\mathbf{X}$ commutes with $\mathbf{I} - \mathbf{P}^{\infty}$, in which case we have

$$\exp\mathbf{R} = \exp\Big(\ln|\lambda|(\mathbf{I} - \mathbf{P}^{\infty})\Big) \cdot \exp\Big((2k+1)\pi\mathbf{X}\Big).$$

We now calculate $\exp\Big(\ln|\lambda|(\mathbf{I} - \mathbf{P}^{\infty})\Big)$ and use the same notations as in the proof of (a), Lemma 3.1:

$$\begin{aligned}\exp\Big(\ln|\lambda|(\mathbf{I} - \mathbf{P}^{\infty})\Big) &= \exp\Big(\mathbf{U}\mathrm{diag}(0, \ln|\lambda|, \ln|\lambda|)\mathbf{U}^{-1}\Big)\\ &= \mathbf{U}\mathrm{diag}(1, |\lambda|, |\lambda|)\mathbf{U}^{-1}\\ &= (1+\lambda)\mathbf{U}\mathrm{diag}(1,0,0)\mathbf{U}^{-1} - \lambda\mathbf{U}\mathrm{diag}(1,1,1)\mathbf{U}^{-1}\\ &= (1+\lambda)\mathbf{P}^{\infty} - \lambda\mathbf{I}.\end{aligned}$$

Hence

$$\begin{aligned}\exp\mathbf{R} &= \big((1+\lambda)\mathbf{P}^{\infty} - \lambda\mathbf{I}\big)(2\,\mathbf{P}^{\infty} - \mathbf{I})\\ &= (1-\lambda)\mathbf{P}^{\infty} + \lambda\mathbf{I},\end{aligned}$$

as $(\mathbf{P}^{\infty})^2 = \mathbf{P}^{\infty}$. □

Theorem 3.3. *Let* $\mathbf{P}(\lambda, \mathbf{P}^{\infty})$ *be as in* (9) *and* (10)*. Then the following conditions are equivalent*:

(a) $\mathbf{P}(\lambda, \mathbf{P}^{\infty})$ *is embeddable.*

(b) *There exist* $k \in \mathbb{N}$ *and* $\mathbf{X} \in \mathbb{R}^{3\times 3}$ *such that* $\mathbf{X}^2 = \mathbf{P}^{\infty} - \mathbf{I}$ *and*

$$\mathbf{R} = \ln|\lambda|(\mathbf{I} - \mathbf{P}^{\infty}) + (2k+1)\pi\mathbf{X} \textit{ is an intensity matrix.}$$

(c) *There exists* $\mathbf{X} \in \mathbb{R}^{3\times 3}$ *such that* $\mathbf{X}^2 = \mathbf{P}^{\infty} - \mathbf{I}$ *and*

$$X_{ij} \geq \frac{1}{\pi}\ln|\lambda|\, p_j, \qquad i \neq j.$$

Proof. (a ⇒ b) Let $\mathbf{P} = \exp\mathbf{R}$, where $\mathbf{R}$ is an intensity matrix. As remarked in the preliminaries section, the eigenvalues of $\mathbf{R}$ are in this case 0, θ_k, and $\overline{\theta_k}$, where θ_k is a complex number satisfying $\mathrm{e}^{\theta_k} = \lambda$, i.e.,

$$\theta_k = \ln|\lambda| + (2k+1)\pi\mathrm{i} \quad (k \in \mathbb{N})$$

and

$$\begin{aligned}\mathbf{R} &= \mathbf{U}\mathrm{diag}(0, \theta_k, \overline{\theta_k})\mathbf{U}^{-1} \quad \text{with } \mathbf{U} \text{ nonsingular}\\ &= \ln|\lambda|\mathbf{U}\mathrm{diag}(0,1,1)\mathbf{U}^{-1} + (2k+1)\pi\mathrm{i}\,\mathbf{U}\mathrm{diag}(0,1,-1)\mathbf{U}^{-1}.\end{aligned} \tag{11}$$

