Proceedings of the World Congress on Engineering 2014 Vol II,
WCE 2014, July 2 - 4, 2014, London, U.K.

On the Embedding Problem for Three-state Markov Chains

Marie-Anne Guerry, Vrije Universiteit Brussel

***Abstract*—The present paper investigates the embedding problem for time-homogeneous Markov chains. A discrete-time Markov chain with time unit 1 is embeddable in case there exists a compatible Markov chain regarding time unit $\frac{1}{m}$ (with $m \in \mathbb{N}, m \geq 2$). An embeddable Markov chain has a transition matrix for which there exists an m-th root that is a probability matrix. The present paper examines the embedding problem for discrete-time Markov chains with three states. Sufficient embedding conditions are presented in case of a diagonalizable transition matrix with all eigenvalues nonnegative.**

***Index Terms*—Markov chain; embedding problem; matrix root.**

I. Introduction

THE use of discrete-time Markov chains as modelling tool is well-known and widespread. Discrete-time Markov models are intensively used in engineering ([13], [3]), and in other fields as there are manpower planning ([1], [5], [10]) and finance ([2]). A discrete-time Markov chain enables to describe the system on subsequent epochs of a discrete set of times $\{0, 1, ..., t, t + 1, ...\}$. A Markov chain is characterized by the transition probabilities regarding time intervals with unit 1 between the states of the system. The transition probabilities are assumed only to depend on the current state, and not in addition on the states in previous epochs. A discrete-time Markov model is considered with a finite number of states $S_1, ..., S_k$ and that is time-homogeneous with transition matrix $\mathbf{P} = (p_{ij})$. A transition matrix has all its elements nonnegative and all the row sums equal to 1 and is therefore a probability matrix. Let us denote $n_i(t)$, the number of items in state S_i at time t, and $n_{ij}(t, t + 1)$, the number of items of state S_i at time t that are in state S_j at time $t + 1$. In case data is available regarding $n_i(t)$ and $n_{ij}(t, t + 1)$ for $t = 0, 1, ..., T - 1$, the transition probability p_{ij} can be estimated by the maximum likelihood estimator $\hat{p}_{ij} = \frac{\sum_{t=0}^{T-1} n_{ij}(t,t+1)}{\sum_{t=0}^{T-1} n_i(t)}$ ([1]). In case for example the time unit is one year and information is available on annual base for some subsequent years, $\hat{p}_{ij}$ provides an estimation for the transition probability p_{ij} from state i to state j in one year. When there is no data available on time intervals of 6 months, but nevertheless insights would be useful regarding the transition probabilities a_{ij} on semi-annual base, the question is whether a probability matrix $\mathbf{A} = (a_{ij})$ does exist that satisfies $\mathbf{A}^2 = \mathbf{P}$. If this is the case, for the Markov chain with time unit 1 and transition matrix $\mathbf{P}$ there does exist a compatible Markov chain with time unit 0.5, and the Markov chain with transition matrix $\mathbf{P}$ is said to be embedded in the Markov chain with transition matrix $\mathbf{A}$ ([4]). More in general, a Markov chain with transition matrix $\mathbf{P}$ is embeddable in case for a natural number $m \in \mathbb{N}$, $m \geq 2$ there does exist a probability matrix $\mathbf{A}$ that is an m-th root of $\mathbf{P}$, i.e. that satisfies $\mathbf{A}^m = \mathbf{P}$. For an embeddable Markov chain with transition matrix $\mathbf{P}$ a probability m-th root provides information on the transition probabilities regarding a time interval with length $\frac{1}{m}$. In this way probability m-th roots of the transition matrix of an embeddable Markov chain are useful in practice.

The embedding problem is first introduced by Elfving ([4]). The embedding problem for discrete-time Markov chains and the existence of probability roots of transition matrices are investigated in [12], [9], [7] and [8]. Especially for Markov chains with two states detailed insights are already known: Necessary and sufficient embedding conditions are formulated and the probability m-th roots are described in analytic form ([8], [6]). For three-state Markov chains so far some insights are published: In He and Gunn (2003) all real root matrices that are functions of a (3×3) transition matrix are presented. Their study focuses on real root matrices without requiring that these roots itself are probability matrices ([8]). Higham and Lie formulated necessary embedding conditions based on the set of all eigenvalues of (3×3) probability matrices ([9]).

In the present paper the embedding problem is investigated for (3×3) transition matrices that are diagonalizable and that have all eigenvalues real and nonnegative. Sufficient embedding conditions are formulated based on the projections and the spectral decomposition of the transition matrix. Illustrations are provided to demonstrate the practical use of the proven properties.

II. Embedding conditions

A. Necessary embedding conditions

In studying sufficient embedding conditions, the discussion on the existence of probability roots can be restricted to transition matrices that satisfy necessary embedding conditions. For this reason, in this section necessary embedding conditions for (3×3) transition matrices are overviewed.
In case m is even, $det(\mathbf{P}) \geq 0$ is a necessary condition to have probability m-th roots for the probability matrix $\mathbf{P}$: In case $\mathbf{P} = \mathbf{A}^m$ then $det(\mathbf{P}) = (det(\mathbf{A}))^m \geq 0$. For a (3×3) matrix $\mathbf{P}$ with eigenvalues $\lambda_1, \lambda_2, \lambda_3$ holds that $det(\mathbf{P}) = \lambda_1\lambda_2\lambda_3$. Besides a probability matrix has 1 as eigenvalue ([11]). By denoting $\lambda_1 = 1$ the determinant of $\mathbf{P}$ can be expressed as $det(\mathbf{P}) = \lambda_2\lambda_3$. In this way the condition $det(\mathbf{P}) \geq 0$ holds for a probability matrix $\mathbf{P}$ with eigenvalues $\lambda_1, \lambda_2, \lambda_3$ satisfying $\lambda_1 = 1$ and $\lambda_2\lambda_3 \geq 0$.

