International Journal of Applied Mathematics

Volume 31 No. 2 2018, 241-249
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)
doi: http://dx.doi.org/10.12732/ijam.v31i2.6

LIMITING PROBABILITY TRANSITION MATRIX OF A CONDENSED FIBONACCI TREE

K.A. Germina

Department of Mathematics
Central University of Kerala
Kasargod, Kerala, INDIA

Abstract: This paper discusses on the construction of condensed Fibonacci trees and present the Markov chain corresponding to the condensed Fibonacci trees. An $n \times n$ finite Markov probability transition matrix for this Markov chain is presented and it is proved that the limiting steady state probabilities are proportional to the first n Fibonacci numbers.

AMS Subject Classification: 05C07, 05C38, 05C75, 05C85
Key Words: Fibonacci numbers, matrix methods, Markov chains, binary search trees, Fibonacci trees

1. Introduction

Most natural / artificial dynamic phenomena are endowed with a nondeterministic description. Stochastic processes provide appropriate mathematical models of such phenomena. Among the class of stochastic processes, Markov chains are highly utilized for many dynamic phenomena because of demonstration of equilibrium behavior. Efficient computation of equilibrium / transient probability distribution of a Discrete Time Markov Chain (DTMC) / Continuous Time Markov Chain (CTMC) is considered to be an interesting research problem. A Markov chain on Ω is a stochastic process $\{X_0, X_1 \ldots, X_t \ldots\}$ with each $X_i \in \Omega$ such that $Pr(X_{t_1} = y/X_t = x, X_{t-1} = x_{t-1}, \ldots,\ X_0 = x_0) =$

Received: January 22, 2018

$Pr(X_{t_1} = y/X_t = x =: P(x,y)$. Hence, the Markov chain in the matrix can be described by a $|\Omega \times \Omega|$ matrix P whose $(i,j)^{th}$ entry is $P(x,y)$. A random variable is said to be discrete, if there exists a finite or countable set $S \subset \mathcal{M}$ such that $P[X \in S] = 1$.

Definition 1. Let $\{X_k\}$ be a discrete-time stochastic process which takes on values in a countable set Ω, called the *state space*. $\{X_k\}$ is called a *Discrete Time Markov chain* or simply a Markov chain, when the discrete nature of the time index is clear, if $P(X_k = i_k|X_{k-1} = i_{k-1}, X_{k-2} = i_{k-2}, \ldots) = P(X_k = i_k|X_{k-1} = i_{k-1})$, where $i_j \in S$. A Markov chain is said to be time homogeneous, if $P(X_k = i|X_{k-1} = j)$ is independent of k.

Associated with each Markov chain there is a matrix called the probability transition matrix, denoted by P, whose $(i,j)^{th}$ element is given by $P_{ij} = P(x_k = j|x_{k-1} = i)$. Let $p[k]$ denote a row vector of probabilities with $p_j[k] = P(X_k = j)$. This vector of probabilities evolves according to the equation $p[k] = p[k-1]P$. Thus, $p[0]$ and P capture all the relevant information about the dynamics of the Markov chain. If there exist a π so that $\pi = \pi P$, then π is called a stationary distribution. With stationary distributions defined, a natural question is whether every Markov chain has a stationary distribution, and whether it is unique. Also, if there exists a unique stationary distribution, is it always $\lim_{k\to\infty} p[k] = \pi$ for all $p[0]$?. In other words, does the distribution of the Markov chain converge to the stationary distribution starting from any initial state?. Unfortunately, not all Markov chains have a unique steady-state distribution.

2. Graphical representation of Markov Chain

Consider a Markov chain with state space Ω and transition matrix P. Then, a corresponding graphical representation is the weighted graph $G = (V, E)$, where $V = \Omega$ and $E = \{(x,y) \in \Omega \times \Omega/P(x,y) > 0\}$. Also, edge (x,y) has weight $P(x,y) > 0$. Self loops are allowed since one can have $P(x,x) > 0$. An edge is present between x and y if and only if the transition probability between x and y is non-zero. The critical point in the theory of Markov chain depends not on the exact values of the entry but rather on whether a particular entry is zero or not. Hence, in terms of graphs, the critical thing is the structure of the graph G rather than the values of its edge weights. Hence, given a homogeneous DTMC, there is a state transition matrix which is naturally associated with its equilibrium / transient behavior, and the state transition diagram, which is a weighted directed graph. The weights are effectively probabilities in such a directed graph. Similarly, a homogeneous CTMC is represented using a state transition rate diagram which is a weighted, directed graph. The weights are *state transition rates*. In general, it is always possible to represent a time-homogeneous Markov chain by a transition graph.

