Summary of Markov Chain facts

23 February 2009

1 Classification of states

For a Markov chain $\{X_n : n = 0, 1, 2, \dots\}$ with a (countable or finite) state space $\mathcal{S}$ define

$$T_i = \inf\{n \geq 1 : X_n = i\} \qquad \text{and} \qquad N_i = \sum\nolimits_{n \in \mathbb{N}} \mathbf{1}\{X_n = i\}.$$

Note that $\mathbb{E}_\mu N_i = \sum_{n \in \mathbb{N}} \mathbb{P}_\mu\{X_n = i\}$.

Define $i \rightsquigarrow j$ to mean $\mathbb{P}_i\{T_j < \infty\} > 0$ and $i \leftrightsquigarrow j$ to mean that both $i \rightsquigarrow j$ and $j \rightsquigarrow i$.

Suppose the chain has transition probabilities $P(i, j)$. A ***stationary measure*** is a set of nonnegative numbers $\{\lambda_i : i \in \mathcal{S}\}$ for which

$$\lambda_k = \sum\nolimits_{j \in \mathcal{S}} \lambda_j P(j, k) \qquad \text{for each } k \in \mathcal{S}.$$

If, in addition, $\sum_{i \in \mathcal{S}} \lambda_i = 1$ then the λ_i's are called a ***stationary probability distribution***. For such a distribution,

$$\mathbb{P}_\lambda\{X_n = i\} = \lambda_i \qquad \text{for all } n \in \mathbb{N} \text{ and all } i \in \mathcal{S}.$$

If $\sum_{i \in \mathcal{S}} \lambda_i < \infty$ then a stationary measure can be rescaled to define a stationary probability distribution. Sometimes there are stationary measures that cannot be standardized (cf.).

1.1 Transience

A state i is said to be ***transient*** if $\mathbb{P}_i\{T_i < \infty\} < 1$.

(i) For a transient state, $\mathbb{E}_i N_i < \infty$ and $\mathbb{P}_i\{N_i < \infty\} = 1$.

(ii) If $i \leftrightsquigarrow j$ and i is transient then so is j.

1.2 Recurrence

A state i is said to be ***recurrent*** if $\mathbb{P}_i\{T_i < \infty\} = 1$.

(i) For a recurrent state, $\mathbb{P}_i\{N_i = \infty\} = 1$ and $\mathbb{E}_i N_i = \infty$.

(ii) If $i \leftrightsquigarrow j$ and i is recurrent then so is j.

(iii) If $\{\pi_j : j \in \mathcal{S}\}$ is a stationary probability distribution with $\pi_i > 0$ then the state i is recurrent. [In fact, i is positive recurrent—see Chang Notes Theorem 1.41.]

(iv) If i is recurrent then $\lambda_j := \mathbb{E}_i\{\text{\# visits to state } j \text{ up to time } T_i\}$ defines a stationary measure with $\lambda_i = 1$. [See Section 3.]

1.3 Null Recurrence

A recurrent state i is said to be ***null recurrent*** if $\mathbb{E}_i T_i = \infty$.

1.4 Positive Recurrence

A recurrent state i is said to be ***positive recurrent*** if $\mathbb{E}_i T_i < \infty$.

(i) If i is a positive recurrent state then there exists a stationary probability distribution with $\pi_i = 1/\mathbb{E}_i T_i$. [See Section 3.] If the state space is irreducible then the stationary distribution is unique.

(ii) If $i \leftrightsquigarrow j$ and i is positive recurrent then so is j. [See Section 2.]

1.5 Periodicity

The period of a state i is defined as the greatest common divisor of

$$\{n \in \mathbb{N} : \mathbb{P}_i\{X_n = i\} > 0\}$$

(i) If $i \leftrightsquigarrow j$ then i and j have the same period. [cf. Chang Notes page 16.]

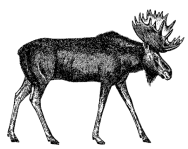
(ii) If state i has period d then $\mathbb{P}_i\{X_{md} = i\} > 0$ for large enough integers m. [cf. Chang Notes page 30.]

1.6 Basic limit theorem

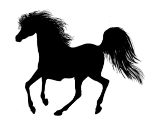
If a chain is irreducible, positive recurrent, and aperiodic then

(i) There exists a unique stationary probability distribution π.

