Applied Mathematics and Computation 241 (2014) 183–192

Contents lists available at ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

Convergence of a second order Markov chain

Shenglong Hu [a,b,*,1], Liqun Qi [b,2]

[a] Department of Mathematics, School of Science, Tianjin University, 300072 Tianjin, China
[b] Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

ARTICLE INFO

Keywords:
Nonnegative tensor
Transition probability tensor
Second order Markov chain

ABSTRACT

In this paper, we consider convergence properties of a second order Markov chain. Similar to a column stochastic matrix being associated to a Markov chain, a *transition probability tensor* P of order 3 and dimension n is associated to a second order Markov chain with n states. For this P, define F_P as $F_P(x) := Px^2$ on the $n-1$ dimensional standard simplex Δ_n. If 1 is not an eigenvalue of ∇F_P on Δ_n and P is irreducible, then there exists a unique fixed point of F_P on Δ_n. In particular, if every entry of P is greater than $\frac{1}{2n}$, then 1 is not an eigenvalue of ∇F_P on Δ_n. Under the latter condition, we further show that the second order power method for finding the unique fixed point of F_P on Δ_n is globally linearly convergent and the corresponding second order Markov process is globally R-linearly convergent.

1. Introduction

Markov chain serves as a fundamental tool for diverse applications [9,17,18]. As a generalization, higher order Markov chain can be used to describe many phenomena in science and engineering, e.g., bioinformatics, genome, speech/text recognition, please refer to [2,10] and references therein. Compared with the sophisticated development of Markov chain based on stochastic matrices, research on higher order Markov chain based on transition probability tensors, is just on the way [2,10,11,13]. Nevertheless, the recent progress in numerical multilinear algebra, especially in tensor decomposition [6,7] and spectral theory of tensors [12,14], introduces many new tools for this topic.

An mth order n dimensional Markov chain is basically characterized by its associated nonnegative tensor P which is an $(m+1)$th order n dimensional tensor with entries $p_{ij_1\ldots j_m}$ for all $i, j_1, \ldots, j_m = 1, \ldots, n$ satisfying:

$$0 \leqslant p_{ij_1\ldots j_m} = \mathrm{Prob}(X_t = i \mid X_{t-1} = j_1, \ldots, X_{t-m} = j_m) \leqslant 1.$$

Here $\{X_t, t = 0, 1, \ldots\}$ represents the stochastic process that takes on n states $\{1, \ldots, n\}$. Obviously,

$$\sum_{i=1}^{n} p_{ij_1\ldots j_m} = 1, \quad (1)$$

for any $j_1, \ldots, j_m = 1, \ldots, n$.

* Corresponding author at: Department of Mathematics, School of Science, Tianjin University, 300072 Tianjin, China.
E-mail addresses: tim.hu@connect.polyu.hk (S. Hu), maqilq@polyu.edu.hk (L. Qi).

[1] This author's work was partially supported by National Natural Science Foundation of China (Grant No. 11101303).
[2] This author's work was supported by the Hong Kong Research Grant Council. (Grant No. PolyU 502510, 502111, 501212 and 501913).

http://dx.doi.org/10.1016/j.amc.2014.05.011
0096-3003/

An $(m+1)$th order n dimensional nonnegative tensor P that satisfies (1) is called a *transition probability tensor*. In this paper, we consider convergence properties of a second order Markov chain which is associated to a third order n dimensional transition probability tensor P. Given such a tensor P, we specially consider the sequence of state distribution vectors generated by the second order Markov process: with two initial state distribution vectors $x^{(0)}, x^{(1)} \in \Delta_n := \{x \in \Re^n_+ \mid \sum_{i=1}^n x_i = 1\}$, the sequence is generated as

$$x^{(s)} := Px^{(s-1)}x^{(s-2)}, \quad \forall s = 2,3,\ldots, \tag{2}$$

where $Px^{(s-1)}x^{(s-2)}$ is an n-vector whose ith element is

$$\sum_{j,k=1}^{n} p_{ijk} x_j^{(s-1)} x_k^{(s-2)},$$

for all $i \in \{1,\ldots,n\}$. If the sequence $\{x^{(k)}\}$ converges to x^*, then we call x^* the stationary probability distribution of the second order Markov chain. Obviously, in this situation,

$$x^* = Px^*x^* =: P(x^*)^2 \tag{3}$$

which is closely related to Z-eigenvalues of tensors introduced in [14].

For the convenience of the subsequent analysis, define a nonlinear map $F_P : \Re^n \to \Re^n$ associated to P as:

$$(F_P(x))_i = \sum_{j,k=1}^{n} p_{ijk} x_j x_k, \tag{4}$$

for all $i \in \{1,\ldots,n\}$ and $x \in \Re^n$, and denote by $\nabla F_P(x)$ the Jacobian matrix of the map F_P at x. Obviously, $\sum_{i=1}^n (F_P(x))_i = 1$ for all $x \in \Delta_n$. Essentially, stationary probability distribution of the second order Markov chain associated to P in (3), whenever it exists, is a fixed point of F_P on Δ_n.

Very recently, under mild conditions, some results like the uniqueness of x^* in (3) and the linear convergence of the power method for finding such a unique x^* were established in [10,11]. Unlike its counterpart of $m = 1$ [18], the convergence of the original Markov process (2) could not be deduced directly from the convergence of the power method for finding x^* in (3). Hence, in this paper, we mainly consider this problem and give an answer.

We give a uniqueness property of the fixed point of F_P on Δ_n by using fixed point index theory in Section 2. We establish globally linear convergence of the second order power method for finding the unique fixed point in Section 3. Globally R-linear convergence of the second order Markov process (2) is proved in Section 4. Some intuitive numerical examples are given in the last section.

2. The uniqueness property

In this section, we discuss the uniqueness of the fixed point of F_P on Δ_n for a given transition probability tensor P, which is parallel to the strong Perron–Frobenius theorem for primitive stochastic matrices [18]. In the first subsection, we give a general result, and we discuss more on third order transition probability tensors in the second subsection.

2.1. General case

We give the general result for mth order n dimensional transition probability tensors. The following concept is useful: an mth order n dimensional nonnegative tensor P is called *reducible* if there exists a nonempty proper index subset $I \subset \{1,\ldots,n\}$ such that

$$P_{ij_1\cdots j_{m-1}} = 0, \quad \forall i \in I, \quad \forall j_1,\ldots,j_{m-1} \notin I. \tag{5}$$

If P is not reducible, then P is called *irreducible* [1]. Obviously, P is irreducible if it is a positive tensor. The concepts of relative interior and relative boundary of a set are in the usual sense [16].

In the subsequent analysis, P is assumed to be an mth order n dimensional transition probability tensor. In order to accomplish the proof, we first introduce briefly the concept of fixed point index, see [4] for a comprehensive discussion. The theory established in [4] applies to general Banach spaces, while we present it in the finitely dimensional cases to match our problem. Intuitive speaking, the theory of fixed point index is just a generalization of the degree theory: the degree theory discusses the fixed points of a map in an open set; while, the theory of fixed point index discusses the fixed points of a map in a relative open set. As our domain Δ_n here is relative open in $\Re^n$, we need the theory of fixed point index to carry out the proof. Here is the analogue of fixed point index to degree theory.

Theorem 1. *Suppose that $U \in \Re^n$ is a bounded relatively open subset with closure $\overline{U}$ and $f : \overline{U} \to \Re^n$ is continuous. If f has no fixed points on the relative boundary ∂U of U, then there exists an integer function, denoted as $Ind(f,U)$, satisfying the following properties:*