18.445 Introduction to Stochastic Processes

Lecture 4: Introduction to Markov chain mixing

Hao Wu

MIT

23 February 2015

Announcement
Midterm : April 6th.(on class)
Final : May 18th.
The tests are closed book, closed notes, no calculators.

Recall
If $(X_n)_n$ is an irreducible Markov chain with stationary distribution π, then

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

Today's goal We will show that X_n converges to π under some "strong sense".

- total variation distance
- the convergence theorem
- mixing times

Three ways to characterize the total variation distance

μ and ν : probability measures on Ω.

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

Lemma

- $$||\mu - \nu||_{TV} = \frac{1}{2} \sum_{x \in \Omega} |\mu(x) - \nu(x)|.$$
- $$||\mu - \nu||_{TV} = \frac{1}{2} \sup\{\mu f - \nu f : f \textit{ satisfying } \max_{x \in \Omega} |f(x)| \leq 1\}.$$
- $$||\mu - \nu||_{TV} = \inf\{\mathbb{P}[X \neq Y] : (X, Y) \textit{ is a coupling of } \mu, \nu\}.$$

Definition

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

The Convergence Theorem

Suppose that $(X_n)_n$ is a Markov chain with transition matrix P. Assume that P is irreducible and aperiodic, then

- there exists r such that $P^r(x, y) > 0$ for all $x, y \in \Omega$;
- there exists a unique stationary distribution π and $\pi(x) > 0$ for all $x \in \Omega$.

Theorem

Suppose that P is irreducible, aperiodic, with stationary distribution π. Then there exist constants $\alpha \in (0, 1)$ and $C > 0$ such that

$$\max_{x \in \Omega} ||P^n(x, \cdot) - \pi||_{TV} \leq C\alpha^n \quad \forall n \geq 1.$$

Mixing time

Definition

$$d(n) = \max_{x \in \Omega} ||P^n(x, \cdot) - \pi||_{TV}$$

$$\bar{d}(n) = \max_{x,y \in \Omega} ||P^n(x, \cdot) - P^n(y, \cdot)||_{TV}$$

Lemma

$$d(n) \leq \bar{d}(n) \leq 2d(n)$$

Lemma

$$\bar{d}(m+n) \leq \bar{d}(m) \cdot \bar{d}(n)$$

Corollary

$$\bar{d}(mn) \leq \bar{d}(n)^m$$

Mixing time

Definition

$$t_{mix} = \min\{n : d(n) \leq 1/4\}, \quad t_{mix}(\epsilon) = \min\{n : d(n) \leq \epsilon\}$$

Lemma

$$t_{mix}(\epsilon) \leq \log(\frac{1}{\epsilon})\frac{t_{mix}}{\log 2}$$

Questions : How long does it take the Markov chain to be close to the stationary measure ?
Lecture 5 : Upper bounds on t_{mix} ; Lecture 6 : Lower bounds on t_{mix} ;
Lecture 7 : Interesting models.

Couple two Markov chains

Definition

A coupling of two Markov chains with transition matrix P is a process $(X_n, Y_n)_{n\geq 0}$ with the following two properties.

- Both (X_n) and (Y_n) are Markov chains with transition matrix P.
- They stay together after their first meet.

Notation : If $(X_n)_{n\geq 0}$ and $(Y_n)_{n\geq 0}$ are coupled Markov chains with $X_0 = x$, $Y_0 = y$, then we denote by $\mathbb{P}_{x,y}$ the law of $(X_n, Y_n)_{n\geq 0}$.

Couple two Markov chains

Theorem

Suppose that P is irreducible with stationary distribution π. Let $(X_n, Y_n)_{n\geq 0}$ be a coupling of Markov chains with transition matrix P for which $X_0 = x$, $Y_0 = y$. Define τ to be their first meet time :

$$\tau = \min\{n \geq 0 : X_n = Y_n\}.$$

Then

$$||P^n(x,\cdot) - P^n(y,\cdot)||_{TV} \leq \mathbb{P}_{x,y}[\tau > n].$$

In particular,

$$d(n) \leq \max_{x,y} \mathbb{P}_{x,y}[\tau > n].$$

Random walk on $N-$cycle : Upper bound on t_{mix}

Lazy walk : it remains in current position with probability $1/2$, moves left with probability $1/4$, right with probability $1/4$.

- It is irreducible.
- The stationary measure is the uniform measure.

Theorem

For the lazy walk on N$-$cycle, we have

$$t_{mix} \leq N^2.$$