Using (8) and Equation (11), and putting $\mathbf{X} = \mathrm{i}\,\mathbf{U}\mathrm{diag}(0,1,-1)\mathbf{U}^{-1}$, we obtain the desired result.
(b $\Rightarrow$ c) Because $\mathbf{R}$ is an intensity matrix, we must have $R_{ij} \geq 0,\ i \neq j$. This means that, by (b), for $i \neq j$

$$\ln|\lambda|(-p_j) + (2k+1)\pi X_{ij} \geq 0,$$

which yields

$$X_{ij} \geq \frac{1}{(2k+1)\pi}\ln|\lambda|p_j \geq \frac{1}{\pi}\ln|\lambda|p_j.$$

(c $\Rightarrow$ a) In this case, by Lemma 3.1 (b) and (c), $\ln|\lambda|(\mathbf{I} - \mathbf{P}^\infty) + \pi\mathbf{X}$ defines an intensity matrix, say $\mathbf{R}$, which has the property $\exp\mathbf{R} = \mathbf{P}^\infty + \lambda(\mathbf{I} - \mathbf{P}^\infty) = \mathbf{P}(\lambda, \mathbf{P}^\infty)$. The proof is now complete. □

A connection with embeddable stochastic matrices with complex conjugated eigenvalues. Another approach can be given starting from the following observation, which is valid for all stochastic matrices.

Lemma 3.4. *A stochastic matrix* $\mathbf{P}$ *is embeddable if and only if it is the square of a stochastic matrix* $\mathbf{Q}$ *that is also embeddable.*

Proof. If $\mathbf{P}$ is embeddable, then an intensity matrix $\mathbf{R}$ exists with $\exp\mathbf{R} = \mathbf{P}$. Now, $\frac{1}{2}\mathbf{R}$ is also an intensity matrix, hence the matrix $\mathbf{Q} = \exp\frac{1}{2}\mathbf{R}$ is stochastic (see [Fre71], p. 125), embeddable, and has the property $\mathbf{Q}^2 = \mathbf{P}$.

Conversely, suppose that an embeddable stochastic matrix $\mathbf{Q}$ exists with $\mathbf{Q}^2 = \mathbf{P}$. Then $\mathbf{Q} = \exp\mathbf{R}$ for some intensity matrix $\mathbf{R}$ and

$$\mathbf{P} = \mathbf{Q}^2 = (\exp\mathbf{R})^2 = \exp(2\,\mathbf{R}),$$

so $\mathbf{P}$ is embeddable. □

This is interesting, because any stochastic 'square root' $\mathbf{Q}$ of $\mathbf{P}(\lambda, \mathbf{P}^\infty)$ must have eigenvalues 1, $\mathrm{i}\sqrt{|\lambda|}$ and $-\mathrm{i}\sqrt{|\lambda|}$. The following characterization of embeddable stochastic matrices of order three with complex conjugate eigenvalues, has been proven in [Joh74].

Theorem 3.5. [Joh74]. *A stochastic matrix* $\mathbf{P}$ *of order three with eigenvalues* $\mathrm{e}^{\alpha+\mathrm{i}\beta}$ *and* $\mathrm{e}^{\alpha-\mathrm{i}\beta}$, $0 < \beta < \pi$, *can be embedded if and only if one of the following conditions holds.*

$$\text{(12)}\quad P_{ij}^{(2)}\Big(\beta(\mathrm{e}^\alpha\cos\beta - 1) - \alpha\mathrm{e}^\alpha\sin\beta\Big) \geq P_{ij}\Big(\beta(\mathrm{e}^{2\alpha}\cos 2\beta - 1) - \alpha\mathrm{e}^{2\alpha}\sin 2\beta\Big) \qquad \textit{for all } i \neq j$$

$$\text{(13)}\quad P_{ij}^{(2)}\Big((\beta - 2\pi)(\mathrm{e}^\alpha\cos\beta - 1) - \alpha\mathrm{e}^\alpha\sin\beta\Big) \geq P_{ij}\Big((\beta - 2\pi)(\mathrm{e}^{2\alpha}\cos 2\beta - 1) - \alpha\mathrm{e}^{2\alpha}\sin 2\beta\Big) \qquad \textit{for all } i \neq j$$

A direct application of this result now yields the following characterization.