A Markov chain is *irreducible* if there is some value n so that for any states x, y, if you start at x and take n steps, there is a positive probability you end up at y. Equivalently, if the Markov chain has transition matrix P, then all entries of P^n are strictly positive. One can think of irreducibility of a Markov chain as a property of the graph that marks which states can transition to which state. To a Markov chain M, associate a directed reachability graph G_M whose nodes are states, with a directed edge for each transition that occurs with positive probability. The digraph $G(P)$, corresponding to an irreducible Markov chain P, is hence strongly connected, and $G(P)$ is connected in the case if the corresponding graphical representation is undirected. A Markov chain M is *ergodic* if and only if (i) the graph G_M is irreducible. and (ii) the graph G_M aperiodic, which means, the gcd of the lengths of positive probability cycles of G_M is 1. Hence, the strongly-connectedness condition means that no matter which state you start in, you eventually reach every other state with positive probability. The aperiodicity condition rules out, that a random walk on a bipartite graph where you necessarily alternate between the two halves on odd and even steps. Together, these conditions suffice to guarantee that there's a unique stationary distribution.

We have the following results proved in [1].

Theorem 2. ([1]) *Any ergodic Markov chain has a unique stationary distribution.*

Theorem 3. ([1]) *A finite state space, irreducible Markov chain has a unique stationary distribution π and if it is aperiodic, $\lim_{k\to\infty} p[k] = \pi$ for all $p[0]$.*

3. Markov chains have a unique steady-state distribution

A number of important properties of the Markov chain (typically derived using matrix manipulations) can be deduced from this graphical representation. Moreover, certain concepts from algebra will also be illuminated by developing this approach. In addition, the graph-theoretic representation immediately suggests several computational schemes for calculating important structural characteristics of the underlying problem. An important concept in the analysis of Markov chains is the categorization of states as either recurrent or transient. The Markov chain, once started in a recurrent state, will return to that state with probability 1. However, for a transient state there is some positive probability that the chain, once started in that state, will never return to it. This same concept can be illuminated using graph-theoretic concepts, which do not involve the numerical values of probabilities. Define node i to be transient if there exists some node j for which $i \to j$ but $j \not\to i$, and node i is called recurrent. Let T denote the set of all transient nodes, and let $R = V - T$ be the set of all recurrent nodes. Thus, in the theory of directed graphs, an irreducible Markov chain M is simply one whose digraph G is strongly connected. In general, the communicating classes of M are just the maximal strongly connected subgraphs of G, the strong components of G. It is known that if nodes i and j lie on a common circuit, then they belong to the same strong component, and conversely. As a consequence, when the vertices within each strong component are combined into a new vertex say, *supervertex*, then the condensed graph G? governing these supervertices can contain no cycles, and hence, G? is an acyclic graph. Hence, the state transition diagram G of the finite chain M allows one to classify each vertex as either recurrent or transient. In addition, these concepts are important in connection with stationary distributions $\pi = [\pi_1, \pi_2, \ldots, \pi_n]$ for a Markov chain having states $N = \{1, 2, \ldots, n\}$. These probabilities represent the long run proportion of time the chain M spends in each state. Such probabilities can be found by solving the linear system $\pi = \pi P$, $\sum_{j \in N} \pi_j = 1$.

Remark 4. If G is strongly connected then there is a unique stationary distribution p for M. Moreover, this distribution satisfies $\pi_j > 0$ for all $j \in V$. If the condensed graph G for M has a single supervertex with no leaving edges then, there is a unique stationary distribution π for M. Moreover, this distribution satisfies $\pi_j > 0$ for all $j \in \mathcal{R}$, and $\pi_j = 0$ for all $j \in T$. One may also note that the vertices in strong component G_i consist of those vertices that are descendants of r_i but are not in G_k, $1 \leq k < i$.

