Adjacent Transpositions Markov Chain

Leonardo Bonanno, Julia Palacios

June 12, 2020

1 Introduction and Definitions

Let τ_n^R denote the set of ranked unlabeled trees with n leaves and O_n denote the set of ordered matchings of size n (defined below). [**?**] shows that there is a bijection between τ_n^R and O_n. Thus, we can define a markov chain (known as the random transpositions chain) on O_n as seen in the section below, first introducing some notation.

Notation 1.1. *An ordered matching of size n is a sequence* $\{(a_i, b_i)\}_{i=1}^{n-1}$ *where* $a_i, b_i \in \{0, \dots n-2\}$, $a_i, b_i < i$, *and* $a_i \leq b_i$ *for all* $i \in \{1, \dots n-1\}$. *Additionally, the elements of* $\{1, \dots n-1\}$ *appear in exactly one pair in the sequence.*

Notation 1.2. *Let* $O \in O_n$. *Then, for* $i \in \{0, \dots n-2\}$ *let* $O[i]$ *denote the ith pair in* O *and for* $j \in \{0, 1\}$ *let* $O[i][j]$ *denote the j-th element of* $O[i]$. *Additionally, let* $C\left(O\right)$ *denote the number of cherries of* $O = \{(a_i, b_i)\}$, *namely let* $C\left(O\right) = |\{i : (a_i, b_i) = (0, 0)\}|$.

2 Markov Chain

In this section we define a markov chain on O_n. Let $X_t = \left(a_1, b_1\right)^1 \left(a_2, b_2\right)^2 \dots \left(a_{n-1}, b_{n-1}\right)^{n-1}$ be the state of the markov chain at time t. Now let $k, k' \in \{1, \dots n-2\}$. Then, if there are any valid ones, choose uniformly at random between valid transpositions of elements from $O[k]$ and $O'[k]$ and perform the chosen transposition to reach X_{t+1}.We now investigate a few properties of this chain. We note here that these conditions guarantee that $(a_1, b_1) = (0, 0)$ and that $|\{i : a_i = 0\}| + |\{i : b_i = 0\}| = 2\left(n-1\right) - \left(n-2\right) = n$

Lemma 2.1. — *The random transpositions chain is reversible with respect to the discrete tajima coalaescent. Namely for* $O \in O_n$, $\pi\left(O\right) = \frac{2^{n-C(O)-1}}{(n-1)!}$

Proof. Let $O, O' \in O_n$. If $P\left(O, O'\right) > 0$ then we can reach O' from O with one transposition. Thus $|C\left(O\right) - C\left(O'\right)| \leq 1$. It is easy to verify that, $\frac{P(O,O')}{P(O',O)} = 1$ if $C\left(O\right) = C\left(O'\right)$, $\frac{P(O,O')}{P(O',O)} = \frac{1}{2}$ if $C\left(O\right) = C\left(O'\right) - 1$, and $\frac{P(O,O')}{P(O',O)} = 2$ if $C\left(O\right) = C\left(O'\right) + 1$. Thus, $\frac{P(O,O')}{P(O',O)} = 2^{C(O)-C(O')} = \frac{\pi(O')}{\pi(O)}$. $\square$

We note here that is is possible to adapt this chain to be reversible with respect to the stationary distribution over O_n if when we select a cherry to transpose, we treat it as a singleton (0).

Lemma 2.2. *The random transpositions chain is ergodic (aperiodic and irreducible).*

Proof. For the case $n = 2, 3, 4$, this is easy to verify by hand. Now let $n \geq 5$. From earlier in this section we have showed that there are thus at least 5 zeros, and that for any $O \in O_n$, $O[0] = (0, 0)$. Thus, there will be at least two $i, j \geq 1$ such that $O[i]$ and $O[j]$ contain zeros and thus

$$P(O, O) \geq \frac{1}{2(n-2)(n-1)} > 0$$

. We now show that the chain is irreducible. To do so, we define a positive probability path between any two states in O_n. Let $O = \{(a_i, b_i)\}_{i=1}^{n-1}$ and $O' = \{(c_i, d_i)\}_{i=1}^{n-1}$. We define a path Γ between O and O' inductively. First, let $\Gamma_1 = 0$ and let Γ_n be the path after n steps. Let $\Gamma_n[i] \in O_n$ denote the ith element of the path Γ_n. If $\Gamma_n[n] = O'$ then stop and let $\Gamma = \Gamma_n$. Now let $k = \min\{i : O[i] \neq O'[i]\}$. Then, for $i \geq k$, there exists $(e_i, f_i) \in \{0, \ldots n-2\}^2$ such that $\Gamma_n[n] = (c_1, d_1)^1 \ldots (c_{k-1}, d_{k-1})^{k-1} (e_k, f_k)^k (e_{k+1}, f_{k+1})^{k+1} \ldots (e_{n-1}, f_{n-1})^{n-1}$. We first consider the case that $e_k = c_k$. In this case, there exists $i \geq k+1$, either $e_i = d_k$ or $f_i = d_k$. If $e_i = d_k$ then let

