Markov Chains

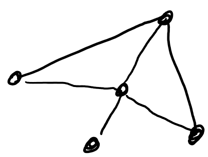
Monday, February 27, 2017 1:48 PM

If $\mathbb{P}[X_t \in A \mid \mathcal{F}_s] = \mathbb{P}[X_t \in A \mid X_s]$.

<u>Discrete time / space</u>

$X_0, X_1, \dots$

- E.G. Random walk $X_i \in V$



- Card Shuffling, $X_i \in S_n$
  - pick random card, move to top

- MCMC: k-Colouring on $G=(V, E)$, $X_i \in [k]^V$

$X_i \to X_{i+1}$ - pick $v \in V$ u.a.r

- colour v uniformly from

  $[k] \setminus \{X_i(u)\}_{\{u : u \sim v\}}$

- Transition matrix

$$P(x, x') = \mathbb{P}[X_{i+1} = x' \mid X_i = x]$$

$$P^t_{x,x'} = \mathbb{P}[X_{t+i} = x' \mid X_i = x]$$

If $X_0 \sim \nu$ then

$$\mathbb{P}[X_t = y] = \sum_x \nu_x P^t_{xy} = (\nu P^t)_y$$

$X_t \sim \nu P^t$

If $\nu P = \nu$ then ν is stationary

When does $X_t \xrightarrow{d} \nu$, how quickly

- Irreducible: If

$\forall x, x' \ \exists t \ \ \mathbb{P}[X_t = x' \mid X_0 = x] > 0.$ 

If $GCD\{t : \mathbb{P}_x[X_t = x]\} = 1$ then x is aperiodic

- If X_t is irreducible + aperiodic it is ergodic.

$\exists T$ s.t. $\forall x, x'$ $\mathbb{P}_x[X_t = x] > 0$

Perron - Frobenius $\Rightarrow$ $\exists$ a stationary distribution

$$\mu_n = \frac{1}{n}\sum_{i=1}^{n} \mu P^t$$

then $\mu_n P = \sum_{i=2}^{n+1} \mu P^t = \mu_n + \underbrace{\frac{1}{n}(\mu P^{t+1} - \mu P)}_{\to 0}$

Take subsequential limit.

Let X_t be an ergodic M.C. Then

$X_t \xrightarrow{d} \mu$, $\quad d_{TV}(X_t, \mu) \le e^{-ct}$

Coupling $\quad X_0 = x, \quad Y_0 \sim \mu.$

$\exists T, \alpha$ such that $\mathbb{P}_i[X_T = j] \geq \alpha$.

X_t, Y_t independent.

Let $t_* = \min\{t : X_t = Y_t\}$

$$\tilde{X}_t = \begin{cases} X_t : & t \le t_* \\ Y_t : & t \ge t_* \end{cases}$$

$$X_t \overset{d}{=} \tilde{X}_t$$

$$d_{TV}(X_t, Y_t) \le \mathbb{P}[\tilde{X}_t \ne Y_t]$$

$$= \mathbb{P}[t_* > t]$$

$$\mathbb{P}[t_* > kT] \le (1-\alpha)^k.$$

Mixing Time

$$t_{mix}(\varepsilon) = \max_{x_0} \min_{t} d_{TV}(\mathbb{P}_{x_0}[X_t \in \cdot], \pi)$$

