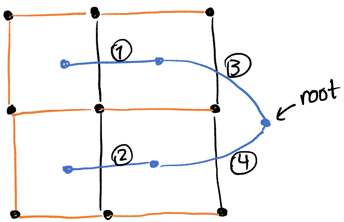
Review

- Counting Bipartite Planar Perfect Matchings

$$\mathrm{per}\left(\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}\right) = \mathrm{det}\left(\begin{bmatrix}1 & -1\\ 1 & 1\end{bmatrix}\right)$$

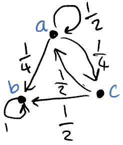
Pólya's scheme

  - * Find Pfaffian orientation = odd fwd edges around every face
  - * First do orange edges arbitrarily, then blue edges, one leaf at a time.



- Intro to Markov Chains
  - * Transition matrix $P \in \mathbb{R}_{\geq 0}^{\Omega \times \Omega}$

    row stochastic



$$\begin{array}{c|ccc} & a & b & c \\ \hline a & \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ b & 0 & 1 & 0 \\ c & \frac{1}{2} & \frac{1}{2} & 0 \end{array}$$

  - * Stationary μ: $\mu P = \mu$
  - * Fundamental Thm: ergodic $\Rightarrow d_{TV}(\nu P^t, \mu) \longrightarrow 0$

Plan for Today

* Get familiar with coupling

- Finish Fundamental Thm
- Mixing Time
- How to Design Markov Chains
- Coloring

if time

Fundamental Theorem

- Irreducibility:
  Possible to reach from every x to every y with >0 prob.
- Aperiodicity:
  For large enough t, possible to cycle $x \to x$ in exactly t steps. $\gcd(\text{lengths}) = 1$

ergodic

Thm: For an ergodic chain P

$\exists$ unique stationary dist μ and for any starting ν

$$\lim_{t\to\infty} d_{TV}(\nu P^t, \mu) = 0$$

Proof of Fundamental Thm

Attempt #1:

Lemma: $d_{TV}(\nu P, \nu' P) \le d_{TV}(\nu, \nu')$

If we could sneak 0.99 we'd be done.

- $\nu, \nu P, \nu P^2, \nu P^3, \dots$ ← Cauchy sequence

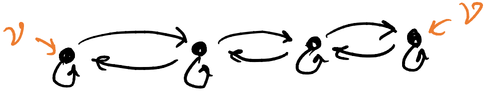
$$d_{TV}(\nu P^n, \nu P^m) \le 0.99^n d_{TV}(\nu, \nu P^{m-n}) \le 0.99^n$$

So it converges.

- Stationary is unique:

$d_{TV}(\mu,\mu') \leftarrow d_{TV}(\mu P, \mu' P) \le 0.99\, d_{TV}(\mu, \mu')$.

Bad Example.



Weak Contraction: $d_{TV}(\nu P, \nu' P) \leq d_{TV}(\nu, \nu')$

Proof: From HW :

$$d_{TV}(\nu, \nu') = \min\left\{ \mathbb{P}[X \neq X'] \;\middle|\; \underbrace{\begin{matrix} X \sim \nu \\ X' \sim \nu' \end{matrix}}_{\text{coupling of } \nu, \nu'} \right\}$$

Let $X_0 \sim \nu$ and $X_0' \sim \nu'$ (not indep.)

Evolve to get $X_1 \sim \nu P$, $X_1' \sim \nu' P$

- If $X_0 = X_0'$ use same transition
- Otherwise evolve independently.

$X_1 \neq X_1' \Rightarrow X_0 \neq X_0'$, so

$$\mathbb{P}[X_1 \neq X_1'] \leq \mathbb{P}[X_0 \neq X_0']$$

Idea: If no matter X_0, X_0', there is $\geq \varepsilon$ chance X_1, X_1' are equal

$\Rightarrow d_{TV}(\nu P, \nu' P) \leq (1-\varepsilon) d_{TV}(\nu, \nu')$

$\Rightarrow$ done!

This is not true in general, but true for t steps of P (P^t):

- Ergodic $\Rightarrow$ for all large enough t, P^t has > 0 entries
- There is >0 chance of collision.
- $d_{TV}(\nu P^t, \nu' P^t) \leq (1-\varepsilon)\, d_{TV}(\nu, \nu')$.

$$\begin{array}{c|cc} & \overset{0}{\downarrow} & \overset{1}{\downarrow} \\ \hline 0 \rightarrow & a & b \\ 1 \rightarrow & c & d \end{array}$$

$\nu = Ber(p)$

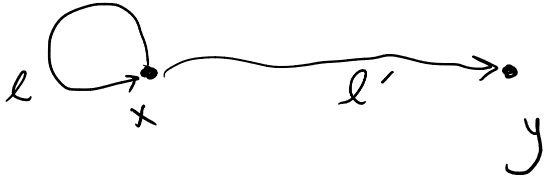
$\nu' = Ber(q)$

$a+b=p \qquad a+c=q$

$c+d=1-p \qquad b+d=1-q$

$$\min\left\{ \underset{\overset{\uparrow}{a}}{c}+b \;\middle|\; \text{subj to } \overset{\overset{q}{\uparrow}}{\text{conditions}} \right\}$$

Why ergodic $\Rightarrow P^t > 0$ for large t



$\exists$ loops of all len $\ell \geq \ell_0$.