Furthermore the fact that $\mathbf{P} = \mathbf{A}^m$ with $1, \mu_2, \mu_3$ the eigenvalues of the probability matrix $\mathbf{A}$, results in eigenvalues of $\mathbf{P}$ equal to $1, \lambda_2 = (\mu_2)^m, \lambda_3 = (\mu_3)^m$. Therefore,

Manuscript received March 7, 2014; revised March 25, 2014.
M.A Guerry is member of the Department Business, Technology and Operations at the Vrije Universiteit Brussel e-mail: maguerry@vub.ac.be

in case m is even and the eigenvalues are real numbers, the eigenvalues λ_2 and λ_3 of $\mathbf{P}$ are both nonnegative.

From Higham and Lin (2011) it is known that for an even number m a real m-th root of $\mathbf{P}$ can only exist in case $\mathbf{P}$ has an even number of Jordan blocks of each size for every negative eigenvalue ([9], Theorem 2.3). Therefore the combination of m even together with a negative eigenvalue of $\mathbf{P}$ can only result in a real root of $\mathbf{P}$ in case of a negative eigenvalue with algebraic multiplicity 2 and Jordan blocks of size (1×1). Consequently in that situation $\mathbf{P}$ is diagonalizable.

Having knowledge of these necessary conditions for the embeddability of probability matrices, the goal is to find supplementary conditions that guarantee that at least one probability root does exist. In the present paper the study is focused on diagonalizable transition matrices with nonnegative eigenvalues.

B. *Sufficient embedding conditions*

Sufficient embedding conditions for a probability matrix $\mathbf{P}$ concern conditions that guarantee the existence of at least one probability m-th root of $\mathbf{P}$. Such a probability m-th root is a probability matrix $\mathbf{A}$ satisfying $\mathbf{A}^m = \mathbf{P}$. Searching for probability roots of $\mathbf{P}$ can be organized in two steps. Firstly roots of $\mathbf{P}$ can be found within the less restrictive set of row-normalized matrices, which are matrices with all row sums equal to 1. Secondly conditions can be formulated under which a row-normalized root of $\mathbf{P}$ has all elements nonnegative. In this section sufficient embedding conditions are investigated for a (3×3) probability matrix $\mathbf{P}$ that is diagonalizable and that has all eigenvalues nonnegative.

For $\mathbf{P}$ with eigenvalues $\lambda_1 = 1 \geq \lambda_2 \geq \lambda_3 \geq 0$ that is diagonalizable, holds that $\mathbf{P} = \mathbf{QDQ^{-1}}$ with diagonal matrix

$$\mathbf{D} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

The transformation matrix $\mathbf{Q}$ and its inverse $\mathbf{Q}^{-1}$ satisfy:
$\mathbf{Q} = (\mathbf{R_1}\, \mathbf{R}_{\lambda_2}\, \mathbf{R}_{\lambda_3})$ with $\mathbf{R_1}, \mathbf{R}_{\lambda_2}, \mathbf{R}_{\lambda_3}$ right eigenvectors of $\mathbf{P}$ with respectively eigenvalue $1, \lambda_2, \lambda_3$ and
$\mathbf{Q}^{-1} = \begin{pmatrix} \mathbf{L}_1' \\ \mathbf{L}_{\lambda_2}' \\ \mathbf{L}_{\lambda_3}' \end{pmatrix}$ with $\mathbf{L}_1', \mathbf{L}_{\lambda_2}', \mathbf{L}_{\lambda_3}'$ left eigenvectors of $\mathbf{P}$ with respectively eigenvalue $1, \lambda_2, \lambda_3$.

The diagonalizable matrix $\mathbf{P}$ has a spectral decomposition $\mathbf{P} = \mathbf{P_1} + \lambda_2 \mathbf{P_2} + \lambda_3 \mathbf{P_3}$. The k-th projection is defined as $\mathbf{P_k} = \mathbf{QI_{kk}Q^{-1}}$ with $\mathbf{I_{kk}}$ the (3×3) matrix with the kk-th element equal to 1 and all the other elements equal to 0. The projections $\mathbf{P_1}, \mathbf{P_2}$ and $\mathbf{P_3}$ satisfy the following properties:

- $\mathbf{P_i P_j} = \mathbf{0}\ \ \forall i \neq j \in \{1, 2, 3\}$
- $\mathbf{P_i P_i} = \mathbf{P_i}\ \ \forall i \in \{1, 2, 3\}$
- $\mathbf{P_1} + \mathbf{P_2} + \mathbf{P_3} = \mathbf{I}$ with $\mathbf{I}$ the identity matrix.

These properties result in the fact that for $m \in \mathbb{N}, m \geq 2$ holds $\mathbf{P}^m = (\mathbf{P_1} + \lambda_2 \mathbf{P_2} + \lambda_3 \mathbf{P_3})^m = \mathbf{P_1} + \lambda_2^m \mathbf{P_2} + \lambda_3^m \mathbf{P_3}$.
In Lemma 1 and Lemma 2 some further properties of the projections are presented.

