ORIE 4580/5580: Simulation Modeling and Analysis

ORIE 5581: Monte Carlo Simulation

Unit 15: Markov chain Monte Carlo

Sid Banerjee

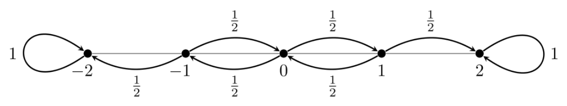
School of ORIE, Cornell University

Markov chains: long-term behavior

- Absorbing MC
- Non-absorbing
  - transient (ie, never settles down)
  - (positive) recurrent

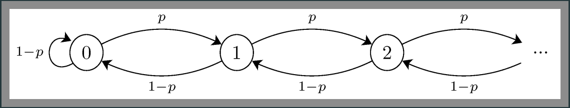
$\mathbb{P}[X(t)=i]$ settles down to some constant π_i for every state i as $t \uparrow \infty$

Markov chains: absorbing chain



- finite absorbing MC always get absorbed

Markov chains: transient/recurrent chain

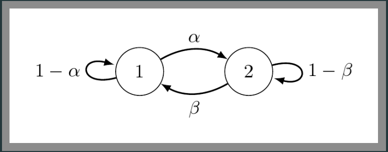


- transient if $p > 1-p$ (goes to ∞)

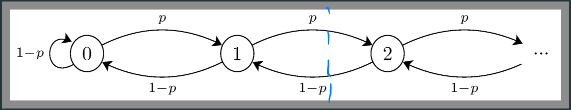
- positive recurrent if $1-p > p$

Markov chains: steady-state behavior

example:



Markov chains: steady-state for infinite chains

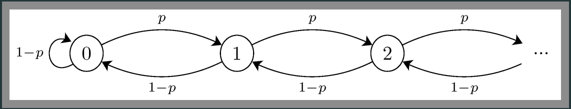


- Steady-state distn obey 'flow-balance'

$$\pi_1 P_{12} = \pi_2 P_{21} , \quad \sum_{i=0}^{\infty} \pi_i = 1$$

$$\Rightarrow \quad \pi_i = \left(\frac{p}{1-p}\right)^i \left(1 - \frac{p}{1-p}\right)$$

Markov chains: reversibility



Idea - If I simulate the MC and then 'play it in reverse'

The reverse MC is also a MC, but may have different transitions.

- Suppose forward chain had transitions P, reverse has Q

For every pair of states x, y

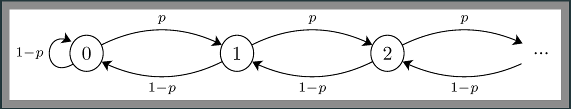
steady-state dist'n

$$\pi(x) P(x,y) = \pi(y) Q(y,x) \quad \circledast$$

(Kelly's Lemma)

Moreover - If you are given Q, π satisfying $\circledast$ $\Rightarrow \pi$ is the steady-state dist for P!

Markov chains: the ergodic theorem



ergodic ≡ averages over time = averages over states

$X(t)$ is a positive recurrent MC

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T f(X(t))\,dt = c = \sum_{x\in\mathcal{X}} \pi_x f(x)$$

steady-state dist $\lim_{t\to\infty} \mathbb{P}[X(t)=x]$

Also

$$\frac{1}{T}\int_0^T f(X(t))\,dt \approx \mathbb{E}_\pi[f(X)]$$

for large T

Markov-chain monte carlo

Idea - $\mathbb{E}_{\pi}[f(x)] \approx \frac{1}{T}\int_0^T f(x(t))\,dt$

Appl 1

DES

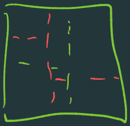
good approx $\longleftarrow$ want to calculate via DES

Appl 2

MCMC

want to compute this $\longrightarrow$ use MC sim s.t. the MC has steady state π

Eg - Gerrymander



the Metropolis algorithm (Nicholas Metropolis)

partition function

- target distribution $P(x) = \underline{\widetilde{P}(x)}/Z$ (what we want to sample)
- proposal distribution(s) $\underbrace{Q(x|y)}_{\text{Markov chain over states}}$, with $\boxed{Q(x|y) = Q(y|x)}\, \forall\, x, y$

Metropolis sampling

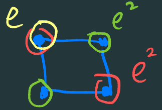
1. choose initial Z_0
2. to obtain sample t, generate $Y_t \sim Q(\cdot|Z_{t-1})$
3. accept $Z_t = Y_t$ with probability $A(Y_t, Z_{t-1}) = \max\left\{1, \frac{\widetilde{P}(Y_t)}{\widetilde{P}(Z_{t-1})}\right\}$
   else reject and set $Z_t = Z_{t-1}$

proposal

current

Eg dist - Given graph G, want to sample 'independent sets' (ie, collection of nodes s.t. no neighbours)

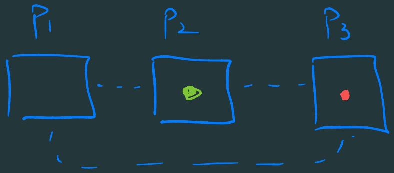
Aim - sample ind set S prop to $e^{|S|}$