Theorem 3.6. $\mathbf{P}(\lambda, \mathbf{P}^\infty)$ *is embeddable if and only if there exists a stochastic matrix* $\mathbf{Q}$ *of order three such that* $\mathbf{Q}^2 = \mathbf{P}(\lambda, \mathbf{P}^\infty)$ *and one of the following conditions holds.*

$$Q_{ij} \geq (1 + \frac{\sqrt{|\lambda|}}{\pi} \ln|\lambda|)p_i \qquad \textit{for all } i \neq j \tag{14}$$

$$Q_{ij} \leq (1 - \frac{\sqrt{|\lambda|}}{3\pi} \ln|\lambda|)p_i \qquad \textit{for all } i \neq j \tag{15}$$

Proof. In view of Theorem 3.5, Lemma 3.4 and the observations made about the eigenvalues of any stochastic square root $\mathbf{Q}$ of $\mathbf{P}(\lambda, \mathbf{P}^\infty)$, we have, with the notations used in Theorem 3.5, $\alpha = \ln\sqrt{|\lambda|}$, $\beta = \frac{\pi}{2}$, and $Q_{ij}^{(2)} = (1-\lambda)p_i$. Inequality (12) then becomes

$$Q_{ij} \geq \frac{1 + \frac{\sqrt{|\lambda|}}{\pi}\ln|\lambda|}{1+|\lambda|}(1-\lambda)p_i = (1 + \frac{\sqrt{|\lambda|}}{\pi}\ln|\lambda|)p_i, \quad i \neq j,$$

and inequality (13)

$$Q_{ij} \leq \frac{1 - \frac{\sqrt{|\lambda|}}{3\pi}\ln|\lambda|}{1+|\lambda|}(1-\lambda)p_i = (1 - \frac{\sqrt{|\lambda|}}{3\pi}\ln|\lambda|)p_i, \quad i \neq j. \quad \square$$

It may be interesting for the sake of consistency to remark the equivalence of the characterizations in Theorems 3.3 and 3.6.
Suppose (14) holds. Then for $i \neq j$,

$$Q_{ij} \geq (1 + \frac{\sqrt{|\lambda|}}{\pi}\ln|\lambda|)p_i$$
$$\Updownarrow$$
$$\frac{1}{\sqrt{|\lambda|}}(Q_{ij} - p_i) \geq \frac{1}{\pi}\ln|\lambda| p_i.$$

Consequently, the matrix $\mathbf{X}$ defined as

$$\mathbf{X} = \frac{1}{\sqrt{|\lambda|}}(\mathbf{Q} - \mathbf{P}^\infty) \tag{16}$$

is the one that satisfies (c) of Theorem 3.3, for it has also the property $\mathbf{X}^2 = \mathbf{P}^\infty - \mathbf{I}$, which remains to be shown:

$$\begin{aligned}
\mathbf{X}^2 &= \frac{1}{|\lambda|}(\mathbf{Q} - \mathbf{P}^\infty)^2 \\
&= \frac{1}{|\lambda|}(\mathbf{Q}^2 - 2\mathbf{Q}\mathbf{P}^\infty + (\mathbf{P}^\infty)^2) \quad (\mathbf{Q} \text{ and } \mathbf{P}^\infty \text{ commute}) \\
&= \frac{1}{|\lambda|}(\mathbf{P}(\lambda, \mathbf{P}^\infty) - 2\mathbf{P}^\infty + \mathbf{P}^\infty) \\
&= \frac{1}{|\lambda|}(\mathbf{P}(\lambda, \mathbf{P}^\infty) - \mathbf{P}^\infty) \\
&= \mathbf{P}^\infty - \mathbf{I} \quad \text{by (9) and } \lambda < 0.
\end{aligned}$$

