6.896: Probability and Computation

Spring 2011

lecture 3

Constantinos (Costis) Daskalakis
costis@mit.edu

recap

Markov Chains

Def: A ***Markov Chain*** on Ω is a stochastic process $(X_0, X_1, \ldots, X_t, \ldots)$ such that

a. $X_t \in \Omega, \ \forall t$

b. $\Pr[X_{t+1} = y \mid X_t = x, X_{t-1} = x_{t-1}, \ldots, X_0 = x_0] \equiv \Pr[X_{t+1} = y \mid X_t = x]$

$$\stackrel{!!}{=} P(x, y)$$

$p_x^{(t)} \in \mathbb{R}_+^{1 \times |\Omega|}$: distribution of X_t conditioning on $X_0 = x$.

then

$$p_x^{(t+1)} = p_x^{(t)} P$$

$$p_x^{(t)} = p_x^{(0)} P^t$$

Graphical Representation

Represent Markov chain by a graph $G(P)$:

- nodes are identified with elements of the state-space Ω
- there is a directed edge between states x and y if $P(x, y) > 0$, with edge-weight $P(x,y)$; (no edge if P(x,y)=0;)
- self loops are allowed (when $P(x,x) > 0$)

Much of the theory of Markov chains only depends on the topology of $G(P)$, rather than its edge-weights.

irreducibility

aperiodicity

Mixing of Reversible Markov Chains

Theorem (Fundamental Theorem of Markov Chains) :

If a Markov chain P is ***irreducible*** and ***aperiodic*** then:

a. it has a unique stationary distribution $\pi = \pi P$;

b. $p^t_x \rightarrow \pi$ as $t \rightarrow \infty$, for all $x \in \Omega$.

examples

Example 1

Random Walk on undirected Graph G=(V, E)

$$P(x,y) = \begin{cases} 1/deg(x), & \text{if } (x,y) \in E \\ 0, & \text{otherwise} \end{cases}$$

P is irreducible iff G is connected;

P is aperiodic iff G is non-bipartite;

P is reversible with respect to $\pi(x) = \deg(x)/(2|E|)$.

Example 2

Card Shufflings

Showed convergence to uniform distribution using concept of doubly stochastic matrices.

Example 3

Designing Markov Chains

Input: A positive weight function $w : \Omega \to R^+$.

Goal: Define MC $(X_t)_t$ such that $p^t_x \to \pi$, where $\pi(x) = \frac{w(x)}{Z}$

The Metropolis Approach

- define a graph $\Gamma(\Omega, E)$ on Ω.
- define a ***proposal distribution*** $\kappa(x, \cdot)$ s.t.

$$\kappa(x, y) = \kappa(y, x) > 0, \text{ for all } (x, y) \in E, x \neq y$$

$$\kappa(x, y) = 0, \text{ if } (x,y) \notin E$$

Example 3

The Metropolis Approach

- define a graph $\Gamma(\Omega,E)$ on Ω.
- define a ***proposal distribution*** $\kappa(x, \cdot\ )$ s.t.

$$\kappa(x, y) = \kappa(y, x) > 0, \text{ for all } (x, y) \in E, x \neq y$$

$$\kappa(x, y) = 0, \text{ if } (x,y) \notin E$$

the Metropolis chain: When chain is at state x

- Pick neighbor y with probability $\kappa(x, y)$
- Accept move to y with probability

$$\min\left\{1, \frac{w(y)}{w(x)}\right\} \equiv \min\left\{1, \frac{\pi(y)}{\pi(x)}\right\}$$

Example 3

the Metropolis chain: When process is at state x

- Pick neighbor y with probability $\kappa(x, y)$
- Accept move to y with probability $\min\left\{1, \frac{w(y)}{w(x)}\right\}$

Transition Matrix:

$$P(x,y) = \kappa(x,y)\min\left\{1, \frac{w(y)}{w(x)}\right\} \equiv \kappa(x,y)\min\left\{1, \frac{\pi(y)}{\pi(x)}\right\}$$

Claim: The Metropolis Markov chain is reversible wrt to $\pi(x) = w(x)/Z$.

Proof: 1point exercise

back to the fundamental theorem

Mixing of Reversible Markov Chains

Theorem (Fundamental Theorem of Markov Chains) :

If a Markov chain P is ***irreducible*** and ***aperiodic*** then:

a. it has a unique stationary distribution $\pi = \pi P$;

b. $p^t_x \rightarrow \pi$ as $t \rightarrow \infty$, for all $x \in \Omega$.

Proof: see notes.