$\Rightarrow \exists$ $x \to y$ paths of all len $\geq \ell_0 + \ell'$.

Note: Irreducible $\Rightarrow$ All x's have same period.

Reminder: Aperiodicity is easy to get.

$$P \Rightarrow \frac{P+I}{2}.$$

Mixing Time

For $\varepsilon > 0$ define

$$t_{mix}(P, \nu, \varepsilon) = \min\{t \mid d_{TV}(\nu P^t, \mu) \leq \varepsilon\}$$

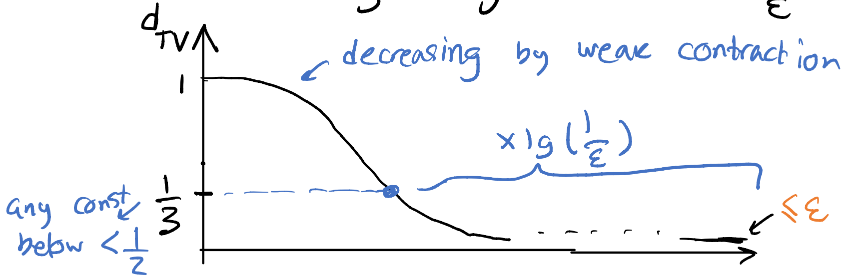
$$t_{mix}(P, \varepsilon) = \max\{t_{mix}(P, \nu, \varepsilon) \mid \nu\}$$

We usually want $t_{mix} = \mathrm{poly}\lg(|\Omega|)$

because this is exponential.

What about dependence on ε?

Thm: It's always logarithmic in $\frac{1}{\varepsilon}$ ☺



This lets us talk t_{mix} without specifying ε.

$$t_{mix} :\simeq t_{mix}(\varepsilon = \underbrace{\tfrac{1}{3}}_{\text{arbitrary} < \frac{1}{2}})$$

Proof: Let $t \geq t_{mix}(P, \frac{1}{3})$.

For any starting points X_0, X_0', if X_t, X_t' are t-step evolutions

$$\Rightarrow d_{TV}(X_t, X_t') \leq d_{TV}(X_t, \mu) + d_{TV}(\mu, X_t') \leq \frac{2}{3}$$

this is why $\geq \frac{1}{2}$ doesn't work

Now take coupling X_0, X_0' of ν, ν':

- If equal evolve identically.
- Else couple evolutions.

$$\mathbb{P}[X_t \neq X_t'] \leq \frac{2}{3} \mathbb{P}[X_0 \neq X_0']$$

In other words

$$d_{TV}(\nu P^t, \nu' P^t) \leq \frac{2}{3} d_{TV}(\nu, \nu').$$

$$\Rightarrow d_{TV}(\nu P^{kt}, \nu' P^{kt}) \leq \left(\frac{2}{3}\right)^k d_{TV}(\nu, \nu')$$

Let $k \simeq \lg \frac{1}{\varepsilon}$ and $\nu' = \mu$.

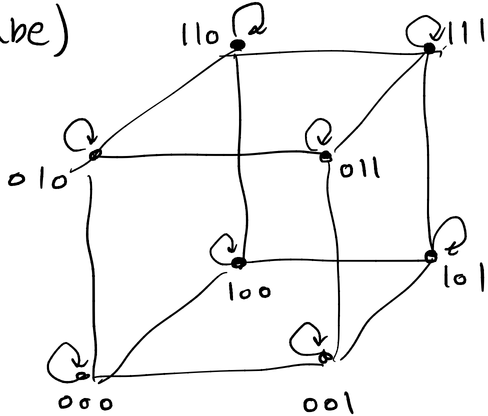
Thm: $t_{mix}(\varepsilon) \leq t_{mix}(\frac{1}{3}) \cdot O(\lg \frac{1}{\varepsilon})$

Remark: Usually in our examples, this is not tight :/

Cutoff phenomenon

Example. (Hypercube)

Replace u.r. coord by $\text{Ber}(\frac{1}{2})$.