$$X_n = (c_1, d_1)^1 \ldots (c_{k-1}, d_{k-1})^{k-1} (c_k, d_k)^k (e_{k+1}, f_{k+1})^{k+1} \ldots (e_{i-1}, f_{i-1})^{i-1} (e_k, f_i)^i (e_{i+1}, f_{i+1})^{i+1} \ldots (e_{n-1}, f_{n-1})^{n-}$$

If $f_i = d_k$ then let

$$X_n = (c_1, d_1)^1 \ldots (c_{k-1}, d_{k-1})^{k-1} (c_k, d_k)^k (e_{k+1}, f_{k+1})^{k+1} \ldots (e_{i-1}, f_{i-1})^{i-1} (e_i, f_k)^i (e_{i+1}, f_{i+1})^{i+1} \ldots (e_{n-1}, f_{n-1})^{n-}$$

In either case let $\Gamma_{n+1}[i] = \Gamma_n[i]$ for $i \leq n$ and $\Gamma_{n+1}[n+1] = X_n$. We now consider the case that $e_k \neq c_k$. There are two subcases to consider. The first is the case that $c_k > \max(e_k, f_k)$, the second that $c_k \leq \max(e_k, f_k)$. We first consider the first case. In this case, there exists $i \geq k+1$ such that either $e_i = d_k$ or $f_i = d_k$. In the first subcase we let,

$$X_n = (c_1, d_1)^1 \ldots (c_{k-1}, d_{k-1})^{k-1} (e_k, d_k)^k (e_{k+1}, f_{k+1})^{k+1} \ldots (e_{i-1}, f_{i-1})^{i-1} (f_k, f_i)^i (e_{i+1}, f_{i+1})^{i+1} \ldots (e_{n-1}, f_{n-1})^{n-}$$

and in the second case we let

$$X_n = (c_1, d_1)^1 \ldots (c_{k-1}, d_{k-1})^{k-1} (e_k, d_k)^k (e_{k+1}, f_{k+1})^{k+1} \ldots (e_{i-1}, f_{i-1})^{i-1} (e_i, f_k)^i (e_{i+1}, f_{i+1})^{i+1} \ldots (e_{n-1}, f_{n-1})^{n-}$$

In either case let $\Gamma_{n+1}[i] = \Gamma_n[i]$ for $i \leq n$ and $\Gamma_{n+1}[n+1] = X_n$. We now consider the second subcase where $c_k \leq \max(e_k, f_k)$. We first consider the subsubcase that $f_k \neq c_k$, then the subsubcase that $f_k = c_k$. Then there exists $i \geq k+1$ such that either $e_i = c_k$ or $f_i = c_k$. In the former we let

$$X_n = (c_1, d_1)^1 \ldots (c_{k-1}, d_{k-1})^{k-1} (c_k, f_k)^k (e_{k+1}, f_{k+1})^{k+1} \ldots (e_{i-1}, f_{i-1})^{i-1} (e_k, f_i)^i (e_{i+1}, f_{i+1})^{i+1} \ldots (e_{n-1}, f_{n-1})^{n-}$$

and in the latter we let

$$X_n = (c_1, d_1)^1 \ldots (c_{k-1}, d_{k-1})^{k-1} (c_k, f_k)^k (e_{k+1}, f_{k+1})^{k+1} \ldots (e_{i-1}, f_{i-1})^{i-1} (e_i, e_k)^i (e_{i+1}, f_{i+1})^{i+1} \ldots (e_{n-1}, f_{n-1})^{n-}$$

