The Asymptotic Equipartition Property for Nonhomogeneous Markov Chains Indexed by Trees

Peng Weicai and Chen Peishu

Department of Mathematics
Chaohu University, Chaohu, Anhui
238000, P.R. China

Abstract. In this paper, the tree T is a general tree. We prove the strong law of large numbers and the asymptotic equipartition property (AEP) for finite nonhomogeneous Markov chains indexed by trees. The results generalize some known results.

Keywords: Nonhomogeneous Markov chain, Tree, Strong law of large numbers.

1 Introduction

By a tree T we mean an infinite, locally finite, connected graph with a distinguished vertex o called the root and without loops or cycles. We only consider trees without leaves. That is, the degree of each vertex (except o) is required to be at least 2. Let σ, τ be vertices of a tree. Write $\tau \le \sigma$ if τ is on the unique path connecting o to σ, and $|\sigma|$ for the number of edges on this path. For any two vertices σ, τ, denote by $\sigma \wedge \tau$ the vertex farthest from o satisfying

$$\sigma \wedge \tau \le \sigma, \sigma \wedge \tau \le \tau$$

The set of all vertices with distance \$n\$ from the root o is called the n-th generation of T, which is denoted by L_n, $L_0 = \{o\}$. We denote by $T^{(n)}$ the subtree of a tree T containing the vertices from level 0 to level n, $T_{(m)}^{(n)}$ the subtree of a tree T containing the vertices from level m to level n.

Let t be a vertex of T, predecessor of the vertex t is another vertex which is nearest from t on the unique path from root o to t. We denote the predecessor of t by 1_t, the predecessor of 1_t by 2_t, the predecessor of N_t by $(N+1)_t$ and 0_t=t, where N=0,1,2,.... We also say that N_t is the N-th predecessor of t.

Definition 1 (Tree-indexed nonhomogeneous Markov chains). Let T be a tree, S be a states space (finite or countable), $\{X_\sigma, \sigma \in T\}$ be a collection of S-valued random variables defined on the probability space $(\Omega, \mathbb{F}, P)$. Let

G. Yang (Ed.): Proceedings of the ICCEAE2012, AISC 181, pp. 673–679.
springerlink.com

$$p=\{p(x), x\in S\} \tag{1}$$

be a distribution on S, and

$$(P_t(y \mid x)), x, y\in S, t\in T \tag{2}$$

be stochastic matrices on S^2. If for any vertices t,

$$\begin{aligned} &P\left(X_t = y \middle| X_t = x \ and\ X_s \ for\ t\wedge s \le 1_t\right) \\ &= P(X_t = y | X_t = x) = P_t(y|x), x, y\in S \end{aligned} \tag{3}$$

and $$P(X_o = x)=p(x) \tag{4}$$

$\{X_t, t\in T\}$ will becalled S-value nonhomogeneous Markov chains indexed by tree with the initial distribution (1) and transition matrix (2). If for all t

$$(P_t(y|x)) = (P(y|x)), x, y\in S \tag{5}$$

$\{X_t, t\in T\}$ will be called S-value homogeneous Markov chains indexed by tree T.

It is easy to see that if $\{X_t, t\in T\}$ is a S-valued Markov chains indexed by a tree defined as above, then

$$P(x^{T^{(n)}}) = P(X^{T^{(n)}} = x^{T^{(n)}}) = p(x_0)\prod_{t\in T^{(n)}_{(N+1)}} P_t(x_t|y_t), \tag{6}$$

$$P(x^{T^{(n-N)}}) = p(x_0)\prod_{t\in T^{(n)}_{(N+1)}} P_t(x_{N_t}|x_{(N=1)_t}) \tag{7}$$

2 Strong Limit Theorems

Lemma 1 (see [4]) Let T be an infinite tree. Let $\{X_t, t\in T\}$ be a T -indexed nonhomogeneous Markov chain with countable states space S defined as before, $\{g_t(x,y), t\in T\}$ be functions defined on S^2. For any given nonnegative integer N,

$$t_{n-N}(\lambda,\omega) = \frac{e^{\lambda\sum_{t\in T^{(n)}_{(N+1)}} g_t(X_{(N+1)t}, X_{N_t})}}{\prod_{t\in T^{(n)}_{(N+1)}} E[e^{\lambda g_t(X_{(N+1)t}, X_{N_t})} | X_{(N+1)_t}]} \tag{8}$$

where λ is a real number. Then $\{t_{n-N}(\lambda,\omega), F_{n-N}, n\ge N+1\}$ is a nonnegative martingale.

