18.445 Introduction to Stochastic Processes

Lecture 3: Markov chains: time-reversal

Hao Wu

MIT

18 February 2015

Recall

Consider a Markov chain with state space Ω and transition matrix P :

$$\mathbb{P}[X_{n+1} = y \,|\, X_n = x] = P(x, y).$$

- A probability measure π is stationary if $\pi = \pi P$.
- If P is irreducible, there exists a unique stationary distribution.

Today's goal

- Ergodic Theorem
- Time-reversal of Markov chain
- Birth-and-Death chains
- Total variation distance

Ergodic Theorem

Theorem

Let f be a real-valued function defined on Ω. If $(X_n)_n$ is an irreducible Markov chain with stationary distribution π, then for any starting distribution μ, we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n} f(X_j) = \pi f, \quad \mathbb{P}_\mu - a.s.$$

In particular,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n} 1_{[X_j=x]} = \pi(x), \quad \mathbb{P}_\mu - a.s.$$

Detailed balance equations

Definition

Suppose that a probability measrue π on Ω satisfies

$$\pi(x)P(x,y) = \pi(y)P(y,x), \quad \forall x, y \in \Omega.$$

These are called detailed balance equations.

Lemma

Any distribution π satisfying the detailed balance equations is stationary for P.

Definition

A chain satisfying detailed balance equations is called reversible.

Simple random walk on graph

Example Consider simple random walk on graph $G = (V, E)$ (which is connected). The measure

$$\pi(x) = \frac{deg(x)}{2|E|}, \quad x \in \Omega$$

satisfies detailed balance equations ; therefore the simple random walk on G is reversible.

Time-reversal of Markov chain

Theorem

Let (X_n) be an irreducible Markov chain with transition matrix P and stationary distribution π. Define $\widehat{P}$ to be

$$\widehat{P}(x,y) = \frac{\pi(y)P(y,x)}{\pi(x)}.$$

- *$\widehat{P}$ is stochastic*
- *Let $(\widehat{X}_n)$ be a Markov chain with transition matrix $\widehat{P}$. Then π is also stationary for $\widehat{P}$.*
- *For any $x_0, ..., x_n \in \Omega$, we have*
$$\mathbb{P}_\pi[X_0 = x_0, ..., X_n = x_n] = \mathbb{P}_\pi[\widehat{X}_0 = x_n, ..., \widehat{X}_n = x_0].$$

We call $\widehat{X}$ the time-reversal of X.

Remark If a chain with transition matrix P is reversible, then $\widehat{P} = P$ and $\widehat{X}$ has the same law as X.

Birth-and-Death chains

A birth-and-death chain has state space $\Omega = \{0, 1, ..., N\}$.
The current state can be though of as the size of some population ; in a single step of the chain there can be at most one birth or death. The transition probabilities can be specified by $\{(p_k, r_k, q_k)_{k=0}^N\}$ where $p_k + r_k + q_k = 1$ for each k and

- p_k is the probability of moving from k to $k+1$ when $0 \le k < N$; $p_N = 0$
- q_k is the probability of moving from k to $k-1$ when $0 < k \le N$; $q_0 = 0$
- r_k is the probability of remaining at k when $0 \le k \le N$.

Theorem

Every birth-and-death chain is reversible.

Total variation distance

Definition

The total variation distance between two probability measures μ and ν on Ω is defined by

$$||\mu - \nu||_{TV} = \max_{A \subset \Omega} |\mu(A) - \nu(A)|.$$

Lemma

The total variation distance satisfies triangle inequality :

$$||\mu - \nu||_{TV} \leq ||\mu - \eta||_{TV} + ||\eta - \nu||_{TV}.$$

Three ways to characterize the total variation distance

Lemma

$$||\mu - \nu||_{TV} = \frac{1}{2} \sum_{x \in \Omega} |\mu(x) - \nu(x)|.$$

Lemma

$$||\mu - \nu||_{TV} = \frac{1}{2} \sup\{\mu f - \nu f : f \textit{ satisfying } \max_{x \in \Omega} |f(x)| \leq 1\}.$$

Three ways to characterize the total variation distance

Definition

A coupling of two probability measures μ and ν is a pair of random variables (X, Y) defined on the same probability space such that the marginal law of X is μ and the marginal law of Y is ν.

Lemma

$$||\mu - \nu||_{TV} = \inf\{\mathbb{P}[X \neq Y] : (X, Y) \textit{ is a coupling of } \mu, \nu\}.$$

Definition

We call (X, Y) the optimal coupling if $\mathbb{P}[X \neq Y] = ||\mu - \nu||_{TV}$.