$$X_0 \longmapsto X_1 \longmapsto \cdots$$

$$x_1 \cdots x_n \longmapsto x_1 \cdots \text{Ber}(\tfrac{1}{2}) \cdots x_n \longmapsto$$

$$\text{Ber}(\tfrac{1}{2}) \cdots \text{Ber}(\tfrac{1}{2}) \cdots x_n \longmapsto \cdots$$

Define τ: First time we touched all coords.

$$\text{Dist}(X_\tau \mid \tau = k) \underset{\uparrow}{=} \text{uniform}$$

this makes τ strong stationary [Aldous-Diaconis]

Lemma: $\mathbb{P}[\tau > t] \le \varepsilon \Rightarrow t_{mix}(\varepsilon) \le t$.

Proof: X_t conditioned on $\tau \le t$ $\sim$ stationary

$$\text{dist}(X_t \mid \tau \le t) = \frac{\mathbb{P}[\tau=0]\,\overset{\text{Stationary}}{\text{dist}(X_0 \mid \tau=0)}P^t + \cdots + \mathbb{P}[\tau=t]\,\overset{\text{Stationary}}{\text{dist}(X_t \mid \tau=t)}}{\mathbb{P}[\tau \le t]}$$

Now couple X_t with μ:

If $\tau \le t$: perfect coupling

Else: arbitrary coupling

$$d_{TV}(X_t, \mu) \le P[\tau > t].$$

Hypercube: $P[\tau > t] \le \underset{\#\text{coords}}{n}\ \underbrace{(1-\tfrac{1}{n})}_{\text{prob of not touching coord } i}{}^{t}$

$\simeq n e^{-\frac{t}{n}}$

$$\Rightarrow t_{mix}(\varepsilon) \le n\lg n + \underbrace{n\lg\tfrac{1}{\varepsilon}}_{\text{much better dependence on } \varepsilon.}$$

Next HW: Show $t_{mix} \ge \Omega(n\lg n)$.

Note: This is NOT proving cutoff!

How to Design Markov Chains

Main Criteria: Correct stationary dist!

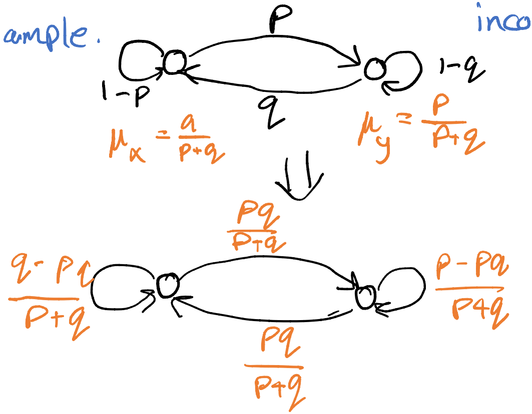
Ergodic Flow: For dist μ and Markov Chain P, define

$$Q(x,y) = \underbrace{\mu(x)P(x,y)}_{\text{prob mass flowing from } x \to y}$$

Lemma. μ stationary $\leftrightarrow$ Q is proper flow

incoming = outgoing

Example.



$\mu_x = \frac{q}{p+q}$ $\mu_y = \frac{p}{p+q}$

Proof. $\sum_x \mu(x)P(x,y) = \sum_z \mu(y)P(y,z) = \mu(y)$ ✓

Time-Reversible / Detailed Balance:

When $Q(x,y) = Q(y,x) \Rightarrow$ flow is proper!

$$\Updownarrow$$

$$\mu(x)P(x,y) = \mu(y)P(y,x)$$

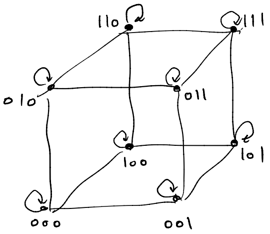
Reversible $\equiv$ Random Walk on Undirected Graph

One step of P: Pick y w.p. $\propto Q(x,y)$ ← Random walk defined by Q

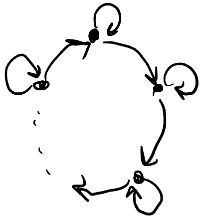
Conversely: If Q is symmetric and P is random walk defined by Q,

$$\mu(x) \propto \sum_y Q(x,y) \Rightarrow \mu(x)P(x,y) \propto Q(x,y).$$

Example.



Non-Example.



Time-Reveral

- Markovian :

$$\cdots \underbrace{X_0 - X_1} - X_2 - \underbrace{X_3 - \cdots}$$

independent
conitioned on X_2

- If X_i's are stationary (equilibrium)

$\underbrace{(X_i, X_{i+1})}$ are distributed the same way.