Lemma 1 For a diagonalizable probability matrix $\mathbf{P} = \mathbf{QDQ^{-1}}$ with spectral decomposition $\mathbf{P_1} + \lambda_2 \mathbf{P_2} + \lambda_3 \mathbf{P_3}$, the projections $\mathbf{P_1}, \mathbf{P_2}, \mathbf{P_3}$ can be expressed, as follows, in terms of left and right eigenvectors:

$$\mathbf{P_1} = \mathbf{R_1 L_1'},\ \mathbf{P_2} = \mathbf{R}_{\lambda_2} \mathbf{L}_{\lambda_2}',\ \mathbf{P_3} = \mathbf{R}_{\lambda_3} \mathbf{L}_{\lambda_3}'$$

Proof: Since the k-th projection is defined as $\mathbf{P_k} = \mathbf{QI_{kk}Q^{-1}}$ with $\mathbf{I_{kk}}$ the (3×3) matrix with the kk-th element equal to 1 and all the other elements equal to 0, the ij-th element of $\mathbf{P_k}$ can be expressed as $(\mathbf{P_k})_{ij} = \sum_{l=1}^{3} \sum_{m=1}^{3} \mathbf{Q}_{il} (\mathbf{I_{kk}})_{lm} \left(\mathbf{Q}^{-1}\right)_{mj} = \sum_{l=1}^{3} \sum_{m=1}^{3} \mathbf{Q}_{il} \delta_{lk} \delta_{mk} \left(\mathbf{Q}^{-1}\right)_{mj} = \mathbf{Q}_{ik} \left(\mathbf{Q}^{-1}\right)_{kj}$. The k-th projection $\mathbf{P_k}$ is therefore the product of the k-th column of the transformation matrix $\mathbf{Q}$ and the k-th row of the inverse matrix $\mathbf{Q}^{-1}$. Which proves the Lemma. ∎

Lemma 2 For a probability matrix P with spectral decomposition $\mathbf{P_1} + \lambda_2 \mathbf{P_2} + \lambda_3 \mathbf{P_3}$, the projections corresponding with an eigenvalue different from 1 have all row sums equal to zero.

Proof: A left eigenvector $\mathbf{L}_\lambda'$ of the matrix $\mathbf{P}$ corresponding with eigenvalue $\lambda \neq 1$ satisfies $\mathbf{L}_\lambda' \mathbf{P} = \lambda \mathbf{L}_\lambda'$. Therefore $\sum_{j=1}^{3} \left(\mathbf{L}_\lambda' \mathbf{P}\right)_j = \sum_{i=1}^{3} \left(\lambda \mathbf{L}_\lambda'\right)_i$ that is equivalent with: $\sum_{j=1}^{3} \sum_{i=1}^{3} \left(\mathbf{L}_\lambda'\right)_i (\mathbf{P})_{ij} = \lambda \sum_{i=1}^{3} \left(\mathbf{L}_\lambda'\right)_i$
$\Leftrightarrow\ \ \sum_{i=1}^{3} \left(\mathbf{L}_\lambda'\right)_i \sum_{j=1}^{3} (\mathbf{P})_{ij} = \lambda \sum_{i=1}^{3} \left(\mathbf{L}_\lambda'\right)_i$
$\Leftrightarrow\ \ \sum_{i=1}^{3} \left(\mathbf{L}_\lambda'\right)_i = \lambda \sum_{i=1}^{3} \left(\mathbf{L}_\lambda'\right)_i$ since the row sums of the probability matrix $\mathbf{P}$ are equal to 1.
Let us conclude that for an eigenvalue $\lambda \neq 1$, the coordinates of a left eigenvector $\mathbf{L}_\lambda'$ sum up to 0: $\sum_{i=1}^{3} \left(\mathbf{L}_\lambda'\right)_i = 0$.
Furthermore according to Lemma 1 the projection $\mathbf{P_k}$ satisfies $\mathbf{P_k} = \mathbf{R}_{\lambda_k} \mathbf{L}_{\lambda_k}'$. For $\lambda_k \neq 1$ this results in:

$$\sum_{j=1}^{3} (\mathbf{P_k})_{ij} = \sum_{j=1}^{3} (\mathbf{R}_{\lambda_k})_i \left(\mathbf{L}_{\lambda_k}'\right)_j = (\mathbf{R}_{\lambda_k})_i \sum_{j=1}^{3} \left(\mathbf{L}_{\lambda_k}'\right)_j = 0$$

Consequently all row sums of the projection $\mathbf{P_k}$ are equal to 0. Which proves the Lemma. ∎

Corollary 1 For a probability matrix $\mathbf{P}$ with spectral decomposition $\mathbf{P_1} + \lambda_2 \mathbf{P_2} + \lambda_3 \mathbf{P_3}$ and nonnegative eigenvalues $\lambda_2 \neq 1, \lambda_3 \neq 1$, the values of the row sums of the matrix $\mathbf{P}(c_2, c_3) = \mathbf{P_1} + c_2 \mathbf{P_2} + c_3 \mathbf{P_3}$ do not depend on the values of $c_2, c_3 \in \mathbb{R}$.