If (15) holds, however, we'll have to put

$$\mathbf{X} = -\frac{1}{\sqrt{|\lambda|}}(\mathbf{Q} - \mathbf{P}^\infty)$$

to arrive at the same conclusion.

On the other hand, one can easily show using the power series definition

$$\exp(\mathbf{A}) = \sum_{p=0}^{\infty} \frac{1}{p!}\mathbf{A}^p$$

that, starting from (c) Theorem 3.3, the relation (16) defines a matrix $\mathbf{Q}$ that is stochastic since it can be written as the exponential of an intensity matrix [1] :

$$\mathbf{Q} = \sqrt{|\lambda|}\mathbf{X} + \mathbf{P}^\infty = \exp\Big(\frac{1}{2}(\ln|\lambda|(\mathbf{I} - \mathbf{P}^\infty) + \pi\mathbf{X})\Big).$$

Furthermore, $\mathbf{Q}^2 = \mathbf{P}(\lambda, \mathbf{P}^\infty)$ by Corollary 3.2. Finally, $\mathbf{Q}$ satisfies (14) by its very definition.

A characterization involving a lower bound for the negative eigenvalue. Instead of seeking $\mathbf{R}$, we put the following question. For a fixed $\mathbf{P}^\infty$ that has the property (10), what are the possible values of $\lambda < 0$ that make the stochastic matrix $\mathbf{P}(\lambda, \mathbf{P}^\infty)$ embeddable? As a consequence of Theorem 3.3 among other things, the following result emerges. Recall the definition of κ from (10).

Theorem 3.7. *Let $\mathbf{P}^\infty$ be defined by* (10)*, then there exists $\Lambda \in]\kappa, 0[$, depending on $\mathbf{P}^\infty$, such that*

$$\mathbf{P}(\lambda, \mathbf{P}^\infty) \text{ embeddable} \quad \Leftrightarrow \quad \Lambda \leq \lambda < 0.$$

Proof. Define the map $F : [\kappa, 0[\to \mathcal{M} : \lambda \mapsto \mathbf{P}(\lambda, \mathbf{P}^\infty)$, where $\mathcal{M}$ is the set of 3×3 stochastic matrices with positive determinant. Let $\mathcal{P}$ be the set of all embeddable stochastic 3×3 matrices. We have then by the nature of the condition (c) of Theorem 3.3

$$\lambda \in F^{-1}(\mathcal{P}) \Rightarrow [\lambda, 0[\subset F^{-1}(\mathcal{P}). \tag{17}$$

Also, if we take an arbitrary real square root $\mathbf{X}$ of $\mathbf{P}^\infty$ and a λ_0 such that $\ln|\lambda_0| \leq \min_{i \neq j} \frac{\pi X_{ij}}{p_j}$, then $\mathbf{P}(\lambda_0, \mathbf{P}^\infty)$ is embeddable, so