(ii) For every initial distribution μ,



$$\sum_{i\in\mathcal{S}} |\mathbb{P}_\mu\{X_n = i\} - \pi_i| \to 0 \qquad \text{as } n \to \infty.$$



2 Proof of assertion 1.4(ii)

Suppose that state i is positive recurrent and the chain starts in state i. We are given that $\gamma := \mathbb{E}_i T_i < \infty$. We need to show that $\mathbb{E}_j T_j < \infty$.

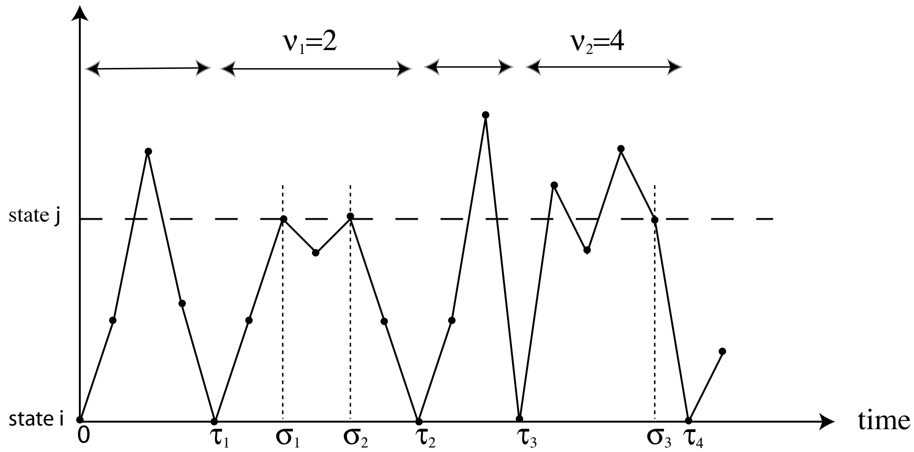
Write $\tau_0 = 0 < \tau_1 < \tau_2 < \dots$ for the times at which the chain is in state i. The first cycle starts at time 1 and ends at time τ_1. The second cycle starts at time $\tau_1 + 1$ and ends at time τ_2. And so on. We could also write each τ_k as a sum $T_i^{(1)} + T_i^{(2)} + \dots + T_i^{(k)}$, where $T_i^{(m)}$ denotes the length of the mth cycle. The random variables $T_i^{(1)}, T_i^{(2)}, \dots$ are independent and identically distributed, each with expected value γ. [The variable $T_i^{(1)}$ is the same as the variable T_i used to define γ.]

Similarly, define $1 \le \sigma_1 < \sigma_2 < \dots$ for the times at which the chain visits state j. Note that σ_1 is what we have also been denoting by T_j.

More subtly, define random variables that pick out the cycles where there are visits to state j:

$$\begin{aligned}
\nu_1 &= \text{ first } m \text{ for which chain visits } j \text{ at least once during } m\text{th cycle}\\
\nu_2 &= \text{ first } m > \nu_1 \text{ for which chain visits } j \text{ at lest once during } m\text{th cycle}
\end{aligned}$$

For the history shown in the picture, note that $\nu_2 = 4$ even though the second visit to state j occurs during the second cycle.



The events $\{$chain visits j during mth cycle$\}$ are independent, each with the same $\mathbb{P}_i$ probability θ. We must have $\theta > 0$, otherwise the chain could never visit j, in violation of the assumption that $i \rightsquigarrow j$. Consequently, ν_1 has a geometric(θ) distribution,

$$\mathbb{P}_i\{\nu_1 = m\} = (1-\theta)^{m-1}\theta \qquad \text{for } m \in \mathbb{N},$$

with expected value $\mathbb{E}_i\nu_1 = 1/\theta$. Similarly, ν_2 is the sum of two independent random variables, each distributed geometric(θ), with expected value $\mathbb{E}_i\nu_2 = 2/\theta$.

The key idea is that during cycles $1, 2, \dots, \nu_2$ there must be at least two visits to state j. That is, we must have $\sigma_2 \leq \tau_{\nu_2}$. Moreover, between times σ_1 and σ_2 the chain makes an excursion that starts and ends in state j. We can hope that a bound on $\mathbb{E}_i\tau_{\nu_2}$ will provide a bound on $\mathbb{E}_jT_j$.