Proof: Since $\lambda_2 \neq 1, \lambda_3 \neq 1$, Lemma 2 results in $\sum_{j=1}^{3} (\mathbf{P_2})_{ij} = \sum_{j=1}^{3} (\mathbf{P_3})_{ij} = 0$. Which proves the corollary. ∎

Corollary 2 For $x \in \mathbb{R}$ and $\mathbf{P}$ probability matrix with nonnegative eigenvalues and spectral decomposition $\mathbf{P_1} + \lambda_2 \mathbf{P_2} + \lambda_3 \mathbf{P_3}$, the matrices $\mathbf{P}(\lambda_2{}^x, \lambda_3{}^x) = \mathbf{P_1} + \lambda_2{}^x \mathbf{P_2} + \lambda_3{}^x \mathbf{P_3}$, $\mathbf{P}(-\lambda_2{}^x, \lambda_3{}^x) = \mathbf{P_1} - \lambda_2{}^x \mathbf{P_2} + \lambda_3{}^x \mathbf{P_3}$, $\mathbf{P}(\lambda_2{}^x, -\lambda_3{}^x) = \mathbf{P_1} + \lambda_2{}^x \mathbf{P_2} - \lambda_3{}^x \mathbf{P_3}$, $\mathbf{P}(-\lambda_2{}^x, -\lambda_3{}^x) = \mathbf{P_1} - \lambda_2{}^x \mathbf{P_2} - \lambda_3{}^x \mathbf{P_3}$ are all row-normalized matrices.

Proof: According to Corollary 1 the row sums of $\mathbf{P} = \mathbf{P_1} + \lambda_2 \mathbf{P_2} + \lambda_3 \mathbf{P_3} = \mathbf{P}(\lambda_2, \lambda_3)$ and $\mathbf{P}(\lambda_2{}^x, \lambda_3{}^x)$, $\mathbf{P}(-\lambda_2{}^x, \lambda_3{}^x)$, $\mathbf{P}(\lambda_2{}^x, -\lambda_3{}^x)$, $\mathbf{P}(-\lambda_2{}^x, -\lambda_3{}^x)$ are equal. Since $\mathbf{P}$ is a probability matrix, all these matrices have row sums equal to 1 and are therefore row-normalized. ∎

For a diagonalizable probability matrix $\mathbf{P}$ with eigenvalues $\lambda_1 = 1 \geq \lambda_2 \geq \lambda_3 \geq 0$, sufficient conditions for the existence of an m-th root of $\mathbf{P}$ are presented in Theorem

2 in the situation that $\lambda_1 = 1 > \lambda_2 > \lambda_3 > 0$. For the special cases with $\lambda_2 = 1$, $\lambda_2 = \lambda_3$ or $\lambda_3 = 0$, it is proved in Theorem 1 that any probability matrix $\mathbf{P}$ (without any condition) has a probability m-th root.

Theorem 1 For a probability matrix $\mathbf{P}$ with spectral decomposition $\mathbf{P} = \mathbf{P_1} + \lambda_2\mathbf{P_2} + \lambda_3\mathbf{P_3}$ and nonnegative eigenvalues that satisfy either $\lambda_2 = 1$, $\lambda_2 = \lambda_3$ or $\lambda_3 = 0$, a probability m-th root of $\mathbf{P}$ exists for all $m \in \mathbb{N}, m \geq 2$.

Proof: Under the condition that either $\lambda_2 = 1$, $\lambda_2 = \lambda_3$ or $\lambda_3 = 0$, the probability matrix $\mathbf{P} = \mathbf{P_1} + \lambda_2\mathbf{P_2} + \lambda_3\mathbf{P_3}$ is of the form $\mathbf{P} = \hat{\mathbf{P}}_1 + \hat{\lambda}\hat{\mathbf{P}}$ with $\hat{\lambda} \geq 0$ since

- $\lambda_2 = 1$ results in $\mathbf{P} = (\mathbf{P_1} + \mathbf{P_2}) + \lambda_3\mathbf{P_3}$;
- $\lambda_2 = \lambda_3$ results in $\mathbf{P} = \mathbf{P_1} + \lambda_2(\mathbf{P_2} + \mathbf{P_3})$ and
- $\lambda_3 = 0$ results in $\mathbf{P} = \mathbf{P_1} + \lambda_2\mathbf{P_2}$.

Furthermore $\left(\hat{\mathbf{P}}_\mathbf{1} + \sqrt[m]{\hat{\lambda}}\,\hat{\mathbf{P}}\right)^m = \hat{\mathbf{P}}_\mathbf{1} + \hat{\lambda}\hat{\mathbf{P}} = \mathbf{P}$ and therefore $\sqrt[m]{\mathbf{P}} = \hat{\mathbf{P}}_\mathbf{1} + \sqrt[m]{\hat{\lambda}}\,\hat{\mathbf{P}}$ is an m-th root of $\mathbf{P}$.
Depending on the situation, the matrix $\hat{\mathbf{P}}$ is equal to either $\mathbf{P_3}$, or $\mathbf{P_2}+\mathbf{P_3}$, or $\mathbf{P_2}$. According to Lemma 2, the matrix $\hat{\mathbf{P}}$ has all row sums equal to zero. The row sums of $\hat{\mathbf{P}}_1$ are equal to the row sums of the probability matrix $\mathbf{P}$. Consequently all row sums of $\sqrt[m]{\mathbf{P}}$ are equal to 1 and therefore $\sqrt[m]{\mathbf{P}}$ is a row-normalized m-th root of $\mathbf{P}$.
Since for each pair $(i, j) \in \{1, 2, 3\} \times \{1, 2, 3\}$ the function $f_{ij}(x) = \left(\hat{\mathbf{P}}_\mathbf{1}\right)_{ij} + \left(\hat{\lambda}\right)^x \left(\hat{\mathbf{P}}\right)_{ij}$ is monotonous (increasing or decreasing) with $f_{ij}(0) = \left(\hat{\mathbf{P}}_\mathbf{1}\right)_{ij} + \left(\hat{\mathbf{P}}\right)_{ij} = I_{ij} \in [0, 1]$ and $f_{ij}(1) = \left(\hat{\mathbf{P}}_\mathbf{1}\right)_{ij} + \hat{\lambda}\left(\hat{\mathbf{P}}\right)_{ij} = \mathbf{P}_{ij} \in [0, 1]$, holds that $\left(\sqrt[m]{\mathbf{P}}\right)_{ij} \in [0, 1]$ for all $m \in \mathbb{N}, m \geq 2$. For these reasons $\hat{\mathbf{P}}_\mathbf{1} + \sqrt[m]{\hat{\lambda}}\,\hat{\mathbf{P}}$ is a probability matrix and an m-th root of $\mathbf{P}$. Which proves the Theorem. ■