Remark 5. M has period d if and only if its digraph G can be partitioned into d sets $C_0, C_1, \ldots, C_{d-1}$ such that (a) if $i \in C_k$ and $(i, j) \in E$ then $j \in C_{(k+1) mod d}$ and (b) d is the largest integer with this property. Define a relation on the nodes of the strongly connected graph G associated with M, assumed to have period d. Namely, define $i \sim j$ if every (i, j)-path has length 0 *modulo d*.

Also since, since G is strongly connected, a path exists between every pair of nodes in G and hence, the relation $i \sim j$ is an equivalence relation. Hence, for every pair of nodes i and j in G, all (i, j)-paths in G have the same length modulo d. Hence, let Q_1 and Q_2 be (i, j_1) and (i, j_2)-paths in G, respectively. If $l(Q_1) \quad mod \quad d = l(Q_2) \quad mod \quad d$, then j_1 and j_2 are elements of the same equivalence class. This establishes the characterization for the period d of an irreducible Markov chain. Consequently, determining d can be reduced to finding an appropriate partition of the nodes in the associated digraph G, which can be accomplished efficiently in conjunction with a breadth-first search of the digraph G.

Consider an unweighted, undirected graph with the symmetric adjacency matrix A. Let the associated diagonal matrix of degree values of each of the nodes be denoted by D. It is clear that $P = D^{-1}A$ is clearly a stochastic symmetric matrix. Thus, we can naturally associate a homogeneous DTMC, based on such a matrix. The state transition diagram of such a DTMC is obtained from the original graph by replacing each edge with two directed edges in opposite directions. Given a simple graph G with M vertices, its Laplacian matrix L is defined as $L = D - A$, where D is the diagonal degree matrix and A is the symmetric adjacency matrix. Now consider the matrix $Q = -L = A - D$. Then, Q satisfies all the properties required by a generator matrix, and thus Q is naturally associated with a homogeneous CTMC. If the transient probability distribution of a homogeneous CTMC, described by the generator matrix Q, such a Q is associated with a state transition rate diagram, which is a directed graph. Then, the transient probability distribution is computed as $\bar{\pi}(t) = \bar{\pi}(0)e^{Qt} = \bar{\pi}(0)e^{(A-D)t}$, where matrix exponential e^{Qt} can be efficiently computed since Q is a symmetric generator matrix. Consider a directed, unweighted graphs. Let D_{in} be the degree matrix associated with edges incident at various nodes and let A_{in} be the adjacency matrix associated with the edges incident at various vertices. Hence, $Q_{in} = A_{in} - D_{in}$ is symmetric, has the properties of a generator matrix, and thus, a homogeneous CTMC can be associated with such a directed, unweighted graph.

Similarly, CTMC / DTMC can be naturally associated with such graphs ignoring the weights on the edges or considering all the edge weights to one. Such an approach can be utilized with weighted graphs that are directed or undirected. More interestingly, a DTMC can be associated with a weighted, undirected graph by normalizing the weight of each edge. That is, consider any vertex, say v_i. Let v_i be connected to vertex v_j. Normalize the weight of such edge by the sum of weights of all edges incident at the vertex v_i. Hence, normalized weight from node v_i to v_j is $\bar{W_{ij}} = \frac{w_{ij}}{\sum_k w_{ik}}$, with $\sum_j \bar{W_{ij}} = 1$, and hence, the matrix of normalized weights is a stochastic matrix. Hence, a DTMC is always associated with a weighted, undirected graph and, the equilibrium and transient probability distribution of such a DTMC can be determined.

4. Fibonacci tree and Markov chain

Following [2], the Fibonacci tree of order k has F_k terminal vertices, where $\{F_k\}$ are the Fibonacci numbers $F_0 = 0, F_1 = 1, F_k = F_{k-1} + F_{k-2}$, and is defined inductively as, if $k = 1$ or 2, the Fibonacci tree of order k is simply the rooted tree. If $k \geq 3$ the left subtree of the Fibonacci tree of order k is the Fibonacci tree of order $k-1$, and the right subtree is the Fibonacci tree of order $k-2$.

