Section 6: Harris Recurrence

Contents

6.1 Harris Recurrent Markov Chains

We have previously developed a fairly complete steady-state theory for discrete state space Markov chains. In particular, we established that when $X = (X_n : n \geq 0)$ is an irreducible Markov chain on discrete state space S, then X has a steady-state (in the sense of the law of large numbers) precisely when X has a stationary distribution. One problem with this theory is that we cannot establish stability without essentially solving the linear system $\pi = \pi P$.

We now wish to develop recurrence theory that permits us to establish stability even for models for which the stationary equations cannot be explicitly solved. Ideally, this theory will also permit us to verify stability for Markov chains evolving in non-compact state spaces.

Definition 6.1.1 A Markov chain $X = (X_n : n \geq 0)$ with state space S is said to be *recurrent in the sense of Harris* if there exists $A \subseteq S$, $\lambda > 0$, $m \geq 1$, and a distribution φ on S such that

i.) $P_x(T_A < \infty) = 1$, $x \in S$, where $T_A = \inf\{n \geq 0 : X_n \in A\}$;

ii.) $P_x(X_m \in \cdot) \geq \lambda\, \varphi(\cdot)$, $x \in A$.

Remark 6.1.1 If S is discrete, then X is Harris recurrent if and only if there exists $z \in S$ such that $P_x(\tau(z) < \infty) = 1$ for $x \in S$, where $\tau(z) = \inf\{n \geq 1 : X_n = z\}$.

Note that condition i.) guarantees that X visits A infinitely often a.s. Using the same coin flip idea, condition ii.) ensures that every time X visits A (not having visited A in the previous m time steps), there is a probability of λ of a successful coin toss (i.e. a "heads"). It follow that there will be infinitely many times $T_1, T_2, \cdots$ at which X has distribution φ. Each such random time T_i initiates a cycle having an identical distribution. Furthermore, $X_{T_i - m}$ is independent of X_{T_i}. Hence, if $m = 1$, the resulting cycles are iid, where if $m > 1$, the cycles are correlated. However, then degree of correlation is modest, and the following theorem is easily established.

Theorem 6.1.1 Let $X = (X_n : n \geq 0)$ is Harris recurrent. Put $\tau_i = T_i - T_{i-1}$. If $\mathrm{E}\tau_1 < \infty$, then X possesses a unique stationary distribution π and for each non-negative f,

$$\frac{1}{n}\sum_{i=0}^{n-1} f(X_1) \to \int_S \pi(dx) f(x) \quad \text{a.s}$$

as $n \to \infty$.

Remark 6.1.2 A sufficient condition guaranteeing $\mathrm{E}\tau_1 < \infty$ is to require that

$$\sup_{x\in A} \mathrm{E}_x \tau_A < \infty, \tag{6.1.1}$$

where $\tau_A = \inf\{n \geq 1 : X_n \in A\}$.

To illustrate what is involved in verifying condition ii.), suppose that

$$P_x(X_m \in B) = \int_B p_m(x,y)\xi(dy) \tag{6.1.2}$$

for some distribution $z(\cdot)$ and transition density p_m. Suppose that

$$\varphi(B) = \int_B \phi(y)\xi(dy)$$

for some density $\phi(\cdot)$. (The Radon-Nikodym theorem actually guarantees that φ must take this form in the presence of condition ii.).) If $p_m(\cdot, y)$ is continuous and positive, with A compact, then we can take

$$\phi(y) = \inf_{x\in A} p_m(x,y),$$

and

$$\lambda = \int_S \inf_{x\in A} p_m(x,z)\xi(dz).$$

Hence in some generality, it follows that any compact set A satisfies condition ii.). (Caution: The above analysis assumes that the m-step transition probabilities $P_x(X_m \in \cdot)$ can be represented as in (6.1.2), with a transition density that is positive and continuous. This must be verified separately for each example. If this fails to be true, one must verify condition ii.) from "first principles".)

6.2 Stochastic Lyapunov Functions

It remains to provide a technique for verifying condition i.) and condition (6.1.1).

Proposition 6.2.1 Let $w : S \to [1,\infty)$ for which there exists $r < 1$ such that

$$\mathrm{E}_x w(X_1) \leq r w(x)$$

for $x \in A^c$. Then,

$$\mathrm{E}_x \tau_A \leq (1-r)^{-1} w(x)$$

for $x \in A^c$.

Proof: Via use of operator/function norms with weight function $w(\cdot)$, we conclude that

$$\mathrm{E}_x \sum_{j=0}^{T_A - 1} w(X_j) \leq (1-r)^{-1} w(x)$$

for $x \in A^c$. But since $w(x) \geq 1$ for $x \in S$,

$$\tau_A \leq \sum_{j=0}^{\tau_A - 1} w(X_j),$$

yielding the result. □

Using "first transition" analysis, we see that for $x \in A$,

$$\mathrm{E}_x \tau_A = 1 + \int_{A^c} P(x, dy) \mathrm{E}_y \tau_A.$$

In view of Proposition 6.2.1, we conclude that

$$\sup_{x \in A} \mathrm{E}_x \tau_A \leq 1 + (1-r)^{-1} \sup_{x \in A} \mathrm{E}_x w(X_1).$$

Proposition 6.2.2 Suppose that there exists $A \subseteq S$, $X > 0$, $m \geq 1$, $\lambda > 0$, a distribution φ, $r < 1$, and $w : S \to [1, \infty)$ such that:

i.)' $\mathrm{E}_x w(X_1) \leq r w(x)$, $x \in A^c$;

ii.)' $\sup_{x \in A} \mathrm{E}_x w(X_1) < \infty$;

iii.)' $P_x(X_m \in \cdot) \geq \lambda \varphi(\cdot)$, $x \in A$.