If we reverse time, both properties still hold. So

$$\cdots - \underbrace{X_3 - X_2} - X_1 - X_0 - \cdots$$

is also a run of a Markov chain at equilibrium.

Recipe for time reveral:

- Revert directions in ergodic flow
- Equivalently define

$$P'(x,y) = \frac{\mu(y)\, P(y,x)}{\mu(x)}$$

prob transition as long as $\mu P = \mu$.

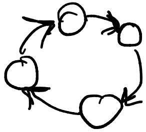
$$\mu(x) P'(x,y) = \mu(y) P(y,x)$$

Under detailed balance, time-reversal results in $P' = P$.

detailed balance ≡ time-reversible ≡ undirected random walk

Observing Markov chain at equilibrium:

- Detailed balance: can't tell direction of time
- Generally: can tell direction



Many Markov chains we study will be time-reversible. Using detailed balance we can design Markov Chains with given stationary distributions.

---

① Metropolis Rule:

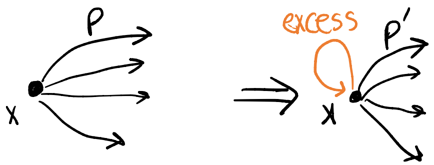
Suppose P doesn't necessarily have μ as stationary dist. Define P':

$$P'(x,y) := \min\left\{1, \frac{\mu(y)P(y,x)}{\mu(x)P(x,y)}\right\} \cdot P(x,y)$$

when $x \neq y$

accept/reject filter

$$\Rightarrow \mu(x)P'(x,y) = \mu(y)P'(y,x)$$

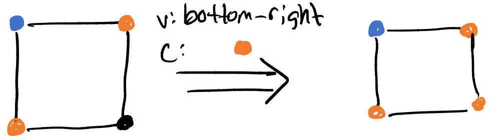


Note: Only need to know μ proportionally.

Example. (Coloring)

μ: uniform over valid colorings

P: pick u.a.r. v and u.a.r. c and color v using c.



Metropolis Rule:

Accept transition w. prob

$$\min\left\{1, \frac{\mu(y)}{\mu(x)} \cdot \frac{\cancel{P(y,x)}}{\cancel{P(x,y)}}\right\} = 1[y \text{ valid}]$$

assume x valid

② Combination with time-reversal

Suppose dist π on $\Omega \times \Omega'$

Want: Marginal on Ω as stationary dist.

we typically describe $X, Y \sim \pi$ by $\text{dist}(X)$ and $\text{dist}(Y|X)$

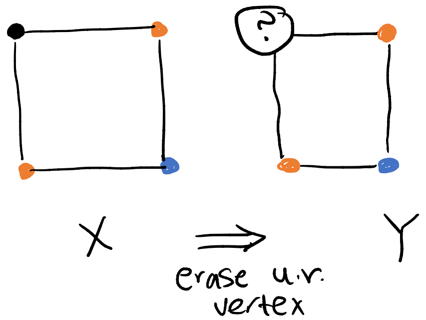
Construction:

$$\begin{tikzcd}
X_0 \arrow[r, "\text{dist}(Y|x_0)"', "\text{if } (X_0,Y)\sim\pi"'{yshift=-12pt}] \arrow[rr, bend left, "\text{one step}"] & Y \arrow[r, "\text{dist}(X_1|Y)"', "\text{if } (X_1,Y)\sim\pi"'{yshift=-12pt}] & X_1
\end{tikzcd}$$

$$Q(x_0,x_1) = \sum_y \left[\underbrace{\sum_{y'} \pi(x_0,y')}_{\text{dist } \mu(x_0)} \cdot \underset{\text{prob}(X_0 \to Y)}{\frac{\pi(x_0,y)}{\sum_{y'} \pi(x_0,y')}} \cdot \frac{\pi(x_1,y)}{\sum_{x'} \pi(x',y)} \right]$$

$$= \sum_y \left[\frac{\pi(x_0,y)\,\pi(x_1,y)}{\sum_{x'} \pi(x',y)} \right] \leftarrow \text{reversible}$$

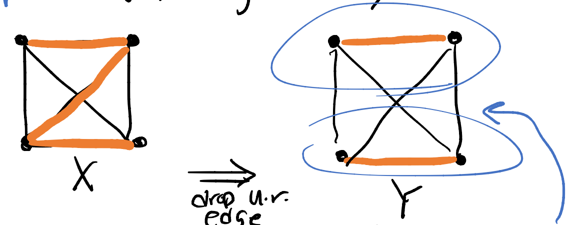
Example (Glauber Dynamics)



Chain: Recolor erased vertex w. prob $\propto \mu(\text{resulting config})$

Bayes Rule

Example (Spanning Trees)



Chain: Add u.r. edge from the cut