- Note $d_{TV}(X_t, \pi)$ is decreasing by coupling

- If $t_{mix}(\frac{1}{2e}) \le L$ then, $k \in \mathbb{N}$

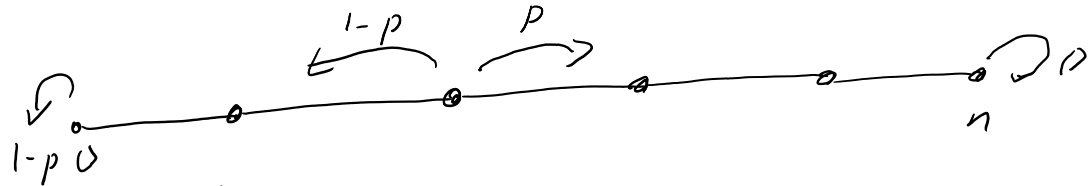
$$d_{TV}(X_{kL}, \pi) \le e^{-k}.$$

For any x, y

$$d_{TV}(\mathbb{P}^L_{x_0}, \mathbb{P}^L_y) \le d_{TV}(\mathbb{P}^L_x, \pi) + d_{TV}(\pi, \mathbb{P}^L_y)$$

- then argument above.

Biased Random walk on line



$p < \frac{1}{2}$.

- Ergodic ✓
- $\pi(i) = C \left(\frac{p}{1-p}\right)^i$
- Let $X_0 = x$, $Y_0 \sim \pi$, $Z_0 = n$

Coupling $W_t = \begin{cases} 1 & \text{w.p. } p \\ -1 & \text{w.p. } 1-p \end{cases}$

$X_{t+1} = X_t + W_t$ (unless $X_{t+1} \notin \{0, \ldots, n\}$ then stay put)

If $X_t, Y_t \le Z_t$

If $Z_t = 0$ then $X_t = Y_t$.

Let $\tau = \min\{t : Z_t = 0\}$

$$\mathbb{P}\left[\tau > \frac{n(1+\delta)}{1-2p}\right] \le \mathbb{P}\left[\sum_{t=1}^{\frac{n(1+\delta)}{1-2p}} Z_t > -n\right]$$

$$\le e^{-\zeta_\delta n} \quad \text{Azuma}$$

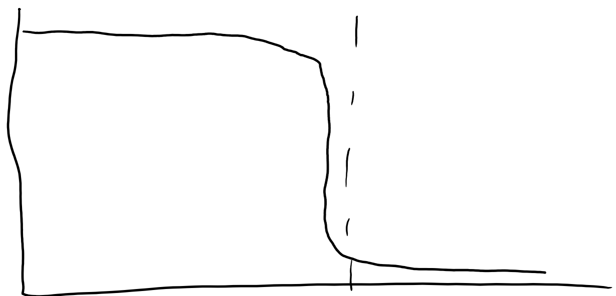
$t_{mix}(\varepsilon) \le \frac{n(1+\delta)}{1-2p}$

$$t_{mix}(\varepsilon) \le \frac{n(1+\delta)}{1-2p}$$

<u>Lower bound.</u>

$$\pi(\{0,\dots,\sqrt{n}\}) \to 1$$

$$\mathbb{P}\left[Z_{\frac{n(1-\delta)}{1-2p}} \le \sqrt{n} \right] \to 0$$



<u>Random walk on $\{0,1\}^n$</u>

- Not aperiodic.

- Lazy - w.p. $\frac{1}{2}$ do nothing

$$P^\alpha = \alpha I + (1-\alpha) P \quad - \ \alpha\text{-Lazy chain}$$

$$\pi P^\alpha = \alpha \pi I + (1-\alpha)\pi P = \pi.$$

Let $X_t = (1,\dots,1)$, $Y_t \sim \pi$.

Coupling: - Choose co-ordinate $I_t \in \{1,\dots,n\}$

- Choose $W_t \in \{0,1\}$ u.a.r.

- Update $X_{t+1}(I_t) = w_t$

$$X_{t+1}(J) = X_t(J) \quad J \neq I.$$

Same for Y_t.

- Coupled once all co-ordinates chosen

$\approx n \log n$ by coupon collector problem.