From Theorem 1 it is known that in the situation where the diagonalizable transition matrix $\mathbf{P}$ has eigenvalues $\lambda_1 = 1 \geq \lambda_2 \geq \lambda_3 \geq 0$ satisfying $\lambda_2 = 1$, $\lambda_2 = \lambda_3$ or $\lambda_3 = 0$, the Markov chain is embeddable. In the rest of this section, sufficient embedding conditions are examined for a diagonalizable transition matrix $\mathbf{P}$ with eigenvalues that satisfy $\lambda_1 = 1 > \lambda_2 > \lambda_3 > 0$.

Lemma 3 For $f : [0, 1] \to \mathbb{R}, x \to a_1 + a_2{\lambda_2}^x + a_3{\lambda_3}^x$ with $f(0), f(1) \in [0, 1]$, $1 > \lambda_2 > \lambda_3 > 0$ and $a_1, a_2, a_3 \in \mathbb{R}$ holds:
The range of f is a subset of $[0, 1]$ iff

1) $a_2a_3 \geq 0$ or
2) $a_2a_3 < 0$ and $x^* \notin]0, 1[$ or
3) $a_2a_3 < 0$ and $x^* \in]0, 1[$ with $f(x^*) \in [0, 1]$

with $x^* = \log_{\frac{\lambda_2}{\lambda_3}}\left(-\frac{a_3 \ln\lambda_3}{a_2 \ln\lambda_2}\right)$

The proof of Lemma 3 is presented in Appendix A.

Remark
Since $f''(x^*) = {\lambda_3}^{x^*} a_3 \ln\lambda_3 (\ln\lambda_3 - \ln\lambda_2)$ and $1 > \lambda_2 > \lambda_3 > 0$, the sign of $f''(x^*)$ corresponds with the sign of a_3. Therefore in case $a_3 > 0$, $f(x^*)$ results in a minimum value for f and the sufficient condition $f(x^*) \in [0, 1]$ can be relaxed to $f(x^*) \geq 0$. In case $a_3 < 0$, $f(x^*)$ results in a maximum value for f and the sufficient condition $f(x^*) \in [0, 1]$ can be replaced by $f(x^*) \leq 1$.

The following Theorem provides in sufficient embedding conditions for a probability matrix $\mathbf{P}$ with eigenvalues $1 = \lambda_1 > \lambda_2 > \lambda_3 > 0$.

Theorem 2 Let $\mathbf{P}$ be a probability matrix with spectral decomposition $\mathbf{P} = \mathbf{P_1} + \lambda_2\mathbf{P_2} + \lambda_3\mathbf{P_3}$ and $1 > \lambda_2 > \lambda_3 > 0$. In case for each pair $(i, j) \in \{1, 2, 3\} \times \{1, 2, 3\}$ and $x^*_{ij} = \log_{\frac{\lambda_2}{\lambda_3}}\left(-\frac{(\mathbf{P_3})_{ij} \ln\lambda_3}{(\mathbf{P_2})_{ij} \ln\lambda_2}\right)$ one of the following conditions holds

1) $(\mathbf{P_2})_{ij}(\mathbf{P_3})_{ij} \geq 0$ or
2) $(\mathbf{P_2})_{ij}(\mathbf{P_3})_{ij} < 0$ and $x^*_{ij} \notin]0, 1[$ or
3) $(\mathbf{P_2})_{ij}(\mathbf{P_3})_{ij} < 0$ and $x^*_{ij} \in]0, 1[$ with $(\mathbf{P_1})_{ij} + \lambda_2^{x^*_{ij}}(\mathbf{P_2})_{ij} + \lambda_3^{x^*_{ij}}(\mathbf{P_3})_{ij} \in [0, 1]$

then $\sqrt[m]{\mathbf{P}} = \mathbf{P_1} + \sqrt[m]{\lambda_2}\,\mathbf{P_2} + \sqrt[m]{\lambda_3}\,\mathbf{P_3}$, with $m \in \mathbb{N}, m \geq 2$, is a probability m-th root of $\mathbf{P}$.

