Lecture 6: Discrete Time Homogenous Markov Chains, IV:
Mean Time to Absorption

John MacLaren Walsh, Ph.D.
February 19, 2014

1 References

Markov chains are a widely taught subject, and hence there are a wide variety of texts. A few that cover this material at a level appropriate for this course are:


The exposition below follows the first reference (which the bookstore has copies of). The section numbers below are reading from this text.

2 Probabilities of Remaining in Transient States for $n$ Transitions

Let the set of transient states of a discrete time homogeneous Markov chain be $T$. From the previous lecture, we know that transient states are not accessible from persistent states, since states which are accessible from persistent states must also be persistent.

In this lecture, we will study the probability that and mean amount of time required for a chain initialized in a transient state is absorbed into a persistent state (thereafter only making transitions between the states in the communicating class that that persistent state is in).

In this vein, let’s begin by defining the probabilities for each transient state $i \in T$

$$\sigma_i^{(n)} = P[X[n] \in T | X[0] = i]$$

(1)

Note that if $X[0] \in T$ and $X[n] \in T$, then $X[n-i] \in T$ for all $i \in \{1, \ldots, n-1\}$, since once a chain enters a persistent state, it must remain forever in the communicating class of persistent state the persistent state it entered is a member of. In other words, if we reorder the states so that all of the persistent states correspond to the last rows and columns of the probability transition matrix $P$, it will take the block form

$$P = \begin{bmatrix} A & 0 \\ S & W \end{bmatrix}$$

(2)

where the $|T| \times |T|$ matrix $[W]_{i,j} = p_{i,j}$ for all $i,j \in T$ is associated with the transitions among transient states. The $n$th order transition probabilities $[P^n]_{i,j}$ between transient states $i, j \in T$ then only involve transitions among transient states, so $[P^n]_{i,j} = [W^n]_{i,j}$.
We can express the probabilities $\sigma_i^{(n)}$ from (1) in terms of this transition distribution between transient states via

$$\sigma_i^{(n)} = \sum_{j \in T} w_{i,j}^{(n)}$$

which is just the sum of the $i$th row of $W^n$. Furthermore, since

$$w_{i,j}^{(n+1)} = \sum_{k \in T} w_{i,k} w_{k,j}^{(n)}$$

$$\sigma_i^{(n+1)} = \sum_{j \in T} w_{i,j}^{(n+1)} = \sum_{j \in T} \sum_{k \in T} w_{i,k} w_{k,j}^{(n)} = \sum_{k \in T} w_{i,k} \sigma_k^{(n)}$$

which can be expressed in a finite chain as

$$\sigma^{(n+1)} = W^{(n)} \sigma^{(n)}$$

Furthermore, we observe that the sequence $\sigma_i^{(n)}$ is monotone non-increasing, since

$$\sigma_i^{(n+1)} = \sum_{j \in T} w_{i,j}^{(n+1)} = \sum_{j \in T} \sum_{k \in T} w_{i,k} w_{k,j}^{(n)} = \sum_{k \in T} w_{i,k} \sum_{j \in T} w_{k,j} \leq \sum_{k \in T} w_{i,k} = \sigma_i^{(n)}$$

Since $\{\sigma_i^{(n)}\}$ is a monotone non-increasing sequence of probabilities, which are bounded below by 0, $\sigma_i^{(n)}$ must converge to a limit as $n \to \infty$. Letting $\sigma_i^* = \lim_{n \to \infty} \sigma_i^{(n)}$ be these limits, we observe

$$\sigma_i^* = \sum_{k \in T} w_{i,k} \sigma_k^*$$

When the chain is finite, the only solution for $\sigma_i^*$ is $\sigma_i^* = 0 \ \forall i \in T$, since we have already proven that for a transient state $j$

$$\lim_{n \to \infty} P_{i,j}^{(n)} = 0$$

so that

$$\lim_{n \to \infty} \sigma_i^{(n)} = \lim_{n \to \infty} \sum_{j \in T} w_{i,j}^{(n)} = \sum_{j \in T} 0 = 0$$

where the latter two inequalities hold if there are a finite number of transient states, but may not hold if the set of transient states is infinite. If there are an infinite number of transient states, it is possible that there exists solutions to (8) other than just $\sigma_i^* = 0$, and in this scenario there is a positive probability that the chain stays in transient states forever.

### 3 Mean Time to Absorption

Now, consider a chain with a finite number of transient states, and given $X[0] = i$, let $T_i := \min_{t \in \{1,2,...\}} X[t] \notin T$, i.e. let $T_i$ be the first time that the chain exits the transient states. Observe then that

$$\sigma_i^{(n)} = P[T_i > n | X[0] = i]$$

so that

$$\sum_{n=0}^{\infty} \sigma_i^{(n)} = \sigma_i^{(0)} + \sum_{n=1}^{\infty} P[T_i > n | X[0] = i] = 1 + \sum_{n=1}^{\infty} \sum_{k \geq n} P[T_i = k | X[0] = i] = 1 + \sum_{k=1}^{\infty} \sum_{n<k} P[T_i = k | X[0] = i]$$

$$= 1 + \sum_{k=1}^{\infty} (k-1)P[T_i = k | X[0] = i] = \sum_{k=1}^{\infty} kP[T_i = k | X[0] = i] = \sum_{k=1}^{\infty} k E[T_i | X[0] = i]$$

We will denote this quantity by $m_i = \sum_{n=0}^{\infty} \sigma_i^{(n)} = E[T_i | X[0] = i]$, which is the mean number of transitions before the chain exits the transient states, given that it started in transient state $i$. 