$$F^{-1}(\mathcal{P}) \neq \emptyset. \tag{18}$$

In addition, it was proven in [Kin62] that $\mathcal{P}$ is closed in $\mathcal{M}$, so by continuity of F, we have that $F^{-1}(\mathcal{P})$ is closed in $[\kappa, 0[$. Hence, in conjunction with (17) and (18), $F^{-1}(\mathcal{P}) = [\Lambda, 0[$, for some $\Lambda \in [\kappa, 0[$. We still have to show that $\Lambda \neq \kappa$. Suppose $\Lambda = \kappa$. Then $\kappa \in F^{-1}(\mathcal{P})$ and $\mathbf{P}(\kappa, \mathbf{P}^\infty)$ is embeddable. But then there exists k such that $P_{kk}(\kappa, \mathbf{P}^\infty) = 0$, so we can apply Ornstein's Theorem ([Chu67], II.5, Theorem 2) which states that whenever a stochastic matrix $\mathbf{Q}$, with $Q_{ij} = 0$ for some i and j, is embeddable, then $Q_{ij}^{(n)} = 0$, $n \geq 1$ and in particular $Q_{ij}^{\infty} = 0$. Consequently $\lim_{n \to +\infty} P_{kk}^{(n)}(\kappa, \mathbf{P}^\infty) = p_k = 0$, giving a contradiction. □

[1] The exponential of an intensity matrix is always stochastic. For a proof, see [Fre71], p. 151.

4. Illustration of the results

Let $p_1 = p_2 = p_3 = 1/3$, then $\kappa = -1/2$. For $\mathbf{P}(\lambda, \mathbf{P}^\infty)$ with $-1/2 < \lambda < 0$ to be embeddable, there must exist by Theorem 3.3 a matrix $\mathbf{X} \in \mathbb{R}^{3\times 3}$ satisfying $\mathbf{X}^2 = \mathbf{P}^\infty - \mathbf{I}$ such that

$$X_{ij} \geq \frac{1}{3}c_\lambda, \quad i \neq j, \quad \text{with } c_\lambda = \frac{\ln|\lambda|}{\pi}. \tag{19}$$

By [Gan60], such a matrix $\mathbf{X}$ assumes the form

$$\mathbf{X} = \mathbf{X}(u,v) = \mathbf{V} \begin{pmatrix} 0 & 0 & 0 \\ 0 & u & -\frac{1+u^2}{v} \\ 0 & v & -u \end{pmatrix} \mathbf{V}^{-1}, \quad u, v \in \mathbb{R},\ v \neq 0, \tag{20}$$

where $\mathbf{V}$ is a nonsingular matrix satisfying

$$\mathbf{P}^\infty - \mathbf{I} = \mathbf{V}\mathrm{diag}(0,-1,-1)\mathbf{V}^{-1}.$$

It is easy to check that we can take

$$\mathbf{V} = \begin{pmatrix} 1 & 1/3 & 1/3 \\ 1 & -1/3 & 0 \\ 1 & 0 & -1/3 \end{pmatrix},$$

whence

$$\mathbf{X}(u,v) = \frac{1}{3} \begin{pmatrix} -\frac{u^2-v^2+1}{v} & -\frac{u^2+2v^2+3uv+1}{v} & \frac{2u^2+v^2+3uv+2}{v} \\ \frac{u^2-uv+1}{v} & \frac{u^2+2uv+1}{v} & -\frac{2u^2+uv+2}{v} \\ u-v & u+2v & -2u-v \end{pmatrix}. \tag{21}$$

We shall now show that $\mathbf{P}(\lambda, \mathbf{P}^\infty)$ cannot be embeddable for values of λ with $-2/\sqrt{5} < c_\lambda < 0$.

Suppose $\mathbf{P}(\lambda, \mathbf{P}^\infty)$ is embeddable. Then according to (19), there exist real numbers u and v, $v \neq 0$ such that

$$X_{ij}(u,v) \geq \frac{1}{3}c_\lambda, \quad i \neq j.$$

Using (21), these conditions become algebraic inequalities in u and v. Among others:

(a) $-u + v \leq -c_\lambda$ for $(i,j) = (3,1)$
(b) $u + 2v \geq c_\lambda$ for $(i,j) = (3,2)$
(c) $X_{12}(u,v) + X_{21}(u,v) \geq 2c_\lambda/3$, yielding $-2u - v \geq c_\lambda$
(d) $|v| \geq 2c_\lambda + 2\sqrt{1+c_\lambda^2}$ for $(i,j) = (1,2)$ and $(2,1)$