Proof: The matrix $\sqrt[m]{\mathbf{P}} = \mathbf{P_1} + \sqrt[m]{\lambda_2}\,\mathbf{P_2} + \sqrt[m]{\lambda_3}\,\mathbf{P_3}$ is an m-th root of $\mathbf{P}$ since $\left(\sqrt[m]{\mathbf{P}}\right)^m = \mathbf{P}$. According to Corollary 2 the m-th root $\sqrt[m]{\mathbf{P}} = \mathbf{P_1} + \sqrt[m]{\lambda_2}\mathbf{P_2} + \sqrt[m]{\lambda_3}\mathbf{P_3}$ is a row-normalized matrix.
Furthermore for each pair $(i, j) \in \{1, 2, 3\} \times \{1, 2, 3\}$ the function $f_{ij}(x) = (\mathbf{P_1})_{ij} + {\lambda_2}^x(\mathbf{P_2})_{ij} + {\lambda_3}^x(\mathbf{P_3})_{ij}$ can be introduced, satisfying $f_{ij}\left(\frac{1}{m}\right) = \left(\sqrt[m]{P}\right)_{ij}$. Since $f_{ij}(0) = (\mathbf{P_1})_{ij} + (\mathbf{P_2})_{ij} + (\mathbf{P_3})_{ij} = \mathbf{I}_{ij}$, holds that $f_{ij}(0) \in [0, 1]$. Besides $f_{ij}(1) = (\mathbf{P_1})_{ij} + \lambda_2(\mathbf{P_2})_{ij} + \lambda_3(\mathbf{P_3})_{ij} = \mathbf{P}_{ij}$ and therefore satisfies $f_{ij}(1) \in [0, 1]$. Consequently for $a_1 = (\mathbf{P_1})_{ij}$, $a_2 = (\mathbf{P_2})_{ij}$ and $a_3 = (\mathbf{P_3})_{ij}$, the conditions of Lemma 3 are fullfilled. All the elements of the matrix $\sqrt[m]{P} = \mathbf{P_1} + \sqrt[m]{\lambda_2}\,\mathbf{P_2} + \sqrt[m]{\lambda_3}\,\mathbf{P_3}$ have therefore a value between 0 and 1. Consequently the matrix $\sqrt[m]{P}$ is a probability m-th root of $\mathbf{P}$. ■

Remarks
(1) Some of the conditions in Theorem 2 can be reformulated. For example the condition $x^*_{ij} = \log_{\frac{\lambda_2}{\lambda_3}}\left(-\frac{(\mathbf{P_3})_{ij} \ln\lambda_3}{(\mathbf{P_2})_{ij} \ln\lambda_2}\right) \in]0, 1[$ for $1 > \lambda_2 > \lambda_3 > 0$ can also be expressed as $\frac{\ln\lambda_2}{\ln\lambda_3} < -\frac{(\mathbf{P_3})_{ij}}{(\mathbf{P_2})_{ij}} < \frac{\lambda_2 \ln\lambda_2}{\lambda_3 \ln\lambda_3}$.
(2) For $\mathbf{P} = \mathbf{P_1} + \lambda_2\mathbf{P_2} + \lambda_3\mathbf{P_3}$ and for m even number, besides $\mathbf{P_1} + \sqrt[m]{\lambda_2}\mathbf{P_2} + \sqrt[m]{\lambda_3}\mathbf{P_3}$, the matrices $\mathbf{P_1} - \sqrt[m]{\lambda_2}\mathbf{P_2} + \sqrt[m]{\lambda_3}\mathbf{P_3}$, $\mathbf{P_1} + \sqrt[m]{\lambda_2}\mathbf{P_2} - \sqrt[m]{\lambda_3}\mathbf{P_3}$ and $\mathbf{P_1} - \sqrt[m]{\lambda_2}\mathbf{P_2} - \sqrt[m]{\lambda_3}\mathbf{P_3}$ are all row-normalized m-th roots of $\mathbf{P}$.

III. Illustrations

In this section for a transition matrix $\mathbf{P}$, it is illustrated how the proven sufficient embedding conditions can be helpful in getting insights on the existence of probability m-th roots. The use of Theorem 2 let us conclude the following: For the first example $\sqrt[m]{\mathbf{P}}$ is not a probability matrix (and this for all $m \in \mathbb{N}, m \geq 2$); for the second example for all $m \in \mathbb{N}, m \geq 2$ holds that $\sqrt[m]{\mathbf{P}}$ is a probability root of $\mathbf{P}$; for the last example $\sqrt[m]{\mathbf{P}}$ is a probability root for all natural numbers m between identified lower and upper bounds.

Example 1 The probability matrix

$$\mathbf{P} = \begin{pmatrix} 0.8 & 0.2 & 0 \\ 0.5 & 0.3 & 0.2 \\ 0.1 & 0.4 & 0.5 \end{pmatrix}$$

ISBN: 978-988-19253-5-0
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

has eigenvalues $\lambda_1 = 1, \lambda_2 = 0.3+\sqrt{0.08}, \lambda_3 = 0.3-\sqrt{0.08}$ and projections (the elements of the projections are mentioned up to three decimals)

$$\mathbf{P_1} = \begin{pmatrix} 0.658 & 0.244 & 0.098 \\ 0.658 & 0.244 & 0.098 \\ 0.658 & 0.244 & 0.098 \end{pmatrix}$$

$$\mathbf{P_2} = \begin{pmatrix} 0.240 & -0.070 & -0.170 \\ -0.260 & 0.076 & 0.184 \\ -0.967 & 0.283 & 0.684 \end{pmatrix}$$

$$\mathbf{P_3} = \begin{pmatrix} 0.102 & -0.174 & 0.072 \\ -0.398 & 0.680 & -0.282 \\ 0.309 & -0.527 & 0.218 \end{pmatrix}$$

For $i = 1$ and $j = 3$ the function $f_{13}(x) = (\mathbf{P_1})_{13} + {\lambda_2}^x(\mathbf{P_2})_{13} + {\lambda_3}^x(\mathbf{P_3})_{13}$ has one critical value namely $x^*_{13} = \log_{\frac{\lambda_2}{\lambda_3}}\left(-\frac{(\mathbf{P_3})_{13}\ln\lambda_3}{(\mathbf{P_2})_{13}\ln\lambda_2}\right) \approx 0.33 < 1$, resulting in a minimum value $f_{13}(x^*_{13}) \approx -0.025$. Moreover $f_{13}(0) = \mathbf{I}_{13} = 0$ and $f_{13}(1) = \mathbf{P}_{13} = 0$. Therefore for each value $x \in]0,1[$ holds that $f_{13}(x) < 0$. In particular for each value $x = \frac{1}{m}$, with $m \in \mathbb{N}, m \geq 2$, holds that $\left(\sqrt[m]{P}\right)_{13} = f_{13}\left(\frac{1}{m}\right) < 0$ so that the matrix $\sqrt[m]{P}$ is not a probability matrix.