Theorem 1. Let T be an infinite tree, $\{X_t, t \in T\}$ be a T -indexed nonhomogeneous Markov chain with countable states space S, $\{g_t(x,y), t \in T\}$ be functions defined on S^2. For any given nonnegative integer N , let

$$H_n(\omega) = \sum_{t \in T^{(n)}_{(N+1)}} g_t(X_{(N+1)_t}, X_{N_t}), \tag{9}$$

$$G_n(\omega) = \sum_{t \in T^{(n)}_{(N+1)}} E[g_t(X_{(N+1)_t}, X_{N_t}) \big| X_{(N+1)_t}] \tag{10}$$

Let $\alpha > 0$, then

$$\lim_{n\to\infty} \frac{H_n(\omega) - G_n(\omega)}{a_n} = 0 \qquad \text{a.e.} \quad \text{on} \quad D(\alpha). \tag{11}$$

where

$$D_{(\alpha)} = \{\limsup_{n\to\infty} \frac{1}{a_n} \sum_{t \in T^{(n)}_{(N+1)}} E[g_t^2(X_{(N+1)_t}, X_{N_t}) e^{a|g_t(X_{(N+1)_t}, X_{N_t})} \big| X_{(N+1)_t}] = M(\omega) < \infty\} \cap B \tag{12}$$

and

$$B = \left\{ \lim_{n\to\infty} a_n = \infty \right\} \tag{13}$$

Proof: By Lemma 1, we have known that $\left\{t_{n-N}(\lambda,\omega), F_{n-N}, n \geq N+1\right\}$ is a nonnegative martingale. According to Doob martingale convergence theorem we have

$$\lim_{n\to\infty} t_{n-N}(\lambda,\omega) = t(\lambda,\omega) < \infty \qquad \text{a.e.} \tag{14}$$

We have by (13) and (14)

$$\limsup_{n\to\infty} \frac{\ln t_{n-N}(\lambda,\omega)}{a_n} \leq 0 \qquad \text{a.e.} \tag{15}$$

By (8),(9) and (15), we get

$$\limsup_{n\to\infty} \frac{1}{a_n} \{\lambda H_n(\omega) - \sum_{t \in T^{(n)}_{(N+1)}} \ln[E[e^{\lambda g(X_{(N+1)_t}, X_{N_t})} \big| X_{(N-1)_t}]]\} \leq 0 \qquad \text{a.e.} \tag{16}$$

Let $\lambda > 0$, dividing two sides of (16) by λ, we have

$$\limsup_{n\to\infty} \frac{1}{a_n} \{H_n(\omega) - \sum_{t \in T^{(n)}_{(N+1)}} \frac{\ln[E[e^{\lambda g(X_{(N+1)_t}, X_{N_t})} \big| X_{(N-1)_t}]]}{\lambda}\} \leq 0 \qquad \text{a.e.} \tag{17}$$

Taking $0 \le \lambda < -\alpha$, we arrive at

$$\limsup_{n\to\infty} \frac{1}{a_n}\{H_n(\omega) - \sum_{t\in T^{(n)}_{(N+1)}} E[g_t(X_{(\mathrm{N+1})_t}, X_{N_t}) \big| X_{(\mathrm{N+1})_t}]\}$$

$$\le \limsup_{n\to\infty} \frac{1}{a_n} \sum_{t\in T^{(n)}_{(N+1)}} \frac{\ln[E[e^{\lambda g(X_{(\mathrm{N+1})_t}, X_{N_t})} \big| X_{(\mathrm{N-1})_t}]]}{\lambda} - E[g_t(X_{(\mathrm{N+1})_t}, X_{N_t}) \big| X_{(\mathrm{N+1})_t}]\}$$

$$\le \limsup_{n\to\infty} \frac{1}{a_n} \sum_{t\in T^{(n)}_{(N+1)}} \frac{E[e^{\lambda g(X_{(\mathrm{N+1})_t}, X_{N_t})} \big| X_{(\mathrm{N-1})_t}] - 1}{\lambda} - E[g_t(X_{(\mathrm{N+1})_t}, X_{N_t}) \big| X_{(\mathrm{N+1})_t}]\}$$

