Class 14

Markov Chains II
Announcements

• No HW at the moment!
• Next week is fall break!
• This Friday is the last day to change to/from CR/NC, and also the last day to withdraw.
Recap: More Markov Chains!

• **Definitions!**
  • A chain is **irreducible** if you can get to anywhere from anywhere else.
  • A state is **recurrent** if you’ll return to it eventually with probability 1.
  • Otherwise it is **transient**.
  • A chain is **periodic** if there’s a state that you can only reach on multiples of c, for some integer $c > 1$.
  • Otherwise it is **aperiodic**.
    • Useful fact: any irreducible chain with a self-loop is aperiodic!
Recap: Fundamental Theorem of Markov Chains

• Any irreducible and aperiodic Markov chain over a finite state space has a unique **stationary distribution** $\pi$.

• As $t$ gets big, $X_t \to \pi$

• $\pi P = \pi$, aka if $X_t \sim \pi$, then $X_{t+1} \sim \pi$

• If you start in state $i$, the expected amount of time to return is $\frac{1}{\pi_i}$
Tie-in to last time...

- Proposition: If a Markov chain has symmetric transitions, and is aperiodic and irreducible, then the stationary distribution is uniform.
Metropolis algorithm and MCMC

• **Markov Chain Monte Carlo:**
  - Set up a Markov Chain with a particular desired stationary distribution.
  - To sample from that distribution, run the chain for a while!

• **Metropolis Algorithm:**
  - A particular way to set up such a chain.

\[
P_{i,j} = \begin{cases} 
0 & \text{if } i, j \text{ not neighbors} \\
\frac{1}{d} \min(1, \frac{\pi(j)}{\pi(i)}) & \text{if } i \neq j \text{ and they are neighbors} \\
1 - \sum_{\ell \neq i} P_{i,\ell} & \text{if } i = j.
\end{cases}
\]
Questions?
Fundamental Theorem, Metropolis Alg? Quiz?
Question 1

• Which states are recurrent?
  • A,B,C

• Which states are transient?
  • D

• Is this chain irreducible?
  • No, it’s reducible into {D}, {A,B,C}.

• Is this chain periodic?
  • No, it’s aperiodic.
Question 2

• What is the stationary distribution?
  • (\(\frac{3}{8}, \frac{1}{4}, \frac{3}{8}\))
Question 3

• Choose a card uniformly at random.
• Move it to the top of the deck.

• Is the MC symmetric? (No)
• Is it irreducible? (Yes)
• Is it periodic? (No)
• Is the stationary distribution uniform? (Yes)
Today: Gibbs Sampling!

• An MCMC algorithm for multivariate distributions.

• Set-up:
  • $\pi$ is a joint distribution on random variables $X, Y$
    • More generally $X_1, X_2, ..., X_m$
  • It’s hard to sample from $\pi$
  • But it’s easy to sample from $\pi(X \mid Y = y)$ or $\pi(Y \mid X = x)$ for any fixed $x, y$. 
Gibbs Sampling
(for two variables)

• Say \((X_t, Y_t) = (x, y)\)
• Draw \(x' \sim \pi(X|Y = y)\)
• Draw \(y' \sim \pi(Y|X = x')\)
• Set \((X_{t+1}, Y_{t+1}) = (x', y')\)
1. Show that the uniform distribution is a stationary distribution.
2. Under what conditions on $\pi$ does FTOMC hold?
3. What is the take-away in the context of MCMC?
4. How would you use Gibbs sampling to sample random colorings?
5. How would you use Gibbs sampling to sample a uniformly random 7-word sentence from the distribution of all reasonable such sentences?
6. Any other applications of Gibbs sampling/MCMC that you’ve encountered?

- Say $(X_t, Y_t) = (x, y)$
- Draw $x' \sim \pi(X | Y = y)$
- Draw $y' \sim \pi(Y | X = x')$
- Set $(X_{t+1}, Y_{t+1}) = (x', y')$
1. Stationary dist is $\pi$

- Show $\pi(x, y) = \sum_{x', y'} \pi(x', y') \Pr[(x', y') \rightarrow (x, y)]$
2. Do the conditions hold?

• Aperiodic:

• Irreducible:

• Finite:
2. Do the conditions hold?

• Not necessarily irreducible:

\[ \text{eg. if } x, y \in \{0, 1\}, \quad \Gamma \text{ is: } \begin{array}{c|c} \hline x & \\hline 0 & 1/2 \\hline 1 & 0 \\hline \end{array} \]

• It is irreducible if this bipartite graph is connected:

One way to get from \((x, y)\) to \((x', y')\):

\((x, y) \rightarrow (x', y) \rightarrow (x'', y') \rightarrow (x', y'') \rightarrow (x''', y') \rightarrow (x, y')\)
3. Why is this useful?

• If we can easily sample from $\pi(X | Y = y)$ or $\pi(Y | X = x)$, then we can sample $(X_t, Y_t)$.

• As $t \to \infty$, this will converge to $\pi$, so eventually we can sample from $\pi$!
  • How long does it take to converge????
4. Graph Coloring

• What is the right multivariate generalization?

• How to apply to sampling a random proper coloring?
5. Sampling 7-word sentences

• What is the algorithm? What task do you need to be able to do?
6. Other examples?

• Y’all come from many different areas – have any of you used Gibbs sampling or any other MCMC method before? For what applications?
Another Example: Image Denoising

• Say you get a noisy (black and white, say) image $X = (x_1, ..., x_N)$.
  • Each pixel $x_p$ is $\pm 1$
• Sample an “un-noisy” version $Y = (y_1, ..., y_N)$, so that the probability
  of $Y$ is proportional to:

$$\exp(\eta \sum_p x_p y_p + \beta \sum_{p \sim p^\prime} y_p y_{p^\prime})$$
Recap

• The fundamental theorem of Markov chains can be useful!
• But it sure would be more useful if we knew how fast we approached the stationary distribution...
  • Next time!