Example 2 The probability matrix

$$\mathbf{P} = \begin{pmatrix} 0.7 & 0.2 & 0.1 \\ 0.1 & 0.8 & 0.1 \\ 0.1 & 0.3 & 0.6 \end{pmatrix}$$

has eigenvalues $\lambda_1 = 1, \lambda_2 = 0.6, \lambda_3 = 0.5$ and projections

$$\mathbf{P_1} = \begin{pmatrix} 0.25 & 0.55 & 0.2 \\ 0.25 & 0.55 & 0.2 \\ 0.25 & 0.55 & 0.2 \end{pmatrix}$$

$$\mathbf{P_2} = \begin{pmatrix} 0.75 & -0.75 & 0 \\ -0.25 & 0.25 & 0 \\ -0.25 & 0.25 & 0 \end{pmatrix}$$

$$\mathbf{P_3} = \begin{pmatrix} 0 & 0.2 & -0.2 \\ 0 & 0.2 & -0.2 \\ 0 & -0.8 & 0.8 \end{pmatrix}$$

Since the elements of the first column of $\mathbf{P_3}$ and the third column of $\mathbf{P_2}$ are all equal to zero, condition 1 holds for all pairs $(i,j) \in \{1,2,3\} \times \{1,3\}$. Condition 1 is also satisfied for $i = 2$ and $j = 2$: $(\mathbf{P_2})_{22}(\mathbf{P_3})_{22} = 0.05 > 0$. For $i \in \{1,3\}$ and $j = 2$ the values for x^*_{ij} are $x^*_{12} = \log_{\frac{\lambda_2}{\lambda_3}}\left(-\frac{(\mathbf{P_3})_{12}\ln\lambda_3}{(\mathbf{P_2})_{12}\ln\lambda_2}\right) \approx -5.58 < 0$ and $x^*_{32} = \log_{\frac{\lambda_2}{\lambda_3}}\left(-\frac{(\mathbf{P_3})_{32}\ln\lambda_3}{(\mathbf{P_2})_{32}\ln\lambda_2}\right) \approx 8.05 > 1$, and therefore condition 2 is satisfied. Theorem 2 let us conclude that $\sqrt[m]{\mathbf{P}}$ is a probability m-th root of $\mathbf{P}$, and this conclusion holds for all $m \in \mathbb{N}, m \geq 2$.

Example 3 The probability matrix

$$\mathbf{P} = \begin{pmatrix} 0.5 & 0.3 & 0.2 \\ 0.3 & 0.3 & 0.4 \\ 0.1 & 0.2 & 0.7 \end{pmatrix}$$

has eigenvalues $\lambda_1 = 1, \lambda_2 = 0.25 + \sqrt{0.0425}$, $\lambda_3 = 0.25 - \sqrt{0.0425}$ and projections (the elements of the projections $\mathbf{P_2}$ and $\mathbf{P_3}$ are mentioned up to three decimals)

$$\mathbf{P_1} = \begin{pmatrix} 0.25 & 0.25 & 0.5 \\ 0.25 & 0.25 & 0.5 \\ 0.25 & 0.25 & 0.5 \end{pmatrix}$$

$$\mathbf{P_2} = \begin{pmatrix} 0.526 & 0.148 & -0.674 \\ 0.148 & 0.041 & -0.189 \\ -0.337 & -0.095 & 0.432 \end{pmatrix}$$

$$\mathbf{P_3} = \begin{pmatrix} 0.223 & -0.398 & 0.175 \\ -0.398 & 0.708 & -0.310 \\ 0.087 & -0.155 & 0.068 \end{pmatrix}$$

The projections $\mathbf{P_2}$ and $\mathbf{P_3}$ have a difference in sign for the corresponding elements $(\mathbf{P_2})_{ij}$ and $(\mathbf{P_3})_{ij}$ in case of $(i,j) \in \{(1,2),(2,1),(1,3),(3,1)\}$. Since the values of x^*_{12} and x^*_{21} are greater than 1, Theorem 2 assures that $\left(\sqrt[m]{P}\right)_{12}, \left(\sqrt[m]{P}\right)_{21} \in [0,1]$, and this for all $m \in \mathbb{N}, m \geq 2$. For $i = 1, j = 3$ and $i = 3, j = 1$, $x^*_{13} = x^*_{31} = 0.012792 \in [0,1]$. The function $f_{13}(x) = (\mathbf{P_1})_{13} + {\lambda_2}^x(\mathbf{P_2})_{13} + {\lambda_3}^x(\mathbf{P_3})_{13}$ has one critical value namely x^*_{13} and is therefore monotone on the interval $[x^*_{13}, 1]$. Since $\frac{1}{38}, \frac{1}{2} \in [x^*_{13}, 1]$ with $f_{13}\left(\frac{1}{2}\right)$ and $f_{13}\left(\frac{1}{38}\right)$ both elements of $[0,1]$, holds that $\left(\sqrt[m]{\mathbf{P}}\right)_{13} \in [0,1]$ for all m values satisfying $2 \leq m \leq 38$. The same conclusions are valid for $i = 3, j = 1$. Besides $f_{13}\left(\frac{1}{39}\right) < 0$ and therefore $\sqrt[m]{P}$ is not a probability matrix for $m \geq 39$. Consequently $\sqrt[m]{P}$ is a probability root of $\mathbf{P}$ for all m values in between 2 and 38.