$t_{mix} \leq n \log n + Cn$

$d_{TV}(X_{(1+\varepsilon)\log n}, \pi) \to 0$

Lower bound

Let $S(X_t) = \sum_{i=1}^{n} X_t(i)$, $U_t = \#$ updated co-ordinates

$$S(X_t) \mid U_t \stackrel{d}{=} Bin(U_t, \tfrac{1}{2}) + n - U_t$$

$$\mathbb{E}\, S(X_t) = \frac{n}{2} + \frac{1}{2}(n - U_t)$$

$$\mathbb{E}(n - U_t) \leq n\, \mathbb{P}\{i \text{ not updated}\}$$

$$= n\left(1 - \frac{1}{n}\right)^t \approx n e^{-t/n}$$

$$Var[n - U_t] = Var[U_t]$$

$$\leq Var\left(\sum_{i=1}^{n} A_i(t)\right)$$

$$= n\, Var(A_i(t)) + \underbrace{(n-1)n\, Cov(A_1(t), A_2(t))}_{\text{neg correlation}}$$

$$\leq n\left(1 - \frac{1}{n}\right)^t$$

$$\leq n e^{-t/n}$$

$S(Y_t) \sim \mathrm{Bin}(n, 1/2)$

TV distance is large is

$$n - U_t \approx n e^{-t/n} \gg \sqrt{n}$$

$$t = \left(\tfrac{1}{2} - \varepsilon\right) n \log n$$

Test function

Let $A = \left\{ x : S(x) \le \frac{n}{2} + \underbrace{\frac{1}{2} n e^{-t/n}}_{\frac{1}{2} n^{\frac{1}{2}+\varepsilon}} \right\}$

$$\pi(A) \to 1.$$

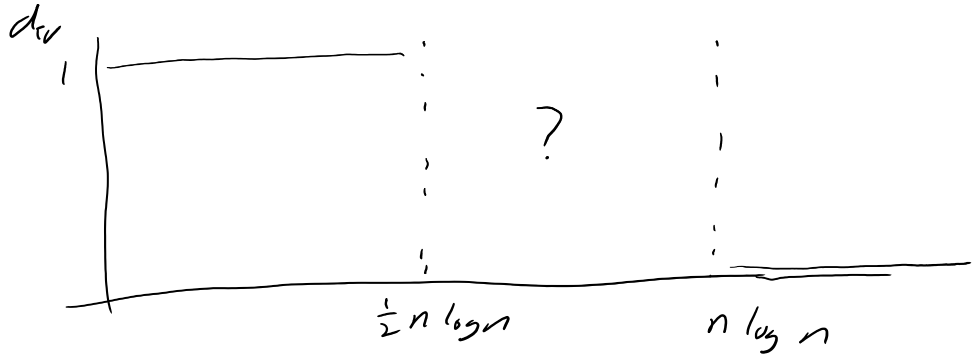
$$\mathbb{P}[X_t \in A]$$

$$= \mathbb{P}\left[S(X_t) \le \frac{n}{2} + \frac{1}{2} n e^{-t/n}\right]$$

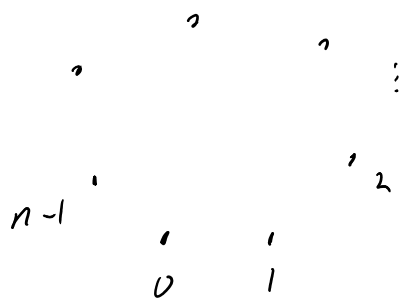
$$\le \mathbb{P}\left[|U_t - \mathbb{E}U_t| \ge \frac{1}{4} n^{\frac{1}{2}+\varepsilon}\right]$$

$$+ \mathbb{P}\left[\left|S(X_t) - \mathbb{E}[S(X_t) \mid U_t]\right| \ge \frac{1}{4} n^{\frac{1}{2}+\varepsilon}\right] \to 0$$

So $t_{mix} \leq (\frac{1}{2} - \varepsilon) n \log n$



Random Walk on the cycle



$\pi \sim$ uniform

$X_0 = 0, \quad Y_0 \sim \pi$

- Coupling: If $X_t = Y_t$ move together.

If $X_t \neq Y_t$ let $W_t = \pm 1$ w.p. $\frac{1}{2}$.

