The drift criterion for recurrence

Let P be an irreducible Markov kernel on $(\mathcal{X}, \mathcal{E})$. Recall that a function $V: \mathcal{X} \to [0,\infty)$ is said to be <u>unbounded off small sets</u> if the sublevel sets

$$\{x \mid V(x) \le c\} \qquad c \in (0,\infty)$$

all are small sets.

<u>Theorem</u> Let P be irreducible. Suppose that there exists a function $V: \mathcal{X} \to [0,\infty)$ which is measurable and unbounded off small sets, and which satisfies that

$$\Delta V(x) \le 0 \qquad \forall x \notin C$$

for some small set $C \in \mathcal{E}^+$. Then P is recurrent.

Proof: Assume that the conclusion is wrong, so that P is transient.

As

$$L_x(C) = P_x(C) + \int_{C^c} L_y(C)\, dP_x(y)$$

we conclude that if

$$(*) \qquad L_y(C) = 1 \qquad \forall\, y \notin C$$

then

$$L_x(C) = 1 \qquad \forall\, x \in C$$

This implies that C is a recurrent $\mathbb{E}^+$-set which is also small. But transience implies that any small $\mathbb{E}^+$-set must be uniformly transient.

Hence we conclude that $(*)$ must be false. In other words: we can find some $x^* \notin C$ such that $L_{x^*}(C) < 1$.

Pick $M > \dfrac{V(x^*)}{1 - L_{x^*}(C)}$

Observe that

$$C_M = \{x \mid V(x) \leq M\}$$

is small. If necessary we can choose an even larger M, and obtain that $C_M \in \mathbb{E}^+$. Combined with the assumption of transience, we have that C_M is uniformly transient.

Define a new Markov kernel $Q = (Q_x)_{x \in \mathcal{X}}$ by

$$Q_x = \begin{cases} P_x & x \notin C \\ \varepsilon_x & x \in C \end{cases}$$

If $X_0, X_1, \ldots$ is a P-Markov Chain and τ is the first hitting time of C, then

$$X_{\tau \wedge 0}, X_{\tau \wedge 1}, X_{\tau \wedge 2}, \ldots$$

is a time homogeneous Markov Chain with onestep transition probability Q.

We note three properties of the Q-kernel

① If $x \notin C$, then

$$\int V(y)\, dQ_x(y) = \int V(y)\, dP_x(y) \leq V(x)$$

If $x \in C$, then

$$\int V(y)\, dQ_x(y) = \int V(y)\, d\varepsilon_x(y) = V(x)$$

so actually

$$\int V(y)\, dQ_x(y) \leq V(x) \qquad \forall x \in \mathcal{X}$$

when the P-kernel only had this property outside C.

② If $x \notin C$ then $Q_x^n(C) \to L_x(C)$ for $n \to \infty$

The easiest proof of this is probably probabilistic: let $X_0, X_1, \ldots$ be a P-chain with $X_0 \equiv x$. Let τ be the first return time to C. Then

$$Q_x^n(C) = P(X_{\tau \wedge n} \in C) = P(\tau \leq n) \to P(\tau < \infty)$$

and this last probability equals $L_x(C)$ per definition.

(3) If $x \notin C$ and $A \subseteq C^c$ we have that

$$Q_x^n(A) \le P_x^n(A)$$

We can again argue probabilistic. Let $X_0, X_1, X_2 \dots$ be a P-chain with $X_0 = x$ and let τ be the first return time to C. Then

$$Q_x^n(A) = P(X_{\tau \wedge n} \in A) = P(X_n \in A, \tau > n)$$
$$\le P(X_n \in A) = P_x^n(A)$$

To finish of the proof and arrive at of contradiction, we iterate on the claim in (1) to get

$$V(x) \ge \int V(y) \, dQ_x^n(y)$$

$$\ge \int_{C^c \cap C_M^c} V(y) \, dQ_x^n(y)$$

$$\ge M \, Q_x^n(C^c \cap C_M^c)$$

$$= M \left(1 - Q_x^n(C \cup C_M)\right)$$

$$= M\left(1 - Q_x^n(C) - Q_x^n(C_M \setminus C)\right)$$

Observe that for $x \notin C$

$$Q_x^n(C_m \setminus C) \leq P_x^n(C_m \setminus C) \leq P_x^n(C_m)$$

As C_m is uniformly transient, there is K s.t.

$$\sum_{n=1}^{\infty} P_x^n(C_m) \leq K \qquad \forall x \in T$$

In particular $P_x^n(C_m) \to 0$.

So for $x \notin C$ we see that

$$Q_x^n(C) \to L_x(C), \quad Q_x^n(C_m \setminus C) \to 0$$

Hence we conclude that

$$V(x) \geq M(1 - L_x(C)) \qquad \forall x \notin C$$

But M is specifically chosen such that this inequality is false for x^*.

This contradiction can only be resolved by refuting the original assumption that P is transient.

□