Math 285A, Spring 2006

Minimum Principle for Markov Chains

Let $X = (X_n)_{n\ge 0}$ be a Markov Chain with state space S (finite or countably infinite) and transition matrix P. Let B be a subset of S, and consider a function $f : S \to [0,\infty)$ that is "superharmonic on B^c" in the sense that

$$f(i) \ge Pf(i) = \sum_{j\in S} p(i,j)f(j), \qquad \forall i \in B^c. \tag{1}$$

Let $\tau := \min\{n \ge 0 : X_n \in B\}$ be the hitting time of B, and let us define a new Markov Chain Y by setting

$$Y_n := X_{n\wedge\tau} = \begin{cases} X_n, & n < \tau, \\ X_\tau, & \tau \le n. \end{cases} \tag{2}$$

Let us compute the transition probabilities $\widetilde{p}(i,j)$ for Y. It is clear that if $i \in B$ and $Y_n = i$, then $X_{n\wedge\tau} = i \in B$, so $n \ge \tau$; from this it follows that $Y_{n+1} = X_{(n+1)\wedge\tau} = X_\tau = i$. That is, each state of B is a trap for Y. Consequently,

$$\widetilde{p}(i,j) = \mathbf{P}[Y_{n+1} = j | Y_n = j] = \begin{cases} 1, & j = i, \\ 0, & j \ne i, \end{cases} \qquad i \in B, j \in S. \tag{3}$$

If $i \in B^c$, then

$$\begin{aligned} \mathbf{P}[Y_{n+1} = j | Y_n = i] &= \frac{\mathbf{P}[X_{(n+1)\wedge\tau} = j, X_n = i, n < \tau]}{\mathbf{P}[X_n = i, n < \tau]} \\ &= \frac{\mathbf{P}[X_{(n+1)} = j, X_n = i, n < \tau]}{\mathbf{P}[X_n = i, n < \tau]} \\ &= \mathbf{P}[X_{(n+1)} = j | X_n = i, n < \tau] \\ &= p(i,j). \end{aligned} \tag{4}$$

(The final equality above holds by the Markov property, because $\{n < \tau\} = \{X_0 \in B^c, X_1 \in B^c, \ldots, X_n \in B^c\}$.) That is, the matrix $\widetilde{P}$ for Y is obtained by replacing the rows of P indexed by B by a matrix of the form $[I|0]$, in which I is an identity matrix whose dimension is the cardinality of B. (I assume the states of S have been partitioned so that those of B are written first and those of B^c second.) It follows from this and (1) that

$$f(i) \ge \widetilde{P}f(i), \qquad \forall i \in S. \tag{5}$$

Applying $\widetilde{P}$ to both sides of (5) repeatedly, we find that

$$f(i) \geq \widetilde{P}^n f(i), \qquad \forall i \in S, n = 1, 2, 3, \ldots. \tag{6}$$

That is,

$$f(i) \geq \mathbf{E}[f(Y_n)|Y_0 = i] = \mathbf{E}[f(X_{n\wedge\tau})|X_0 = i], \qquad \forall i \in S, n = 1, 2, 3, \ldots. \tag{7}$$

Moving everything in (7) to the right side, we have, equivalently,

$$0 \geq \mathbf{E}\left[[f(X_{n\wedge\tau}) - f(i)]|X_0 = i\right], \qquad \forall i \in S, n = 1, 2, 3, \ldots. \tag{8}$$

Suppose now that i is a state at which f attains its minimum value, and that $i \in B^c$. Then $f(X_{n\wedge\tau}) - f(i) \geq 0$, so (8) implies that

$$\mathbf{P}\left[f(X_{n\wedge\tau}) = f(i)|X_0 = i\right] = 1, \qquad \forall i \in S, n = 1, 2, 3, \ldots. \tag{9}$$

We have now, in effect, proved the following result.

Theorem. *Let $f : S \to [0, \infty)$ satisfy (1) for some set $B \subset S$. Suppose the following "irreducibility" conditions hold:*

$$\mathbf{P}[X_n = j, n < \tau|X_0 = i] > 0 \text{ for some } n \geq 1, \qquad \forall i \in B^c, j \in B^c, \tag{10}$$

$$\mathbf{P}[X_\tau = j, \tau < \infty|X_0 = i] > 0, \qquad \forall i \in B^c, j \in B. \tag{11}$$

If f attains its minimum value at some point of B^c, then f is constant on S.

Proof. Assume, as before, that i is a point of B^c at which f attains its minimum value. By (10), if $j \in B^c$,

$$\mathbf{P}[X_n = j, n < \tau|X_0 =] > 0 \tag{12}$$

and then (9) forces $f(j) = f(i)$. Likewise, if $j \in B$ then (11) implies that $\mathbf{P}[X_\tau = j, \tau \leq n] > 0$ for large enough n, and then (9) again implies that $f(j) = f(i)$. Thus $f(j) = f(i)$ for all $j \in S$. □

The case in which B is empty is especially simple.

Corollary 1. *Suppose that X is irreducible, and let $f : S \to [0, \infty)$ be superharmonic on all of S. If f attains its minimum value at some point of S, then f is constant on S.*

The following proposition is a useful consequence of the idea behind the theorem.

Proposition. *Let $f : S \to [0, \infty)$ satisfy (1) for some set $B \subset S$. Suppose that*

$$\mathbf{P}[\tau < \infty] > 0, \qquad \forall i \in B^c \tag{13}$$

Then

$$\inf\{f(i) : i \in B\} = \inf\{f(i) : i \in S\}. \tag{14}$$

Proof. Clearly $\inf\{f(i) : i \in B\} \geq \inf\{f(i) : i \in S\}$. If this is a strict equality, then there exists $i \in B^c$ such that $f(i) < f(j)$ for all $j \in B$. In view of the hypothesis (13) (which implies that $\mathbf{P}[X_\tau = j, \tau \leq n | X_0 = i] > 0$ for some $n \geq 1$ and some $j \in B$), we thereby obtain a contradiction of (9). ∎

The following corollary of the proposition was used in class to show that for the simple symmetric random walk on $\{0, 1, 2, \dots, N\}$, the probability of hitting N before 0 (when the walk is started at $i \in \{1, 2, \dots, N-1\}$) is i/N.

Corollary 2. *Let $f : S \to [0, \infty)$ and $g : S \to [0, \infty)$ be "harmonic" on B^c:*

$$Pf(i) = f(i), Pg(i) = g(i), \qquad \forall i \in B^c, \tag{15}$$

and suppose that $\mathbf{P}[\tau < \infty | X_0 = i] > 0$ for all $i \in B^c$. If $f(j) = g(j)$ for all $j \in B$, then $f(j) = g(j)$ for all $j \in S$.