With probability $\frac{1}{2}$ $\quad X_{t+1} = X_t + W_t$

$Y_{t+1} = Y_t$

" " " $\quad X_{t+1} = X_t$

$Y_{t+1} = Y_t + W_t.$

.

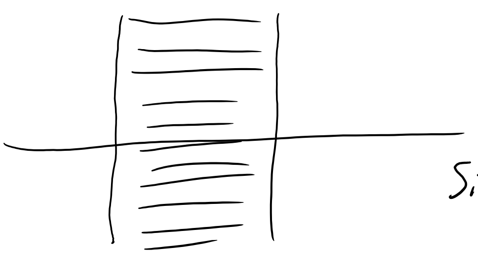
Now couple $Y_t - X_t$ and Z_t SRW on $\mathbb{Z}$

so that $Y_t - X_t = Z_t \pmod{n}$

$$\limsup_n \mathbb{P}[\, Z_{Cn^2} \in [1, n-1] \,] \le \mathbb{P}\left[0 \le N(0,1) \le \frac{1}{\sqrt{C}}\right]$$

So $t_{mix}(\varepsilon) \le Cn^2$.

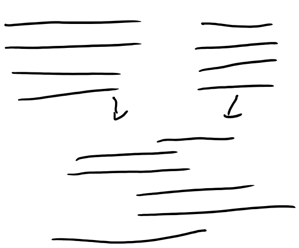
Riffle Shuffle Gilbert - Shannon - Reid



Split into two decks

Size $Bin(n, 1/2)$

- Interlace drop with probability

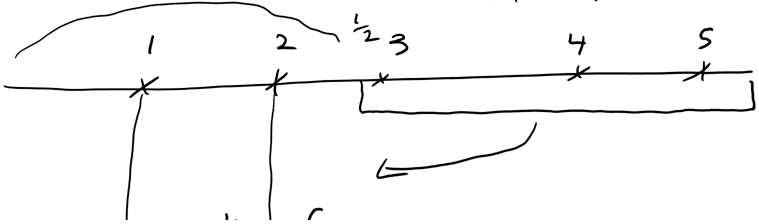


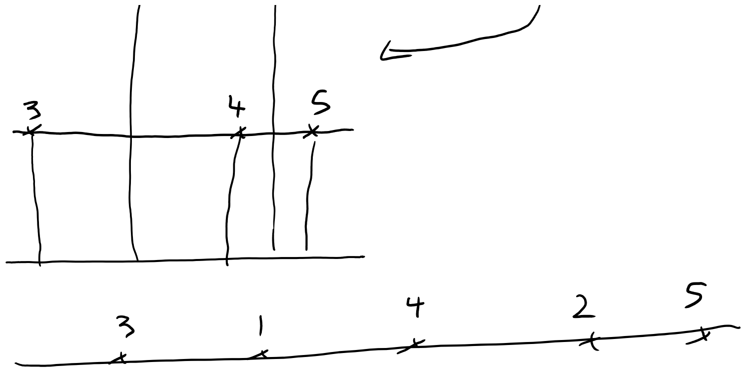
$$\frac{N_L}{N_L + N_R}$$

Can couple to following system

$U_1(0), \dots, U_n(0)$ IID $[0,1]$

Pick $Bin(n, \frac{1}{2})$ $\quad U_i(k+1) = 2U_i(k) \mod (1)$





Inverse Shuffle

Mark each card 0/1. IID

A	1
B	0
C	0
D	1
E	0
F	0
G	0

Move 0's to top

Keeping order the same.