We now consider the case that $f_k = c_k$. Then, there exists some $i \geq k+1$ such that either $e_i = d_k$ or $f_i = d_k$. In the former we let

$$X_n = (c_1, d_1)^1 \dots (c_{k-1}, d_{k-1})^{k-1} (c_k, d_k)^k (e_{k+1}, f_{k+1})^{k+1} \dots (e_{i-1}, f_{i-1})^{i-1} (e_k, f_i)^i (e_{i+1}, f_{i+1})^{i+1} \dots (e_{n-1}, f_{n-1})^{n-}$$

and in the latter we let

$$X_n = (c_1, d_1)^1 \dots (c_{k-1}, d_{k-1})^{k-1} (c_k, d_k)^k (e_{k+1}, f_{k+1})^{k+1} \dots (e_{i-1}, f_{i-1})^{i-1} (e_i, f_k)^i (e_{i+1}, f_{i+1})^{i+1} \dots (e_{n-1}, f_{n-1})^{n-}$$

In either case let $\Gamma_{n+1}[i] = \Gamma_n[i]$ for $i \leq n$ and $\Gamma_{n+1}[n+1] = X_n$. It is clear from this process that within $2n$ steps the path will reach O', because at each step, at least one more similarity is being added and none are being destroyed. Thus the chain is irreducible. The path we constructed here will be used when we discuss the use of the path method for the upper bound on the mixing time of this chain below. $\square$

3 Example: n = 5

In this section we consider the case where $n = 5$. There are 5 matchings for the case $n = 5$. Namely,

1. $(0,0)^1 (0,0)^2 (1,2)^3 (0,3)^4$
2. $(0,0)^1 (0,0)^2 (0,2)^3 (1,3)^4$
3. $(0,0)^1 (0,0)^2 (0,1)^3 (2,3)^4$
4. $(0,0)^1 (0,1)^2 (0,2)^3 (0,3)^4$
5. $(0,0)^1 (0,1)^2 (0,0)^3 (2,3)^4$

The transition matrix for the chain in this case is

$$\begin{bmatrix} \frac{1}{3} & \frac{1}{6} & \frac{1}{6} & \frac{1}{3} & 0 \\ \frac{1}{6} & \frac{1}{3} & \frac{1}{6} & \frac{1}{3} & 0 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{2} & 0 & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & 0 & \frac{1}{3} & \frac{1}{6} \\ 0 & 0 & \frac{1}{6} & \frac{1}{3} & \frac{1}{2} \end{bmatrix}$$

The stationary distribution is $\pi = \begin{bmatrix} \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{3} & \frac{1}{6} \end{bmatrix}$, exactly the discrete tajima coalescent. The second largest eigenvalue for this matrix is 0.953 and the relaxation time is about 21.27.

4 Lower Bound on Mixing Time

It will be useful for further purposes to have a recursive formula for the number of trees with n leaves and c cherries.

Lemma 4.1. *Let $T_{n,c}$ denote the set of ranked unlabeled trees with n leaves and c cherries. Then*

$$|T_{n,c}| = |T_{n-1,c}| \cdot (c) + |T_{n-1,c-1}| \cdot (n - 2c + 1)$$

with $|T_{a,1}| = 1$ and $|T_{a,0}| = 0$ for all $a \geq 2$ and $|T_{n,c}| = 0$ when $2c > n$.

Proof. This follows directly from how you recursively build up a tree. It would be great to find a limiting distribution for this quantity. □

Conjecture 4.2. *For the random transpositions chain $t_{rel} \geq O\left(n^2\right)$.*

Proof. For $O \in O_n$, let $C(O)$ denote the number of cherries of O as above. [] establishes that $Var_\pi(O) = \frac{2n}{45}$

$$t_{rel} = \sup_{f:O_n \to \mathbb{R}} \frac{Var_\pi(f)}{\varepsilon(f)} \geq \frac{2}{45} \frac{n}{\varepsilon(C)}$$

To prove the $O\left(n^2\right)$ bound we would need the following result.

Conjecture 4.3. $\varepsilon(C) \leq O\left(\frac{1}{n}\right)$

Proof. We compute directly,

$$\varepsilon(C) = \frac{1}{2} \sum_{x,y \in O_n} \pi(x) P(x,y) (C(x) - C(y))^2 = \frac{1}{2} \frac{2^{n-1}}{(n-1)!} \sum_{x \in O_n} 2^{-c} |\{y \in O_n ; P(x,y) > 0, |C(x) - C(y)| = 1\}|$$

$$\leq \frac{1}{2} \frac{2^{n-1}}{(n-1)!} \sum_{c=1}^{\lfloor \frac{n}{2} \rfloor} \frac{4(c-1)(n-c-1)}{n^2} 2^{-c} |T_{n,c}|$$

□

After calculating the cardinality of the $|T_{n,c}|$'s experimentalally, it seems unlikely that this conjecture is true □

5 Upper bound on Mixing Time

6 References