It is a straightforward matter to see that (a), (b), (c) and (d) together imply

$$2c_\lambda + 2\sqrt{1+c_\lambda^2} \leq -c_\lambda.$$

The contradiction lies in the fact that this inequality cannot be satisfied if $-2/\sqrt{5} < c_\lambda < 0$. Hence for $\mathbf{P}(\lambda, \mathbf{P}^\infty)$ to be embeddable, one must necessarily have $c_\lambda \leq -2/\sqrt{5}$ or $\lambda \geq -\mathrm{e}^{-2\pi/\sqrt{5}}$. In this case, the lower bound Λ from Theorem 3.7 must satisfy $\Lambda \geq -\mathrm{e}^{-2\pi/\sqrt{5}}$. In fact, $\Lambda = -\mathrm{e}^{-\pi\sqrt{3}}$, which is a consequence of the following more general result that provides a formula to express Λ in terms of p_1, p_2 and p_3:

$$\boxed{\left(\frac{\ln|\Lambda|}{\pi}\right)^{-2} = \begin{cases} 4p/s^2 - 1 & \text{if } p_m(1-p_m) > s/2 \\ p_m/(1-p_m) & \text{otherwise} \end{cases}}$$

where m is an index such that $p_m = \min_i p_i$, $p = p_1p_2p_3$ and $s = p_1p_2 + p_1p_3 + p_2p_3$. The proof uses the characterization in Theorem 3.3 and is very technical, so it will be omitted here.

Acknowledgement. The author is indebted to the referee for his suggestions resulting in an enhancement of the manuscript.

References

[Bar82] D. J. Bartholomew, *Stochastic Models for Social Processes*, third ed., John Wiley & Sons, Chichester, 1982.

[Chu67] K. L. Chung, *Markov Chains with Stationary Transition Probabilities*, second ed., N. Y., Springer, 1967.

[Cut73] James R. Cuthbert, *The logarithm function for finite-state Markov semi-groups*, J. London Math. Soc. (2) **6** (1973), 524–532.

[Fre71] David Freedman, *Markov Chains*, Holden-Day, 1971.

[Fry80] Halina Frydman, *The embedding problem for Markov chains with three states*, Math. Proc. Camb. Phil. Soc. **87** (1980), 285–294.

[Fry83] Halina Frydman, *On a number of Poisson matrices in bang-bang representations for* 3×3 *embeddable matrices*, J. Multivariate Analysis **13(3)** (1983), 464–472.

[FS79] Halina Frydman and Burton Singer, *Total positivity and the embedding problem for Markov chains*, Math. Proc. Camb. Phil. Soc. **86** (1979), 339–344.

[Gan60] F. R. Gantmacher, *The Theory of Matrices*, vol. one, Chelsea Publishing Company, N. Y., 1960.

[Kin62] J. F. C. Kingman, *The imbedding problem for finite Markov chains*, Z. Wahrscheinlichkeitstheorie **1** (1962), 14–24.

[JR79] Søren Johansen and Fred L. Ramsey, *A bang-bang representation for* 3×3 *embeddable stochastic matrices*, Z. Wahrscheinlichkeitstheorie verw. Gebiete **47** (1979), 107–118.

[Joh74] Søren Johansen, *Some results on the imbedding problem for finite Markov chains*, J. London Math. Soc. (2) **8** (1974), 345–351.

[Spe67] J. M. O. Speakman, *Two Markov chains with a common skeleton*, Z. Warscheinlichkeitstheorie **7** (1967), 224.

[SS77] Burton Singer and Seymour Spilerman, *Fitting stochastic models to longitudinal survey data – some examples in the social sciences*, Bull. Int. Statist. Inst. (3) **47** (1977a), 283–300.

Centrum voor Statistiek en Operationeel Onderzoek, Vrije Universiteit Brussel, Belgium

phcarett@vub.ac.be