B	1 0	C	0 0
C	0 0	G	0 0
E	1 0	D	0 1
F	1 0	B	1 0
G	0 0	E	1 0
A	1 1	F	1 0
D	0 1	A	1 1

Random when all distinct,

$\mathbb{P}[\text{string } i = \text{string } j] = 2^{-t}$

$2^{-t}\binom{n}{2} = o(1)$

$t_{mix} \le (2+\varepsilon)\log_2 n$

RW on group: $\mu(g)$ $g \in S_n$

$X_n = G_1 G_2 \dots G_n$

Inverse $Y_n = G_1^{-1} G_2^{-1} \dots G_n^{-1}$

$$\mathbb{P}[X_n = x] = \sum_{g_1, \dots, g_n :} I(g_1 \dots g_n = x) \prod \mu(g_i)$$

$$= \sum_{g_n \dots g_1} I(g_1 \dots g_n = x) \prod \mu(g_i)$$

$$= \sum_{g_n \dots g_1} I(g_n^{-1} \dots g_1^{-1} = x^{-1}) \prod \mu(g_i)$$

$$= \mathbb{P}[Y_n = x^{-1}]$$

$$d_{TV}(X_n, \pi) = \frac{1}{2} \sum_{x \in G} \left| \mathbb{P}(X_n = x) - \frac{1}{|G|} \right|$$

$$= \frac{1}{2} \sum_{x \in G} \left| \mathbb{P}(X_n = x^{-1}) - \frac{1}{|G|} \right|$$

$$= d_{TV}(Y_n, \pi)$$

Reversibility A Markov chain with S.D. π is reversible if

$$\forall x, y, \quad \pi(x) P(x,y) = \pi(y) P(y,x) \quad \text{Detailed Balance}$$

$$\Rightarrow \quad \mathbb{P}[X_1 = y] = \sum_x \pi(x) P(x,y) = \sum_x \pi(y) P(y,x)$$
$$= \pi(y)$$

$$(X_1, \dots, X_n) \overset{d}{=} (X_n, \dots, X_1)$$

Examples: Colouring, random walks

MCMC: Example Glauber dynamics:

- $(\sigma_1, \sigma_2, \dots, \sigma_n) \sim \pi$

Pick $i \in \{1, \dots, n\}$

Reset σ_i according to $\pi(\sigma_i \mid \{\sigma_j\}_{j \neq i})$

σ_{i^c}

I.E. set σ_i to x

w.p. $$\frac{\pi(\sigma_1, \dots, x, \dots, \sigma_n)}{\sum_{x'} \pi(\sigma_1, \dots, x', \dots, \sigma_n)}$$

Only need ratios not π itself.

- If $\underline{\sigma}$ uniform colouring then

- Check π is stationary

If σ, σ' differ only at i,

$$\pi(\sigma) P(\sigma, \sigma') = \pi(\sigma') P(\sigma, \sigma')$$

$$= \pi(\sigma_{i^c}) \cdot \pi(\sigma_i \mid \sigma_{i^c}) \cdot \frac{1}{n} \pi(\sigma_i' \mid \sigma_{i^c})$$

Metropolis Hastings

- Target distribution $\pi(\sigma)$

- Kernel $Q(\sigma, \sigma')$

$$A(\sigma, \sigma') = \max\left\{1, \frac{\pi(\sigma') Q(\sigma', \sigma)}{\dots}\right\}$$

$\pi(\sigma)\, Q(\sigma,\sigma')$

$$P(\sigma,\sigma') = A(\sigma,\sigma')\, Q(\sigma,\sigma') \qquad \sigma \neq \sigma'$$

Assume $A(\sigma,\sigma') = 1$, $A(\sigma',\sigma) = \dfrac{\pi(\sigma)\, Q(\sigma,\sigma')}{\pi(\sigma')\, Q(\sigma',\sigma)}$

$$\pi(\sigma)\, P(\sigma,\sigma') = \pi(\sigma)\, Q(\sigma,\sigma')$$

$$\pi(\sigma')\, P(\sigma',\sigma) = \pi(\sigma')\, Q(\sigma',\sigma) \cdot A(\sigma',\sigma)$$

$$= \pi(\sigma)\, Q(\sigma,\sigma') \quad \checkmark$$

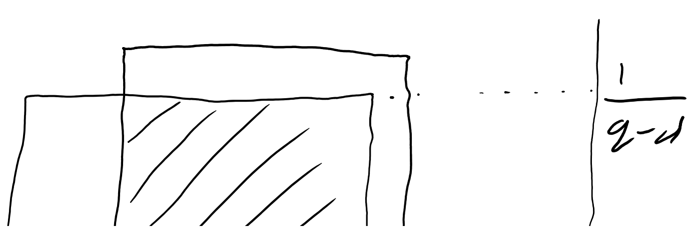
Need only know $\pi(\sigma)/\pi(\sigma')$.