To get the bound on $\mathbb{E}_i\tau_{\nu_2}$, first note that

$$\tau_{\nu_2} = \sum\nolimits_{k\in\mathbb{N}} T_i^{(k)}\mathbf{1}\{k \leq \nu_2\}$$

You should check that if $\nu_2 = m$ then the sum on the right-hand side reduces to $T_i^{(1)} + T_i^{(2)} + \cdots + T_i^{(m)} = \tau_m$. Take expectations then condition.

$$\mathbb{E}_i\tau_{\nu_2} = \sum\nolimits_{k\in\mathbb{N}} \mathbb{E}_i(T_i^{(k)} \mid k \leq \nu_2)\mathbb{P}_i\{k \leq \nu_2\}.$$

The information $k \leq \nu_2$ tells us precisely that among cycles $1, 2, \dots, k-1$ there has been at most one during which there was one or more visits to state j; The event $\{k \leq \nu_2\}$ gives only information about the first $k-1$ cycles, whereas $T_i^{(k)}$ is the length of the kth cycle. Independence of what happens from one cycle to the next therefore lets us discard the conditoning information, leaving

$$\mathbb{E}_i(T_i^{(k)} \mid k \leq \nu_2) = \mathbb{E}_i(T_i^{(k)}) = \gamma.$$

The expression for $\mathbb{E}_i\tau_{\nu_2}$ simplifies to

$$\gamma \sum\nolimits_{k\in\mathbb{N}} \mathbb{P}_i\{k \le \nu_2\} = \gamma\mathbb{E}_i(\sum\nolimits_{k\in\mathbb{N}} \mathbf{1}\{k \le \nu_2\}) = \gamma\mathbb{E}_i\nu_2 = 2\gamma/\theta < \infty,$$

because $\sum_{k\in\mathbb{N}} \mathbf{1}\{k \le \nu_2\} = \nu_2$. Thus $\mathbb{E}_i\sigma_2 \le \mathbb{E}_i\tau_{\nu_2} < \infty$.

Now we have only to extract the excursion from j back to j by conditioning on the value of σ_1, which we can also write as T_j.

$$\mathbb{E}_i\sigma_2 = \sum\nolimits_{n\in\mathbb{N}} \mathbb{P}_i\{T_j = n\}\mathbb{E}_i(\sigma_2 \mid T_j = n).$$

Notice that we don't need a contribution from $T_j = \infty$ because

$$\mathbb{P}_i\{T_j = \infty\} = \mathbb{P}_i\{\text{chain never visits } j\ \} = 0.$$

If you write $\mathbb{E}_i(\sigma_2 \mid T_j = n)$ as

$$\mathbb{E}_i(\sigma_2 \mid X_1 \ne j,\ X_2 \ne j, \dots, X_{n-1} \ne j,\ X_n = j)$$

you should see, by the Markov property, that $\mathbb{E}_i(\sigma_2 \mid T_j = n) = n + \mathbb{E}_j T_j$. Note how the first n steps contribute to σ_2. Thus

$$\mathbb{E}_i\sigma_2 = \mathbb{E}_j T_j \sum\nolimits_{n\in\mathbb{N}} \mathbb{P}_i\{T_j = n\} + \sum\nolimits_{n\in\mathbb{N}} \mathbb{P}_i\{j = n\}n = \mathbb{E}_j T_j + \mathbb{E}_i T_j.$$

Not only can we now conclude that $\mathbb{E}_j T_j < \infty$, so that state j is positive recurrent, but also that $\mathbb{E}_i T_j < \infty$.

3 Stationary distributions for Markov chains

Suppose i is a recurrent state for a Markov chain. For each j in the state space $\mathcal{S}$ define

$$\begin{aligned}\lambda_j &= \mathbb{E}_i\{\#\text{ visits to state } j \text{ up to time } T_i\} \\ &= \mathbb{E}_i \sum\nolimits_{n\in\mathbb{N}} \mathbf{1}\{X_n = j,\ n \le T_i\} \\ &= \sum\nolimits_{n\in\mathbb{N}} \mathbb{P}_i\{X_n = j,\ n \le T_i\}\end{aligned}$$

Note that $\lambda_i = 1$ because $X_n \ne i$ for $1 \le n < T_i$ and $X_n = i$ when $n = T_i$.

<1> THEOREM. $\sum_{j\in\mathcal{S}} \lambda_j P(j,k) = \lambda_k$ *for each state* k.