2
Recalling the definition of $\sigma_i^{(n)}$ we observe that

$$m_i = \sum_{n=0}^{\infty} \sum_{j \in T} w_{i,j}^{(n)} = \sum_{n=0}^{\infty} \sum_{j \in T} [W^N]_{i,j} = \sum_{n=0}^{\infty} [W^n]_i = \left[ \left( \sum_{n=0}^{\infty} W^n \right) 1 \right]_i = [(I - W)^{-1}]_i$$

so that if we form the vector $m = [m_i | i \in T]$ of mean times to absorption given one starts at the various transient states, we can obtain it as

$$m = (I - W)^{-1} 1.$$  \hspace{1cm} (14)

4 Recurrent Class Absorption Probabilities

Given that a chain starts in a transient state, if it exits the class of transient states (which will happen with probability one in a finite chain), it enters into a communicating class of recurrent/persistent states, then stays there. In particular, let the collection of disjoint communicating classes of recurrent/persistent states be $C_p, p \in \mathcal{P}$.

Because we have shown that for any aperiodic non-null persistent state $j$ and for any state $i$

$$\lim_{n \to \infty} p_{i,j}^{(n)} = \frac{f_{i,j}}{\mu_j}$$

and because, by the Chapman-Kolmogorov equations

$$p_{i,j}^{(n+1)} = \sum_k p_{i,k} p_{k,j}^{(n)}$$

we can show by taking the $\lim_{n \to \infty}$ of (16)

$$f_{i,j} = \sum_k p_{i,k} f_{k,j} \mu_j$$

which is true if and only if

$$f_{i,j} = \sum_k p_{i,k} f_{k,j}$$

Now, states in two different persistent communicating classes are not reachable from one another in either direction, for otherwise they would communicate with one another and be in the same communicating class. Furthermore, if $k, j \in C_p$, then $f_{k,j} = 1$.\(^1\) Hence if $j \in C_p$, the only $f_{k,j} > 0$ are those for which $k \in C_p$ or $k \in T$ the class of all transient states, so that we can rewrite (18) as

$$f_{i,j} = \sum_{k \in T} p_{i,k} f_{k,j} + \sum_{k \in C_p} p_{i,k}$$

Further defining

$$\alpha^{(p)}_i = \sum_{k \in C_p} p_{i,k}$$

we can rewrite (19) as

$$f_{i,j} = \sum_{k \in T} p_{i,k} f_{k,j} + \alpha^{(p)}_i$$

for any $j \in C_p$ a communicating class of recurrent states. If we create the vector $f_j = [f_{i,j} | i \in T]$ and the vector $\alpha^{(p)}$, we observe that (21) can be rewritten as

$$f_j = W f_j + \alpha^{(p)}$$

\(^1\)To see why, let $\alpha$ be the conditional probability that, given a chain starts in state $i$, it travels to state $j$ without first returning to $i$ again. Since $i$ and $j$ communicate, this must be positive. Then, since the probability of not returning to $i$ given you are currently in $j$ is $1 - f_{j,i}$, the probability of not returning to $i$ $1 - f_{i,i} \geq \alpha(1 - f_{j,i})$. But, since $i$ is a persistent state $1 - f_{i,i} = 0$ so $0 \geq \alpha(1 - f_{j,i})$ and since $\alpha > 0$, we must have $f_{j,i} = 1$. 

3
so that

\[ f_j = (I - W)^{-1} \alpha^{(p)} \quad (23) \]

Now, observe that the right hand side of this equation (23) is the same for any \( j \in C_p \), so that \( f_{i,j} = \beta_i^{(p)} \) for all \( j \in C_p \). Thus, we can rewrite (19) as

\[ \beta_i^{(p)} = \sum_{k \in T} p_{i,k} \beta_k^{(p)} + \alpha_i^{(p)} \quad (24) \]

which will, like any linear system of equations, have a unique solution if and only if the homogeneous part

\[ \beta_i^{(p)} = \sum_{k \in T} p_{i,k} \beta_k^{(p)} \quad (25) \]

has as its only solution \( \beta_i^{(p)} = 0 \). Recognizing (25) as (4), we observe that this will be the only solution for a finite chain. Furthermore, for a finite chain, stacking these into the column vectors \( \beta^{(p)} = [\beta_i^{(p)} | i \in T] \) and \( \alpha^{(p)} = [\alpha_i^{(p)} | i \in T] \), this equation becomes

\[ \beta^{(p)} = W \beta^{(p)} + \alpha^{(p)} \quad (26) \]

This yields the solution for the absorption probabilities \( \beta_i^{(p)} \) given by

\[ \beta^{(p)} = (I - W)^{-1} \alpha^{(p)} \quad (27) \]

in which we can obtain \( \alpha^{(p)} \) from the first order transition matrix using (20).