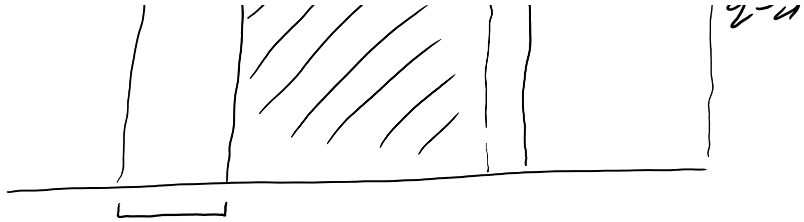
Coupling: Let G be a graph, d max degree, k colouring $k > 2d$.

Contraction: $X_t = x_0$, $Y_t \sim \pi$.

$$D_t = \{v : X_t(v) \neq Y_t(v)\}, \quad \rho_t = |D_t|$$

- Rule: Pick v, update same if possible
  - Probability we pick a disagreement $\frac{\rho_t}{n}$.
- S_v = prob new disagreement at v.





neighbors with disagreements

$$\mathbb{E}[\rho_{t+1} \mid \mathcal{F}_t] = \rho_t - \frac{\rho_t}{n} + \frac{1}{n} \sum_v s_v$$

$$\sum_{v \in V} s_v \le \sum_v \sum_{u \sim v} I(u \in D_t) / (q-d)$$

$$\le \frac{1}{q-d} \sum_u I(u \in D_t) \cdot \sum_{v : u \sim v} 1$$

$$\mathbb{E}[\rho_{t+1} \mid \mathcal{F}_t] \le \rho_t \left(1 - \frac{1}{n}\left(1 - \frac{d}{q-d}\right)\right)$$

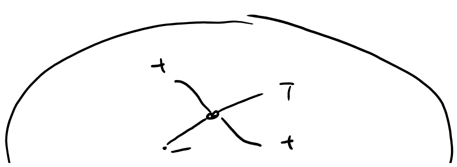
$$\rho_0 \le n \quad \text{so} \quad \mathbb{E}\rho_t \le n \exp\left(-\frac{t}{n}\left(1 - \frac{d}{q-d}\right)\right)$$

$$\mathbb{E}\rho_{cn\log n} \le n \cdot n^{-c\frac{q-2d}{q-d}}$$

$$c > \frac{q-d}{q-2d}.$$

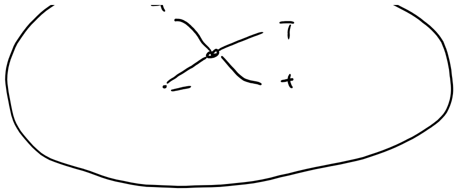
Ising model: $\sigma \in \{+1, -1\}^n$ $\quad \mathbb{P}[\sigma] = \dfrac{\exp(\beta \sum \sigma_i \sigma_j)}{Z}.$