Metropolis-Hastings:example

generate r.v. $X \in \{1, 2, \ldots, n\}$ s.t. $\mathbb{P}[X = i] = p_i$ for each $i \in [n]$

$n=3$



$Q \equiv$ random neighbor with wraparound - steady state $\equiv$ uniform over $\{1, 2, \ldots, n\}$

Metropolis - Given $Z_0 = 2$, proposal $Y_1 = 3$

- Accept Y_1 with prob $\max\{1, p_3/p_2\}$. Set $Z_1 = 3$

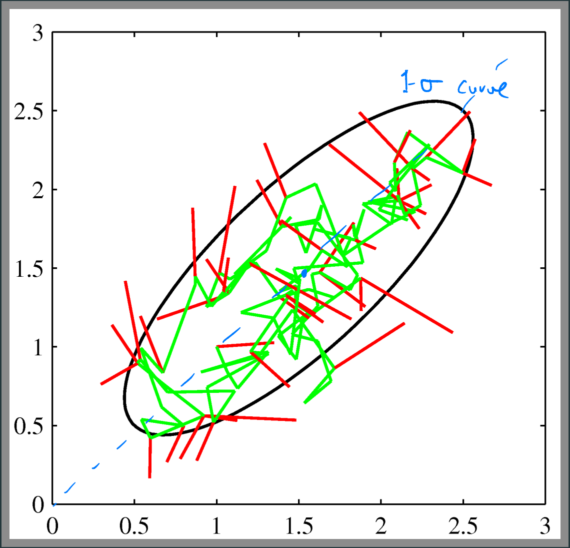
Else set $Z_1 = 2$

Guarantee - In steady-state, sample from $\{p_1, p_2, \ldots, p_n\}$

Metropolis-Hastings:example

generate r.v. $X \in \{1, 2, \ldots, n\}$ s.t. $\mathbb{P}[X = i] = p_i$ for each $i \in [n]$

Metropolis for 2-d Gaussian



Metropolis algorithm: proof of correctness

For any x, y, $P(x,y) = Q(x,y)\min\left\{1, \frac{\tilde{P}(y)}{\tilde{P}(x)}\right\}$

set $\Pi(x) \propto \tilde{P}(x)$, set reverse transitions = forward transitions

- $$\Pi(x)\,P(x,y) = \frac{\tilde{P}(x)}{\cancel{Z}} \cdot \cancel{Q}(x,y) \cdot \min\left\{\cancel{1}, \frac{\tilde{P}(y)}{\tilde{P}(x)}\right\}$$

$\|$ under uniform random neighbor

$$\Pi(y)\,P(y,x) = \frac{\tilde{P}(y)}{\cancel{Z}} \cdot \cancel{Q}(y,x) \cdot \min\left\{1, \cancel{\frac{\tilde{P}(x)}{\tilde{P}(y)}}\right\}$$

Metropolis-Hastings

- target distribution $P(x) = \widetilde{P}(x)/Z$
- proposal distribution(s) $Q(x|y)$

Metropolis-Hastings sampling

1. choose initial Z_0
2. to obtain sample t, generate $Y_t \sim Q(\cdot|Z_{t-1})$
3. accept $Z_t = Y_y$ with prob $A(Y_t, Z_{t-1}) = \max\left\{1, \frac{\widetilde{P}(Y_t)Q(Z_{t-1}|Z_t)}{\widetilde{P}(Z_{t-1})Q(Z_t|Z_{t-1})}\right\}$
   else reject and set $Z_t = Z_{t-1}$

Gibbs sampling

- target distribution $P(x(1), x(2), \ldots, x(n))$

Gibbs sampling

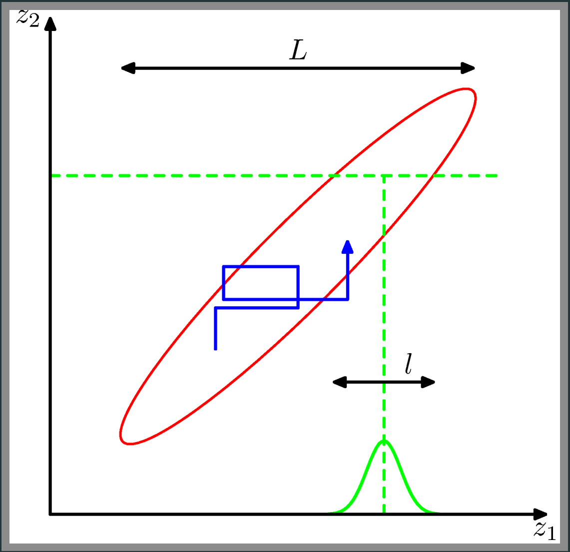
1. choose initial $X_0 = (X_0(1), X_0(2), \ldots, X_0(n))$
2. to obtain sample t:

   pick I_t uniformly at random

   set $X_t(i) = X_{t-1}(i)$ for $i \neq I_t$

   set $X_t(I_t) \sim P(\cdot | X_{t-1} \setminus X_{t-1}(I_t))$

Gibbs sampling for 2-d Gaussian



Gibbs sampling: proof of correctness