$$\le \frac{\lambda}{2} \limsup_{n\to\infty} \frac{1}{a_n} E[g_t^{\,2}(X_{(\mathrm{N+1})_t}, X_{N_t}) e^{\lambda |(X_{(\mathrm{N+1})_t}, X_{N_t})} \big| X_{(\mathrm{N+1})_t}]$$

$$= \frac{\lambda}{2} M(\omega) \qquad a.e. \quad \omega \in D(a) \tag{18}$$

where the first inequality follows by (17) and the fact that

$$\limsup_{n\to\infty}(c_n + b_n) \le \limsup_{n\to\infty} c_n + \limsup_{n\to\infty} b_n$$

the second follows by the inequality $\ln x \le x - 1 (x > 0)$

the third follows by the inequality $0 \le e^x - x - 1 \le \frac{1}{2} x^2 e^{|x|}$

Letting $\lambda \to 0^+$ $\lambda \to 0^-$ in (18), by (10) we have

$$\limsup_{n\to\infty} \frac{H_n(\omega) - G_n(\omega)}{a_n} \le 0 \qquad a.e. \quad \omega \in D(a) \tag{19}$$

Let $-\alpha \le \lambda < 0$. By (16), we similarly get

$$\liminf_{n\to\infty} \frac{H_n(\omega) - G_n(\omega)}{a_n} \ge \frac{\lambda}{2} M(\omega) \qquad a.e. \quad \omega \in D(a) \tag{20}$$

Letting $\lambda \to 0^-$, we can arrive at

$$\liminf_{n\to\infty} \frac{H_n(\omega) - G_n(\omega)}{a_n} \ge 0 \qquad a.e. \quad \omega \in D(a) \tag{21}$$

Combing (19) and (21), we obtain (11) directly. This completes the proof of Theorem 1.

Let $k \in S$, $S_n(k)$ be the number of k in the set of random variables $X^{T^{(n)}}$, and $S^N{}_n(k)$ be the number of k's N-th descendants in the set of random variables $X^{T^{(n)}}$, that is

$$S_n(k) = \sum_{t \in T^{(n)}} \delta(X_t), \tag{22}$$

$$S_n^N(k) = \sum_{t \in T^{(n)}} \delta_k(X_{N_t}) \tag{23}$$

Corollary 1. For any nonnegative integer N, $S^N{}_n(k)$ be the number of k's N-th descendants. If moreover

$$\lim_{|t| \to \infty} P_t(k|l) = P_t(k|l) \tag{24}$$

then

$$\lim_{n \to \infty} \frac{1}{|T^{(n)}|} \{S_n^N(k) - \sum_{l \in S} P_t(k|l) S_n^N(l)\} = 0 \tag{25}$$

Further more, for any positive integer m <n-N, we have

$$\lim_{n \to \infty} \frac{1}{|T^{(n)}|} \{S_n^N(k) - \sum_{l \in S} P^m(k|l) S_n^{N+m}(l)\} = 0 \tag{26}$$

Where $P^m(k \mid l)$ is the m-step transition probability determined by the transition matrix $(P_t(y \mid x))$.

Proof. Let $g_t(X_{(\mathrm{N}+1)_t}, X_{N_t}) = \delta_k(X_{\mathrm{N}_t})$ $a_n = |T^{(n)}|$ in Theorem 1, then (25) and (26) follow obviously.

Theorem 2. If (24) holds, then

$$\lim_{n \to \infty} \frac{S_n^N(k)}{|T^{(n)}|} = \pi(k) \tag{27}$$

$$\lim_{n \to \infty} \frac{S_n(k)}{|T^{(n)}|} = \pi(k) \tag{28}$$

Where $\pi = (\pi(0), \ldots, \pi(b-1))$ is the uniquestationary distribution determined by transition matrix $(P(i|j)_{i,j \in S})$.

