Mixing Time

CS 70, Summer 2019

Bonus Lecture, 8/9/19

Disclaimer:

Much handwaving today!!

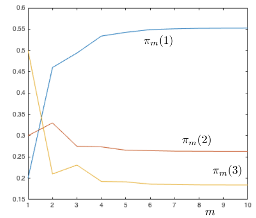
Goal is to get a high level intuition / picture for the concept of mixing time and applications

Emphasis is on **heuristics** rather than rigor

What We Know...

Every irreducible, aperiodic Markov chain has a **unique** stationary distribution.

$$\pi_m = \pi_0 P^m = \pi_0 \begin{bmatrix} 0.8 & 0.2 & 0 \\ 0 & 0.3 & 0.7 \\ 0.6 & 0.4 & 0 \end{bmatrix}^m.$$



Q: How long does it take to get **close** to the stationary distribution?

Total Variation Distance

Just one of many ways to measure how **close** two distributions are.

Let P_1, P_2 be two PMFs. Their TV distance is:

Mixing Time: Definition

I have an **irreducible, aperiodic** Markov chain. Notation: $\mu^{(n)}$ is the distribution **at time** n, and π is its **(unique) stationary distribution**.

I want to keep running my chain until:

The mixing time $t_{mix}(\varepsilon)$ is the first time this happens.

(Omitted fact: The TV distance between $\mu^{(n)}$ and π decreases as n increases.)

Complete Graph With Loops

Mixing time analysis sometimes direct!

Take a random walk on a complete graph **with loops**. What is the stationary distribution?

What does the transition matrix look like?

Mixing Time? Dependence on ε?

Random-To-Top Shuffling I

$\mathcal{S}$: All orderings of n cards in a deck.
Transitions: Choose card randomly in the deck.
Move it to the top.

Different strategy called **coupling:**

Random-To-Top Shuffling II

At each time, each card is labeled **coupled** or not.

Initially, all cards start **uncoupled**.
Pick a random card C.

In both decks, take card C and move it to the top.
If C isn't already coupled, mark it as **coupled**.

What happens when we look at each deck individually?

Random-To-Top Shuffling III

Time until all cards get coupled = time until Deck 1 is fully random.

For all ϵ: $t_{mix}(\varepsilon) =$

Random Hypercube Walk I

Take an n-dimensional hypercube.
Stationary distribution of hypercube walk:

Try coupling again:

Random Hypercube Walk II

Each **coordinate** is labeled **coupled** or not.

Initially all coordinates are **uncoupled**.
Pick a random coordinate i.
Flip a coin to set the i-th coordinate to 0 or 1.
If i-th coordinate is **uncoupled**, set it to **coupled.**

Random Hypercube Walk III

Use Coupon Collector again!

For all ε, $t_{mix}(\varepsilon) =$

Conditions for Fast Mixing?

Complete graph K_n?

Path on n vertices?

Dumbbell?

Bottlenecks

We use a measurement called the **conductance** to quantify the notion of a bottleneck.

The conductance of a set $A \subseteq \mathcal{S}$ is:

$$\Phi(A) =$$

The conductance of the chain M is:

$$\Phi(M) =$$

How To Measure Conductance?

Measuring conductance = looking at all subsets of states with $\text{vol}(A) \leq \frac{1}{2}$.
How many subsets, potentially?

Alternative: Get **lower bound** on $\Phi(M)$ using **second largest eigenvalue** of transition matrix.

$$\Phi(M) \geq$$

Eigenvalues are much faster to compute!

Markov Chain Monte Carlo

Monte Carlo: randomized algorithm where the output is allowed to be incorrect

Use cases:

- sampling from complicated distributions
- counting combinatorial objects
- Bayesian inference
- statistical physics
- volume estimation, integration

Markov Chain Monte Carlo

Key idea: Design a Markov chain so that its stationary follows the distribution that you want to sample from. Run the chain, wait for it to mix.

Runtime depends on...