Statistics and Probability Letters 84 (2014) 247–252

Contents lists available at ScienceDirect

Statistics and Probability Letters

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

A note on the relaxation time of two Markov chains on rooted phylogenetic tree spaces

David A. Spade [a,*], Radu Herbei [b], Laura S. Kubatko [b]

[a] University of Missouri–Kansas City, Kansas City, MO, 64110, United States
[b] The Ohio State University, Columbus, OH, 43210, United States

ARTICLE INFO

Article history:
Received 8 May 2012
Received in revised form 19 August 2013
Accepted 18 September 2013
Available online 29 October 2013

Keywords:
Markov chains
Phylogenetic trees
Relaxation time
Distinguished paths

ABSTRACT

Phylogenetic trees are commonly used to model the evolutionary relationships among a collection of biological species. Over the past fifteen years, the convergence properties for Markov chains defined on phylogenetic trees have been studied, yielding results about the time required for such chains to converge to their stationary distributions. In this work we derive an upper bound on the relaxation time of two Markov chains on rooted binary trees: one defined by nearest neighbor interchanges (NNI) and the other defined by subtree prune and regraft (SPR) moves.

1. Introduction

In biology, it is often of interest to study the patterns of evolution among a collection of species. Typical assumptions are that the species have evolved from a common ancestor and that the process of speciation results in the formation of two new species at a single point in time. To visually describe these assumptions, biologists commonly use a rooted, binary tree called a phylogenetic tree. This undirected, acyclic graph has n external vertices, called *leaves*, and $n-2$ internal nodes of degree 3. The graph also has one node of degree 2 that shall be termed the *root*. Markov chain Monte Carlo methods are frequently used to estimate the distribution of these trees given DNA sequences at the leaves. Therefore, understanding rates of convergence of Markov chains widely used in phylogenetics is important in efficiently estimating phylogenetic trees.

While a considerable amount of work has been done in the area of mixing times for Markov chains, the application of these techniques to phylogenetics has only been considered in the past fifteen years. Diaconis and Stroock (1991) develop inequalities that are integral to our study of relaxation and mixing times. In particular, they establish bounds on the spectral gap of the transition matrix of an irreducible, aperiodic, and reversible Markov chain. Aldous (2000) explores the idea of using chain coupling to establish bounds on the mixing time of a Markov chain on unrooted phylogenies. His work gives an $O(n^3)$ upper bound on the relaxation time of a chain where a step of the chain consists of removing a leaf from a tree and then attaching it to another edge. Aldous (2000) also expands the concepts brought forth by Diaconis and Stroock (1991), proving that the relaxation time of his chain is bounded below by $O(n^2)$. He conjectures that the relaxation time is also bounded above by $O(n^2)$, and Schweinsberg (2002) later proves Aldous's conjecture using the method of distinguished paths.

Randall and Tetali (1999) investigate the time required for a Markov chain on rooted phylogenetic trees to converge to its stationary distribution. Their chain moves about the set T_n of n-leaf rooted phylogenetic trees by performing at each step of the chain a tree rearrangement similar to those in one of the Markov chains we describe below. They also establish that the mixing time of the chain under study is $O(n^5 \log n)$.

* Corresponding author. Tel.: +1 816 235 2853.
E-mail addresses: spaded@umkc.edu (D.A. Spade), herbei@stat.osu.edu (R. Herbei), lkubatko@stat.osu.edu (L.S. Kubatko).

0167-7152/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.spl.2013.09.017

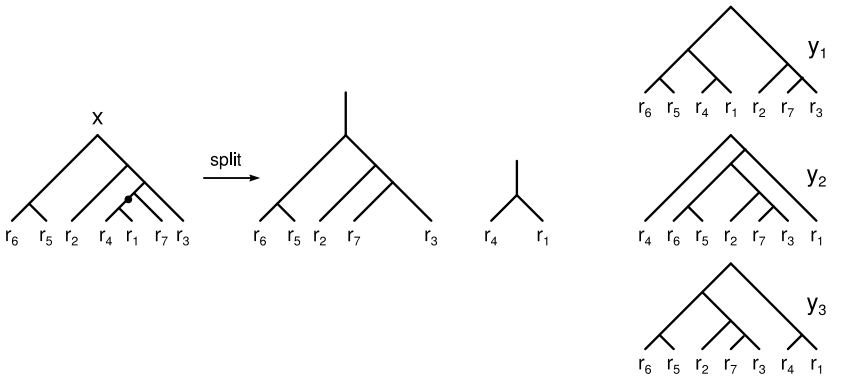

Fig. 1. Example SPR moves. Tree x is split into two subtrees shown to the right along the branch marked with a black dot. Tree y_1 is formed by re-attaching the rightmost subtree to the branch ancestral to the clade containing leaves r_5 and r_6. Tree y_2 is formed by re-attaching the leftmost subtree to the branch ancestral to leaf r_1. The tree y_3 is formed by re-attaching the two branches extending back from the roots of the trees formed by splitting x.