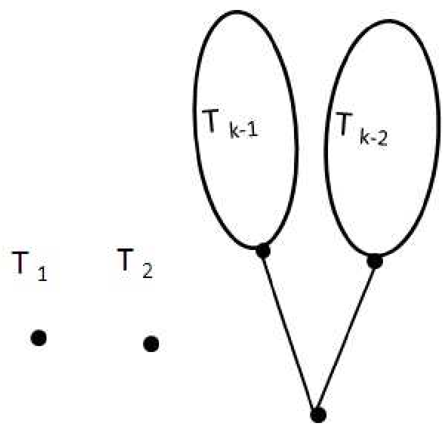

Figure 1: Fibonacci tree

The Fibonacci tree of order k will be denoted by T_k. This definition will serve as a recursive definition for an infinite sequence of trees $T_1, T_2 \ldots, T_k$. Then the tree T_k with $k > 2$ will have F_k leaf-vertices, since the Fibonacci number sequence is defined in the same recursive way, namely $F_k = F_{k-1} + F_{k-2}$ with starting values $F_1 = 1$ and $F_2 = 1$. Notice that successive numbers of leaf-vertices in the trees, after T_1 and T_2, are $2, 3, 5, 8, \ldots$ which are Fibonacci numbers. Note also that each tree after the second is constructed by mounting the two previous trees on a fork as in Figure 1. If there are a total of N nodes, any particular node is landed upon with probability $\frac{1}{N}$. By the recursive definition of T_k, the total number N_k of nodes in the k^{th} Fibonacci tree T_k, we see that $N_k = 1 + N_{k-l} + N_{k-2}$. This implies $N_k + 1 = (N_{k-l} + 1) + (N_{k-2} + 1)$, which implies $N_k = 2F_k - 1$, with $N_1 = N_2 = 1$.

One may notice that, in the theory of DTMCs based on doubly stochastic state transition matrix, the equilibrium distribution of DTMC is given by $[\frac{1}{M}, \frac{1}{M}, \frac{1}{M}, \ldots, \frac{1}{M},]$, where the matrix P is of dimension $M \times M$. That is, the equilibrium distribution achieves maximum entropy possible (uniform distribution).

From [1] it is known the following:

Lemma 6. ([1]) *The transient probability distribution of DTMC associated with an undirected, unweighted graph is given by* $\bar{\pi}(n+1) = \bar{\pi}(0)D^{-n}A^n = \bar{\pi}(0)A^n D^{-n}$.

Thus, the higher powers of adjacency matrix and higher powers of degree matrix play an important role in computing the transient probability distribution. We have, $\bar{\pi}(n+1) = \bar{\pi}(0)p^n = \bar{\mu}(0,n)A^n$. But, since A is symmetric, it has real eigen values and the corresponding eigenvectors are orthonormal. Thus, $A^n = \sum_{i=1}^{N} \rho^n \bar{f}_1 \bar{g}_1$.

Each stage of the rooted Fibonacci tree may be considered as a finite one dimensional random walk with equal probabilities of one step to the right or two step to the left except near the endpoints. There is also a probability of a movement directly to zero. We now construct a condensed Fibinacci tree whose markov chain or equivalently, the corresponding digraph, can be explained as follows. The states of the Markov chain as $\{T_1, T_2, T_3, \ldots, T_n\}$ as the vertex set of the digraph G. That is, $V(G) = \{T_1, T_2, T_3, \ldots, T_n\}$ and the arc set $D(G)$ of G is, $D(G) = \{(T_1, T_2)\} \bigcup \{(T_2, T_1)\} \bigcup \{(T_i, T_i)\} \bigcup \{(T_i, T_j)/\ i < j, j = i+1\} \bigcup \{(T_j, T_i)/j > i, j = i+2\} \bigcup \{(T_n, T_n)\}$.

The respective transition matrix could be considered as

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}.$$

So this Markov chain could represent a population with limited capacity which increases with a birth, decreases with a death, and allows a mass migration of everyone out of the present location, which would lower the population to zero. There is a Fibonacci connection to the Fibinacci tree Markov matrix. We can find the limiting probability (steady state) vector for this Markov chain and show that it has Fibonacci type entries.