Then, X possesses a unique stationary distribution π and for each non-negative f,

$$\frac{1}{n} \sum_{i=0}^{n-1} f(X_i) \to \int_S \pi(dx) f(x) \quad \text{a.s}$$

as $n \to \infty$.

Remark 6.2.1 The function $w(\cdot)$ is called a *stochastic Lyapunov function.*

Example 6.2.1 Suppose that $(S_n : n \geq 0)$ is a sequence of rv's describing reservoir storage in the presence of a linear release rule, so that

$$S_{n+1} = S_n + Z_{n+1} - a S_{n+1}$$

for $a > 0$. Assume that $(Z_n : n \geq 1)$ is a sequence of iid non-negative rv's for which $\mathrm{E} Z_1 < \infty$. Put $w(x) = 1 + x$ for $x \geq 0$. Then,

$$\begin{aligned} \mathrm{E}_x w(S_1) &= 1 + \mathrm{E}_x S_1 \\ &= 1 + \frac{\mathrm{E}(x + Z_1)}{1+a} \\ &\leq \left(1 + \frac{a}{2}\right)^{-1} w(x) + 1 + \frac{\mathrm{E} Z_1}{(1+a)} - a x (1+a)^{-1} (2+a)^{-1} \\ &\leq \left(1 + \frac{a}{2}\right)^{-1} w(x) \end{aligned}$$

for $x \geq (1+a)(2+a)/a + \mathrm{E} Z_1 (2+a)/a$. Put $A = [0, (1+a)(2+a)/a + \mathrm{E} Z_1 (2+a)/a]$. Condition ii.)' of Proposition 10.2.2 is easily verified here. Hence, it remains only to show that condition iii.)' is satisfied. But this is straightforward to carry out if we assume that Z_1 has a continuous positive density on $[0, \infty)$.

A weaker Lyapunov condition is offered by our next result.

Theorem 6.2.1 Suppose that there exists $A \in S$, $\lambda > 0$, $m \geq 1$, a distribution φ, $\epsilon > 0$, and $g : S \to [0, \infty)$ such that:

I.) $\mathrm{E}_x g(X_1) \leq g(x) - \epsilon$, $x \in A^c$;

II.) $\sup_{x \in A} \mathrm{E}_x g(X_1) < \infty$;

III.) $P_x(X_m \in \cdot) \geq \lambda \varphi(\cdot)$, $x \in A$.

Then, X possesses a unique stationary distribution π and for each non-negative f,

$$\frac{1}{n} \sum_{i=0}^{n-1} f(X_i) \to \int_S f(x) \pi(dx) \quad \text{a.s.}$$

as $n \to \infty$.

Proof: Let $B = (B(x, dy) : x, y \in A^c)$ be the restriction of $P = (P(x, dy) : x, y \in S)$ to A^c, so that $B(x, dy) = P(x, dy)$ for $x, y \in A^c$. Put $e(x) = 1$ for $x \in A^c$. Note that condition i.) implies that

$$\int_{A^c} B(x, dy) g(y) \;\; = \;\; \int_{A^c} P(x, dy) g(y) \leq \int_S P(x, dy) g(y) = \mathrm{E}_x s(X_1) \leq g(x) - \epsilon e(x),$$

so that

$$Bg \leq g - \epsilon e.$$

Hence,

$$\epsilon e \leq g - Bg.$$

Since B is a non-negative operator, it follows that

$$\epsilon B^j e \leq B^j g - B^{j+1} g.$$

Summing over $j \in \{0, 1, \cdots, n\}$, we get

$$\epsilon \sum_{j=0}^{n} B^j e \leq g - B^{n+1} g \leq g.$$

Sending $n \to \infty$, we conclude that

$$\epsilon \sum_{j=0}^{n} B^j e \leq g.$$

But

$$(B^j e)(x) = P_x(\tau_A > j),$$

so

$$\sum_{i=0}^{\infty} (B^j e)(x) = \sum_{j=0}^{\infty} P_x(\tau_A > j) = \mathrm{E}_x \tau_A,$$

yielding the inequality

$$\mathrm{E}_x \tau_A \leq g(x)/\epsilon$$

for $x \in A^c$. Consequently, $P_x(\tau_A < \infty) = 1$ for $x \in A^c$ (and hence for all $x \in S$). Furthermore,

$$\sup_{x \in A} \mathrm{E}_x \tau_A \leq 1 + \sup_{x \in A} \mathrm{E}_x g(X_1).$$

The conclusion of the theorem therefore follows as for Proposition 6.2.2.

Remark 6.2.2 A nice physical way to think about g is to view $g(x)$ as representing the "potential energy" associated with x. Condition i.) asserts that, in expectation, the potential energy has a tendency to decrease by ϵ on A^c. Hence, since the system wishes to move to points of lower potential energy, it follows that the system should eventually enter A.

Remark 6.2.3 A high level of ingenuity may be required to find a suitable Lyapunov function. One approach to finding a suitable g is to try some candidate functions. When $S \subseteq \mathbb{R}^d$, typical candidates to try are:

a.) $\|x\|^p$;

b.) $\exp(a\|x\|^p)$;

c.) $(\log(1 + \|x\|))^p$.