In the work we present here, we establish upper bounds on the relaxation time of two particular Markov chains on rooted binary trees. One of these Markov chains moves about T_n in a fashion similar to the chain of Aldous (2000) and Schweinsberg (2002). We add some conventions to these tree rearrangements in order to handle situations that arise with rooted trees, but that do not occur with unrooted trees. The other Markov chain moves about the tree space by making moves similar to those in the work of Randall and Tetali (1999).

2. Background and notation

Let T_n be the set of n-leaf rooted trees, having cardinality $c_n \equiv (2n-3)!!$ (Felsenstein, 2004). A homogeneous Markov chain $\{X_t\}_{t\geq 0}$ on T_n is defined via a transition probability matrix of $\mathbf{P}(x, y) = \Pr(X_{m+1} = y \mid X_m = x)$, for each $x, y \in \mathsf{T}_n$. We assume that $\mathbf{P}(\cdot, \cdot)$ satisfies the usual regularity conditions ensuring existence and uniqueness of a stationary distribution $\pi(\cdot)$. In this paper we focus on two types of transitions on T_n: (1) subtree prune and regraft (SPR) and (2) nearest neighbor interchange (NNI).

An SPR transition consists of choosing an edge uniformly at random and pruning the subtree that descends from this edge. The pruned subtree is viewed as being rooted at the node that descends from the selected edge, keeping the edge extending out from this root. The remaining tree is viewed as being rooted at the most recent common ancestor (MRCA) of the remaining leaves, with an added edge that extends back from this tree. An edge is randomly selected (from either of the two subtrees) and the edge extending back from the root of the other subtree is attached to this edge. The root of the resulting tree is either the root of the subtree to which the edge is attached (when the randomly selected edge for re-attachment is not the one extending back from the root) or is the node formed by re-attaching the subtree (when the edge selected for re-attachment is the one extending back from the root). Fig. 1 gives an example of three possible SPR moves.

An NNI transition is performed by first choosing an internal node (other than the root node) to be the *target node*. The target node has two child nodes and a sibling node. Two of these three nodes, along with their descendant subtrees, will be selected to become the new children of the target node. With probability 0.5, we select the two current children (and the tree does not change), and with probability 0.5 we select the sibling node and one of the two children at random. In this situation, the child that is not selected becomes the new sibling of the target node. Let $\{\mathsf{X}_t\}_{t\geq 0}$ be the Markov chain resulting from SPR transitions and $\{\mathsf{Y}_t\}_{t\geq 0}$ be the chain resulting from NNI transitions. The following lemma describes the transition probability matrices for the two chains.

Lemma 1. *Let $\mathbf{P}_1$ and $\mathbf{P}_2$ be the transition probability matrices for the SPR and NNI chains, respectively. Then*

(i) *For each $x, y \in \mathsf{T}_n$, such that $\mathbf{P}_1(x, y) > 0$, $\mathbf{P}_1(x, y) \geq \frac{1}{(2n-2)^2}$.*

(ii) *For each $x, y \in \mathsf{T}_n$ such that $\mathbf{P}_2(x, y) > 0$, $\mathbf{P}_2(x, y) = \frac{1}{2}$ if $x = y$ and $\mathbf{P}_2(x, y) = \frac{1}{4(n-2)}$, otherwise.*

Part (i) of this lemma follows from the fact that there are $2n-2$ choices for the edge that is cut and, given the edge that is cut, there are $2n-2$ ways to re-attach the two resulting subtrees. Some of these may give the same transition between trees. The proof of part (ii) of this lemma is straightforward and is thus omitted from this manuscript. Note that both the NNI and the SPR transitions are reversible and symmetric and that the resulting Markov chains are aperiodic and irreducible (Karlin and Taylor, 1975), thus ergodic. In both cases, the unique stationary distribution is the uniform distribution, $\pi(x) = 1/c_n$, for each $x \in \mathsf{T}_n$.

Often, it is of interest to study the rate at which an ergodic Markov chain converges to its stationary distribution. This is typically done by obtaining upper and lower bounds on the *mixing time* of the Markov chain. The mixing time of the process $\{X_m\}_{m\geq 0}$ is defined as $\tau_{\text{mix}}(\epsilon) := \min\left\{m : \max_{x\in\mathsf{T}_n} \|\mathbf{P}^m(x, \cdot) - \pi(\cdot)\|_{TV} < \epsilon\right\}$, where $\|\mu(\cdot) - \nu(\cdot)\|_{TV}$ denotes the total