Proof. Obviously, (28) follows by setting N=0 in (32). Hence, we only need to proof (27). By Corollary 1,

$$\begin{aligned}
&\limsup_{n \to \infty} \frac{S_n^N(k)}{|T^{(n)}|} - \pi(k) \\
&= \limsup_{n \to \infty} \frac{1}{|T^{(n)}|} \left| S_n^N(k) - \sum_{l \in S} P^m(k|l) S_n^{N+m}(l) + \sum_{l \in S} P^m(k|l) S_n^{N+m}(l) - \pi(k) \right| \\
&\le \sum_{l \in S} \left| P^m(k|l) - \pi(k) \right|
\end{aligned} \tag{29}$$

Since

$$P^m(k|l) \to \pi(k), m \to \infty, \tag{30}$$

by (30), the right hand side of (29) is arbitrary small for enough large m, which implies the equation (27). We complete the proof.

Let T be a tree,

$$f_n(\omega) = -\frac{1}{|T^{(n)}|} \ln P(X^{T^{(n)}}), \tag{31}$$

will be called the entropy density. If $\{X_t, t \in T\}$ is a T-indexed nonhomogeneous Markov chain with states space S , we have by (6),

$$f_n(\omega) = -\frac{1}{|T^{(n)}|}[\ln P(X_0) + \sum_{t \in T^{(n)}_{(N+1)}} \ln P(X_t | X_{1t})] \tag{32}$$

The convergence of $f_n(\omega)$ to a constant in a sense (L_1 convergence , convergence in probability, a.e. convergence) is called the Shannon-McMillan theorem or the entropy theorem or the AEP in information theory.

Theorem 3. Let T be an infinite tree, then

$$\lim_{n\to\infty} f_n(\omega) = \sum_{j \in S} \pi(j) H[P(0|j), ...P(b-1|j)] \tag{33}$$

.

Proof. By (9), (10) and (32),

$$\frac{H_n(\omega)}{|T^{(n)}|} = -\frac{1}{|T^{(n)}|} \sum_{t \in T^{(n)}_{(N+1)}} \ln P_t(X_t | X_{1t}) = f_n(\omega) + \frac{\ln p(X_0)}{|T^{(n)}|} \tag{34}$$

$$\frac{G_n(\omega)}{|T^{(n)}|} = -\frac{1}{|T^{(n)}|} \sum_{t \in T^{(n)}_{(N+1)}} E[\ln P_t(X_t | X_{1t}) | X_{1t}]$$

$$= -\frac{1}{|T^{(n)}|} \sum_{t \in T^{(n)}_{(N+1)}} H[P_t(0|X_{1t}), ...P_t(b-1)|X_{1t}] \tag{35}$$

If (18) holds, we can prove easily

$$\lim_{n\to\infty} \frac{1}{|T^{(n)}|} \sum_{t \in T^{(n)}_{(N+1)}} \{H[P_t(0|X_{1t}), ..., P(b-1|X_{1t}) - H[P(0|X_{1t}), ..., P(b-1|X_{1t})\} = 0 \tag{36}$$

By Theorem 2, we get

$$\lim_{n\to\infty}\frac{1}{\left|T^{(n)}\right|}\times\sum_{t\in T_{(N+1)}^{(n)}} H[P_t(0\left|X_{1t}\right.),.....,P(b-1\left|X_{1t}\right.)]$$

$$=\lim_{n\to\infty}\frac{1}{\left|T^{(n)}\right|}\sum_{t\in T_{(1)}^{(n)}}\sum_{j\in S} H[P(0\left|j\right.),.....,P(b-1\left|j\right.)]\delta_j(X_{1t})$$

$$=\sum_{j\in S}\pi(j)H[P(0\left|j\right.),.....,P(b-1\left|j\right.)] \tag{37}$$

by (34), (35), (36) and (37), (33) holds. This is the end of the proof.

Acknowledgement. The work is partly supported by Yong Talents Funds of Anhui (2010SQ RL129) and Chaohu University Scientific Research Foundation.

References

[1] Benjamini, I., Peres, Y.: Markov chains indexed by trees. Ann. Probab. 22, 219–243 (1994)

[2] Berger, T., Ye, Z.-X.: Entropic aspects of random fields on trees. IEEE Trans. Inform. Theory 36(5), 1006–1018 (1990)

[3] Yang, W.-G., Ye, Z.-X.: The Asympotoic equipartition property for nonhomogeneous Markov chains indexed by a Homogeneous tree. IEEE Trans. Inform. Theory 53(9), 3275–3280 (2007)

[4] Huang, H.-L., Yang, W.-G.: Strong law of large numbers for Markov chains indexed by an infinitetree with uniformly bounded degeree. Science in China Series A: Mathematics 51(2), 195–202 (2008)