Lemma 7. *The Markov chain of condensed Fibonacci tree graph is irreducible and ergotic.*

Proof. The proof follows since the irreducibility of the the Markov chain is the property of the associated digraph G_M, which is strongly connected. Since, G_M has no leaves and each vertex has a unique neighbour set, G_M is irreducible and hence, the Markov Chain is ergodic. $\square$

We have the following result from [1].

Theorem 8. ([1]) *An ergodic Markov chain has a unique stationary distribution.*

Invoking Theorem 8 we get the following:

Lemma 9. *An ergodic Markov chain of condensed Fibonacci tree graph has a unique stationary distribution*

The following theorem gives the limiting probability vector for the transition matrix of the condensed Fibonacci tree.

Theorem 10. *The matrix*
$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & \dots & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & \dots & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} & \dots & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \vdots & \vdots & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & \vdots & \vdots & 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & \dots & \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}$$
is the $n \times n$ transition matrix for the condensed Fibonacci tree for a discrete time Markov chain. Also the limiting probability is

$$\underline{\pi} = (\frac{F_n}{F_{n+2}-1}, \frac{F_{n-1}}{F_{n+2}-1}, \frac{F_{n-2}}{F_{n+2}-1}, \dots, \frac{F_1}{F_{n+2}-1}).$$

Proof. From the construction of the digraph of the condensed Fibonacci tree with the states of the Markov chain as $\{T_1, T_2, T_3, \dots, T_n\}$ as the vertex set and the arc set $D(G)$ of G as $D(G) = \{(T_1, T_2)\} \bigcup \{(T_2, T_1)\} \bigcup \{(T_i, T_i)\} \bigcup \{(T_i, T_j)/$

$i < j, j = i+1\} \bigcup \{(T_j, T_i)/j > i, j = i+2\} \bigcup \{(T_n, T_n)\}$. Clearly, the matrix P is the transition matrix. The balance equation for state $n-1$ comes from the last column, which represents the initial condition for Fibonacci recursion. Thus, $\pi_{n-1} = b\pi_{n-2} + (1-b)\pi_{n-1}$, which implies, $\pi_{n-2} = \pi_{n-1}$. The balance equation for the state $n-2$ comes from the second last column and hence, $\pi_{n-3} = 2\pi_{n-2} = 2\pi_{n-1}$. The balance equations for states $1, 2, 3, \ldots, n-3$ come from columns $2, 3, \ldots, n-2$ and are of the type, $\pi_i = \frac{1}{2}\pi_{i-1} + \frac{1}{2}\pi_{i+2}$. This implies, $\pi_{i+2} = 2\pi - \pi_{i-1}$. Since, Fibonacci numbers satisfy the equation $F_{i+1} = 2F_1 + 1 - F_{1+2}$ for all, i, we conclude that, π's give the Fibonacci ratio in the reverse order. The initial conditions are implied by the final two columns. But, $F_1 + F_2 + \ldots + F_n = F_{n+2} - 1$ and the limiting probabilities must sum to 1, we get the limiting probability form as $\underline{\pi} = (\frac{F_n}{F_{n+2}-1}, \frac{F_{n-1}}{F_{n+2}-1}, \frac{F_{n-2}}{F_{n+2}-1}, \ldots, \frac{F_1}{F_{n+2}-1})$. □

5. Conclusion

Since the recursion relationship for the Lucas numbers is the same as that of the Fibonacci numbers, we should be able to construct the condensed Lucas tree and get the corresponding probability transition matrix by modifying the last two columns of P. Hence, one can find a whole class of probability transition matrices with the same limiting probability vector by using similar methods which convert probability transition matrices for discrete time Markov chains into infinitesimal generators for continuous time Markov processes and vice-versa.

References

[1] X. Cao, J. Zhang, Event-based optimization of Markov systems, *IEEE Trans Automatic Control*, **53**, No 4 (2008), 1076–1082.

[2] Y. Horibe, An entropy view of Fibonacci trees, *The Fibonacci Quarterly*, **20**, No 2 (1982), 168–178.

[3] F. Spitzer, *Principles of Random Walk*, Springer, New York (1964).