Glauber Dynamics



$$\mathbb{P}[\sigma_v = +1 \mid \sigma_{v^c}]$$

$$= \frac{e^{\beta \sum_{i \sim v} \sigma_i \cdot 1}}{\cdots}$$



$$= \frac{e^{\beta \sum_{i \sim v} \sigma_i \cdot 1}}{e^{\beta \sum_{i \sim v} \sigma_i \cdot 1} + \exp^{\beta \sum \sigma_i \sim (-1)}}$$

$$= \varphi(\beta \textstyle\sum \sigma_i)$$

$$\varphi(x) = e^x / (e^x + e^{-x})$$

Single disagreement at v

$\rho_0 = 1.$

$$\rho_1 = 1 - \frac{1}{n} + \frac{1}{n} \sum_{u \in N_v} \max\Big(\mathbb{P}[\sigma_u = + \mid \{\sigma_w\}_{w \notin \{u,v\}}, \sigma_v = +]$$
$$- \mathbb{P}[\sigma_u = + \mid \quad '' \quad , \sigma_v = -] \Big)$$

$$\leq 1 - \frac{1}{n}(1 - d \tanh \beta)$$

$d \tanh\beta < 1.$ Actually $(d-1)\tanh\beta < 1$

$$\max_x \; \varphi(x) - \varphi(x - 2\beta) = \varphi(\beta) - \varphi(-\beta)$$
$$= \tanh\beta$$

Path Coupling

Suppose we can find a coupling on neighbouring states that contract:

Joint Distribution for all x, x' s.t.

$$Q_{x,x'}(y, y')$$

$= \mathbb{P}[X_1 = y, X_1' = y' \mid X_0 = x, X_0 = x']$

$\mathbb{E}[d(X_1, X_1') \mid X_0 = x, X_0' = x'] \le 1 - \alpha$

Coupling of initial states z, z'; $d(z, z') = k$.

Find $z = z_0, z_1, \ldots, z_k = z'$

such that $d(z_i, z_{i+1}) = 1$.

Want to find $Z_i \sim P_{z_i}$

So that $\mathbb{E}\, d(Z_0, Z_n) \le (1-\alpha) k$.

Choose $Z_0 \sim P_{z_0}$

$\mathbb{P}[Z_1 = \cdot \mid Z_0] = \mathbb{P}[X_1' = \cdot \mid X_0 = z_0, X_0' = z_1, X_1' = Z_0]$

$= Q_{z_0, z_1}(\cdot \mid Z_0)$

Check $\mathbb{P}(Z_1 = y']$

$= \sum_y \underset{\shortparallel}{\mathbb{P}[Z_0 = y]} \cdot \mathbb{P}(X_1' = y' \mid X_0 = z_0, X_0' = z_1, X_1 = y]$

$= \Sigma\, \mathbb{P}[X_1 = y] \cdot \mathbb{P}[\;\ldots\;]$

$= \mathbb{P}[X_1' = y' \mid X_0 = x, X_0' = x'] = P_{x', y'}.$

Set

$\mathbb{P}[Z_{i+1} = y' \mid Z_i] = Q_{z_i, z_{i+1}}(y' \mid Z_i)$

Each $(Z_{i-1}, Z_i) \sim Q_{Z_{i-1}, Z_i}$

$$\mathbb{E}\, d(Z_0, Z_k) \leq \sum_{i=1}^{k} \mathbb{E}\, d(Z_{i-1}, Z_i)$$

$$\leq k(1-\alpha)$$

So $\mathbb{E}\rho_1 \leq (1-\alpha)\rho_0$, $\mathbb{E}\rho_t \leq (1-\alpha)^t \rho_0$

If $n = \max_{x, x'} d(x, x')$

$$t_{mix}\left(\tfrac{1}{2e}\right) \leq \frac{1}{\alpha} \log\left(n/2e\right)$$

Back to Ising

$t_{mix} \sim C_\beta n \log n$ if $d \tanh \beta < 1$.

Approximating Partition function Z.