PROOF By definition,

$$(*) = \sum_{j\in\mathcal{S}} \lambda_j P(j,k) = \sum_{j\in\mathcal{S}} \sum_{n\in\mathbb{N}} \mathbb{P}_i\{X_n = j,\, n \le T_i\} P(j,k)$$

The summand is equal to

$$\mathbb{P}_i\{X_n = j,\, X_{n+1} = k,\, n \le T_i\}$$

because $\mathbb{P}_i\{X_{n+1} = k \mid X_n = j,\, n \le T_i\} = P(j,k)$. Thus

$$\begin{aligned}(*) &= \sum_{n\in\mathbb{N}} \sum_{j\in\mathcal{S}} \mathbb{P}_i\{X_n = j,\, X_{n+1} = k,\, n \le T_i\} \\ &= \sum_{n\in\mathbb{N}} \mathbb{P}_i\{X_{n+1} = k,\, n \le T_i\}.\end{aligned}$$

Split the last summand into two pieces, corresponding to the decomposition of the event $\{T_i \ge n\}$ into the union of disjoint events $\{T_i = n\}$ and $\{T_i \ge n+1\}$.

Now note that

$$\begin{aligned}&\sum_{n\in\mathbb{N}} \mathbb{P}_i\{X_{n+1} = k,\, T_i = n\} \\ &\qquad = \sum_{n\in\mathbb{N}} \mathbb{P}_i\{T_i = n\} \mathbb{P}_i\{X_{n+1} = k \mid T_i = n\} \\ &\qquad = \left(\sum_{n\in\mathbb{N}} \mathbb{P}_i\{T_i = n\}\right) P(i,k) \\ &\qquad = P(i,k) \qquad \text{because } \mathbb{P}_i\{T_i \ge 1\} = 1.\end{aligned}$$

For the contribution from $\{T_i \ge n+1\}$ replace the variable of summation by $m = n+1$:

$$\begin{aligned}\sum_{n\in\mathbb{N}} \mathbb{P}_i\{X_{n+1} = k,\, T_i \ge n+1\} &= \sum_{m\ge 2} \mathbb{P}_i\{X_m = k,\, T_i \ge m\} \\ &= \lambda_j - \mathbb{P}_i\{X_1 = k, T_i \ge 1\}.\end{aligned}$$

Once again note that $\mathbb{P}_i\{X_1 = k, T_i \ge 1\} = \mathbb{P}_i\{X_1 = k\} = P(i,k)$ to conclude that

$$\begin{aligned}(*) &= \sum_{n\in\mathbb{N}} \mathbb{P}_i\{X_{n+1} = k,\, n \le T_i\} \\ &= \mathbb{P}_i\{X_1 = k,\, T_i \ge 1\} + \sum_{n\ge 2} \mathbb{P}_i\{X_n = k,\, T_i \ge n\} \\ &= \lambda_k.\end{aligned}$$

The Theorem is proved.

□

The stationary measure can be standardized to give a stationary probability distribution if the λ_j's have a finite sum. By definition,

$$\begin{aligned}\sum\nolimits_{j\in\mathbb{S}}\lambda_j &= \mathbb{E}_i\sum\nolimits_{n\in\mathbb{N}}\sum\nolimits_{j\in\mathbb{S}}\mathbf{1}\{X_n=j,\, n\le T_i\}\\ &= \mathbb{E}_i\sum\nolimits_{n\in\mathbb{N}}\mathbf{1}\{n\le T_i\}\\ &= \mathbb{E}_iT_i \qquad \text{because } T_i=\sum\nolimits_{n\in\mathbb{N}}\mathbf{1}\{n\le T_i\}.\end{aligned}$$

If $\mathbb{E}_iT_i<\infty$, it follows that there is a stationary probability distribution defined by

$$\pi_j=\lambda_j/\mathbb{E}_iT_i \qquad \text{for all } j\in\mathbb{S}.$$

In particular, $\pi_i=1/\mathbb{E}_iT_i$.

It might seem that we should have many different stationary distributions for a positive recurrent chain, one for each possible choice of the state i in the preceding construction. However, if the chain is also irreducible and aperiodic, the Basic Limit Theorem ensures that there can be at most one stationary π. It follows in that case that

$$\pi_j=1/\mathbb{E}_jT_j \qquad \text{for every } j \text{ in } \mathbb{S}.$$

Thanks to SC and XY for proofreading.