IV. Conclusion

The present paper investigates probability roots for a transition matrix $\mathbf{P}$ that is diagonalizable and that has all eigenvalues nonnegative. In this way sufficient embedding conditions are obtained for a Markov chain with such a transition matrix $\mathbf{P}$. For an embeddable Markov chain a probability m-th root provides information on the transition probabilities regarding time intervals with length $\frac{1}{m}$.

According to Theorem 1 for a diagonalizable transition matrix with either $\lambda_2 = 1$, $\lambda_2 = \lambda_3$ or $\lambda_3 = 0$, for all $m \in \mathbb{N}, m \geq 2$, a probability m-th root of $\mathbf{P}$ exists. For all other cases of a diagonalizable transition matrix $\mathbf{P}$ with nonnegative eigenvalues, Theorem 2 provides sufficient conditions for the existence of a probability m-th root of $\mathbf{P}$.

Further research should investigate sufficient conditions under which probability roots do exist either in case the transition matrix is diagonalizable with not all eigenvalues nonnegative or in case the transition matrix is not diagonalizable.

Appendix A
Proof of Lemma 3

In case $a_2a_3 \geq 0$, $f(x) = a_1 + a_2{\lambda_2}^x + a_3{\lambda_3}^x$ is a monotonous function.

In case $a_2a_3 < 0$, $f(x)$ has one critical value:

$$f'(x) = a_2{\lambda_2}^x\ln\lambda_2 + a_3{\lambda_3}^x\ln\lambda_3 = 0$$

$$\Leftrightarrow \quad x = x^* = \log_{\frac{\lambda_2}{\lambda_3}}\left(-\frac{a_3\ln\lambda_3}{a_2\ln\lambda_2}\right)$$

ISBN: 978-988-19253-5-0
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

In case $a_2a_3 < 0$ and $x^* \notin]0,1[$, the function $f(x)$ has no critical value in $]0,1[$ and consequently $f(x)$ is monotonous on $[0,1]$.

Therefore in case $a_2a_3 \geq 0$ and in case $a_2a_3 < 0$ with $x^* \notin]0,1[$, the range of $f(x)$ is a subset of $[0,1]$ since $f(x)$ is a monotonous function on $[0,1]$ with $f(0), f(1) \in [0,1]$.

In case $a_2a_3 < 0$ and $x^* \in]0,1[$ the function $f(x)$ has x^* as critical value in $[0,1]$. In case the value $f(x^*)$ belongs to the interval $[0,1]$, the range of $f(x)$ is a subset of $[0,1]$. Which proves the Lemma.

Acknowledgment

The author thanks the reviewers for their remarks and valuable suggestions.

References

[1] D. J. Bartholomew, A. F. Forbes and S. I. McClean, *Statistical techniques for manpower planning*, Wiley, Chichester, 1991.
[2] N. Bauerle and U. Rieder, *Markov decision processes with applications to finance*, Springer, 2011.
[3] H. Berthiaux and V. Mizonov, Applications of Markov chains in particulate process engineering: a review, *The Canadian Journal of Chemical Engineering*, vol. 82, pp. 1143–1168, 2004.
[4] G. Elfving, Zur theorie der Markoffschen ketten, *Acta Social Sci. Fennicae n. Ser. A 2*, vol. 8, pp. 1–17, 1937.
[5] A. C. Georgiou and N. Tsantas, Modelling recruitment training in mathematical human resource planning, *Applied Stochastic Models in Business and Industry*, vol. 18, pp. 53–74, 2002.
[6] M. A. Guerry, On the embedding problem for discrete-time Markov chains, *J. of Appl. Probab.*, vol. 50, no. 4, pp. 918–930, 2013.
[7] M. A. Guerry, Some results on the embeddable problem for discrete-time Markov models in manpower planning, *Commun. in Stat. - Theory and Methods,* to appear, 2014.
[8] Q. -M. He and E. Gunn, A note on the stochastic roots of stochastic matrices, *J. of Syst. Sci. and Syst. Eng.*, vol. 12, no. 2, pp. 210–223, 2003.
[9] N. J. Higham and L. Lin, On pth roots of stochastic matrices, *Linear Algebra and its Appl.*, vol. 435, pp. 448–463, 2011.
[10] K. Nilakantan and B. G. Raghavendra, Control aspects in proportionality Markov manpower systems, *Applied Mathematical Modelling* vol. 29, no. 1, pp. 85–116, 2005.
[11] E. Seneta, *Non-negative Matrices. An Introduction to Theory and Applications*, London: George Allen and Unwin, 1973.
[12] B. Singer and S. Spilerman, Social mobility models for heterogeneous populations, in: H.L. Costner (Ed.), *Soc. Methodol.* vol. 5, pp. 356–401, 1974.
[13] A. Tamir, *Applications of Markov Chains in Chemical Engineering*, Elsevier, 1998.