Linear Algebra and its Applications 473 (2015) 114–125

Contents lists available at ScienceDirect

Linear Algebra and its Applications

www.elsevier.com/locate/laa

Stationary probability vectors of higher-order Markov chains

Chi-Kwong Li [a,*], Shixiao Zhang [b]

[a] *Department of Mathematics, College of William and Mary, Williamsburg, VA 23187, USA*
[b] *Department of Mathematics, University of Hong Kong, Hong Kong*

A R T I C L E   I N F O

Article history:
Received 12 March 2013
Accepted 26 March 2014
Available online 13 April 2014
Submitted by L.-H. Lim

MSC:
15A69
15A18

Keywords:
Transition probability tensor
Higher-order Markov chains

A B S T R A C T

We consider the higher-order Markov chain, and characterize the second order Markov chains admitting every probability distribution vector as a stationary vector. The result is used to construct Markov chains of higher-order with the same property. We also study conditions under which the set of stationary vectors of the Markov chain has a certain affine dimension.

1. Introduction

A discrete-time Markov chain is a stochastic process with a sequence of random variables

$$\{X_t,\ t = 0, 1, 2 \ldots\},$$

* Corresponding author.
E-mail addresses: ckli@math.wm.edu (C.-K. Li), praetere@gmail.com (S. Zhang).

http://dx.doi.org/10.1016/j.laa.2014.03.043

which takes on values in a discrete finite state space

$$\langle n \rangle = \{1, \ldots, n\}$$

for a positive integer n, such that time independent probability

$$\begin{aligned} p_{ij} &= \Pr(X_{t+1} = i \mid X_t = j, X_{t-1} = i_{t-1}, X_{t-2} = i_{t-2}, \ldots, X_1 = i_1, X_0 = i_0) \\ &= \Pr(X_{t+1} = i \mid X_t = j) \end{aligned}$$

holds for all $i, j, i_0, \ldots, i_{t-1}$. The nonnegative matrix $P = (p_{ij})_{1 \leqslant i,j \leqslant n}$ is the transition matrix of the Markov process being column stochastic, i.e., $\sum_{i=1}^{n} p_{ij} = 1$ for $j = 1, \ldots, n$. Denote by

$$\Omega_n = \left\{ \mathbf{x} = (x_1, \ldots, x_n)^t \colon\ x_1, \ldots, x_n \geqslant 0,\ \sum_{i=1}^{n} x_i = 1 \right\} \tag{1}$$

the simplex of probability vectors in $\mathbf{R}^n$. A nonnegative vector $\mathbf{x} \in \Omega_n$ is a stationary probability vector (also known as the distribution) of a finite Markov chain if $P\mathbf{x} = \mathbf{x}$. By the Perron–Frobenius theory (e.g., see [3,13]) every discrete-time Markov chain has a stationary probability vector, and the vector is unique if the transition matrix is primitive, i.e., there is a positive integer r such that all entries of P^r are positive. The uniqueness condition is useful when one uses numerical schemes to determine the stationary vectors. With the uniqueness condition, any convergent scheme would lead to the unique stationary vector; e.g., see [5].

More generally, one may consider an mth order Markov chain such that

$$\begin{aligned} p_{i,i_1,\ldots,i_m} &= \Pr(X_{t+1} = i \mid X_t = i_1, X_{t-1} = i_2, \ldots, X_1 = i_t, X_0 = i_{t+1}) \\ &= \Pr(X_{t+1} = i \mid X_t = i_1, \ldots, X_{t-m+1} = i_m), \end{aligned}$$

where $i, i_1, \ldots, i_m \in \langle n \rangle$; see [1,2]. In other words, the current state of the process depends on m past states. Observe that

$$\sum_{i=1}^{n} p_{i,i_1,\ldots,i_m} = 1, \quad 1 \leqslant i_1, \ldots, i_m \leqslant n.$$

When $m = 1$, it is just the standard Markov chain. There are many situations that one would use the Markov chain models. We refer the readers to the papers [1,2,8,11,12] and the references therein. Note that $P = (p_{i,i_1,\ldots,i_m})$ is an $(m+1)$-fold tensor of $\mathbf{R}^n$ governing the transition of states in the mth order Markov chain according to the following rule

$$x_i(t+1) = \sum_{1 \leqslant i_1,\ldots,i_m \leqslant n} p_{i,i_1,\ldots,i_m} x_{i_1}(t) \cdots x_{i_m}(t), \quad i = 1, \ldots, n.$$