1 Markov kernel as operator

Suppose we are given a Markov chain \((X_n)\) w.r.t. a filtration \(\mathcal{F}_n\), taking values in a state space \((S, \mathcal{S})\) and an associated transition kernel \(P\). Formally, \(P\) is a function from \(S \times S\) to \(\mathbb{R}\) such that

- for each \(s \in S\), \(A \mapsto P(s, A)\) is a probability measure on \((S, \mathcal{S})\).
- for each \(A \in \mathcal{S}\), \(s \mapsto P(s, A)\) is measurable.
- \(P(X_n, A)\) is a version of \(\mathbb{P}(X_{n+1} \in A | \mathcal{F}_n)\).

\(P\) can be regarded as an operator taking functions on \(S\) to functions on \(S\). Specifically, suppose \(f : S \to \mathbb{R}\), then define

\[(Pf)(x) = \int_S f(y)P(x, dy)\]

This operator gives a succinct way of expressing conditional expectations:

\[E[f(X_{n+1})|X_1, \ldots, X_n] = (Pf)(X_n)\]

To see this, note that this holds if \(f\) is an indicator:

\[E[1(X_{n+1} \in A)|X_1, \ldots, X_n] = \mathbb{P}(X_{n+1} \in A|X_n) = \int_S 1(y \in A)P(X_n, dy)\]

Now by linearity the claim holds for simple functions, and by monotone convergence and its conditional expectation counterpart it holds for nonnegative, bounded measurable functions. By considering positive and negative parts the claim holds for bounded measurable functions.

Consider the fixed points of the operator \(P\); i.e. functions \(f\) satisfying \(Pf = f\), or equivalently, \((P - I)f = 0\). By the expression for conditional expectation above, if \(f\) is a fixed point of \(P\), then \((f(X_n))_{n \geq 1}\) is a martingale relative to \(\mathcal{F}_n\).
Similarly, if \( f \geq Pf \), then \( f(X_n) \) is a supermartingale, and if \( f \leq Pf \), \( f(X_n) \) is a submartingale.

A function \( f \) is said to be subharmonic (relative to \( P \)) if \( Pf \geq f \), superharmonic if \( Pf \leq f \), and harmonic if \( Pf = f \). The following example provides motivation for this definition.

1.1 The fair coin-tossing random walk

Let \( P(x,y) = 1/2 \) if \( |x-y| = 1 \) and 0 else. Then
\[
(Pf)(x) = \frac{1}{2}f(x+1) + \frac{1}{2}f(x-1)
\]
and
\[
(P-I)f(x) = \frac{1}{2}f(x+1) + \frac{1}{2}f(x-1) - f(x)
\]
We can modify the random walk so that we take steps of size \( \delta \) instead of size 1. Then
\[
(P-I)f(x) = \frac{1}{2}f(x+\delta) + \frac{1}{2}f(x-\delta) - f(x)
\]
If \( f \) is twice differentiable, then
\[
\frac{f(x+\delta) - f(x)}{\delta} \to f'(x)
\]
\[
\frac{\frac{1}{2}f(x+\delta) + \frac{1}{2}f(x-\delta) - f(x)}{\delta^2} \to f''(x)
\]
so in this case \((P-I)\) is a second-order difference operator, formally similar to \( \frac{d^2}{dx^2} \).

In fact,
\[
\frac{P^\delta - I}{\delta^2} \to \frac{1}{2} \frac{d^2}{dx^2}.
\]
Thus, functions harmonic relative to \( P \) are analogs of solutions of \( \frac{d^2}{dx^2} f = 0 \). In the field of Analysis, functions satisfying \( \frac{d^2}{dx^2} f = 0 \) are called harmonic, which provides motivation for the Markov chain definition of harmonic.

2 Hitting Times

We now turn our attention to the study of hitting times. Suppose \( A \subseteq B \subseteq S \). Let
\[
T = \inf\{n : X_n \in B\}.
\]
Our task is to compute
\[
h_A(x) = \mathbb{P}_x(T < \infty \text{ and } X_T \in A).
\]
WLOG we may assume that $B$ is absorbing, i.e. that $P(x, x) = 1$ if $x \in B$. Otherwise replace $X_n$ by $X_{T\wedge n}$. Then

$$h_A(x) = \mathbb{P}_x(X_n \in A \text{ eventually}) = \lim_{n \to \infty} \mathbb{P}_x(X_n \in A)$$

The last equality follows because $A$ is absorbing. In other notation, the last expression is

$$\lim_{n \to \infty} P^n(x, A)$$

where $P^n(x, A) = (P^n 1_A)(x)$ and $P^n$ denotes composition of the kernel with itself.

**Theorem 1** $h_A(x)$ is the minimal nonnegative solution $h$ of the boundary value problem $Ph = h$ and

$$h_A(j) = \begin{cases} 1 & \text{if } j \in A \\ 0 & \text{if } j \in B \setminus A \end{cases}$$

**Proof:** The boundary value requirement is clear. We’ve seen that

$$h_A(x) = \lim_{n \to \infty} (P^n 1_A)(x)$$

Note that by definitions and the bounded convergence theorem,

$$(Ph_A)(x) = (P \lim_{n \to \infty} (P^n 1_A))(x) = \int (\lim_{n \to \infty} (P^n 1_A)(y)) P(x, dy) = \lim_{n \to \infty} \int (P^n 1_A)(y) P(x, dy)$$

$$= \lim_{n \to \infty} (P^{n+1} 1_A)(x) = h_A(x)$$

As for minimality, suppose $h \geq 0$ also satisfies both the boundary condition and $Ph = h$. Then $h \geq 1_A$, so $Ph \geq P1_A$, and $P^n h \geq P^n 1_A$. By the fact that $h$ is harmonic and induction on $n$,

$$h = P^n h \geq P^n 1_A \uparrow 1 h_A$$

Note that there may be more than one solution to the boundary value problem in the last theorem. For example, consider the escape probability $e_B(x) = \mathbb{P}_x(T_B = \infty)$, which is the probability that the chain never hits $B$. By conditioning on $X_1$, it can be shown that $e_B = Pe_B$, i.e. that escape probabilities are harmonic. Clearly $e_B = 0$ on $B$. $P$ is a linear operator so if the escape probability is not identically zero and $c$ is a constant, $h_A + ce_B$ is also harmonic and satisfies the boundary conditions.