$\frac{1}{Z} \exp(\beta \sum \sigma_i \sigma_j)$, $\frac{1}{Z} I(\sigma \text{ colouring})$

Idea: $v_1, \dots, v_n$ verticies

$\frac{1}{Z} \exp(\beta |E|) = \mathbb{P}[\sigma_{v_1} = +, \dots, \sigma_{v_n} = +]$ - very rare

$= \mathbb{P}[\sigma_{v_1} = +] \cdot \mathbb{P}[\sigma_{v_2} = + \mid \sigma_{v_1} = +] \cdot \dots \cdot \mathbb{P}[\sigma_{v_n} = + \mid \sigma_{v_1}, \dots, \sigma_{v_{n-1}} = +]$

If for each i we can $\hat{p}_i$ so that

$$\mathbb{P}\left[\left|\frac{\hat{p}_i}{\mathbb{P}[\sigma_{v_i}=+\mid\sigma_{v_1},\dots,\sigma_{v_{i-1}}=+]}-1\right|>\frac{1}{n^2}\right]\le\frac{1}{n^2}.$$

Sample from $\sigma \mid \sigma_{v_1}\dots\sigma_{v_{i-1}}=+$ many times.

Spectral Properties

- Useful for R.W. on groups - representation theory
- Comparison
- Block dynamic

If P is reversible w.r.t. π.

$A = D_\pi^{1/2} P D_\pi^{-1/2}$ is symmetric.

$\exists\ U : U^T U = I,\ A = U^T \Lambda U$

$\exists\ 1 = \lambda_1 \geq \lambda_2 \cdots \geq \lambda_n \geq -1$ eigenvalues

$\lambda_* = \max_{i\geq 2} |\lambda_i| < 1$ if ergodic

$\gamma = 1 - \lambda_*$ gap $\quad t_{rel} = 1/\gamma.$

Eigenvectors φ_i of A, $\quad f_i = D_\pi^{-1/2}\varphi_i$

$$\frac{P^t(x,y)}{\pi(y)} = \sum_{j=1}^{n} f_j(x) f_j(y) \lambda_j^t$$

$$1 + \sum^{\hat{n}} f\ (x)\ f\ \ t$$

$$= 1 + \sum_{j=2}^{n} f_j(x) f_j(y) \lambda_j^t$$

- $d_{TV}(X_t, \pi) \sim C \lambda_*^t$

Inner product space

$$\langle f, g \rangle_\pi = \sum_x f(x) g(x) \pi(x)$$

f_j are orthonormal basis

Theorem: $t_{mix}(\varepsilon) \leq \log\left(\frac{1}{\varepsilon \pi_{min}}\right) t_{rel}$

$$d_{TV}(X_t, \pi) = \frac{1}{2} \sum_y |P^t(x,y) - \pi(y)|$$

$$= \frac{1}{2} \sum_y \left| \frac{P^t(x,y)}{\pi(y)} - 1 \right| \pi(y)$$

$$\left| \frac{P^t(x,y)}{\pi(y)} - 1 \right| \leq \sum_{j=2}^{n} |f_j(x) f_j(y)| \lambda_*^t$$

$$\leq \lambda_*^t \sqrt{\sum_{j=2}^{n} f_j(x)^2} \cdot \sqrt{\sum_{j=2}^{n} f_j(y)^2}$$

$$\delta_x = \sum f_j(x) \pi(x) f_j$$

so

$$\pi(x) = \langle \delta_x, \delta_x \rangle_\pi = \sum f_j(x)^2 \pi^2(x)$$

$$\pi(x)^{-1} = \sum f_j(x)^2$$

$| P^t(x,y) \quad .1 \qquad \lambda_*^t \qquad , t$

$$\left| \frac{P^t(x,y)}{\pi(y)} - 1 \right| \le \frac{\lambda_*^t}{\sqrt{\pi(x)\pi(y)}} \le \lambda_*^t / \pi_{min}$$

$$d_{TV}(X_t, \pi) = \frac{1}{2} \sum \left| \frac{P^t(x,y) - \pi(x)}{\pi(x)} \right| \cdot \pi(x)$$

$$\le \frac{1}{2} \lambda_*^t / \pi_{min} \le \varepsilon.$$

Note: For Ising $\log(\pi_{min}) \approx Cn$.

For small β, $\gamma \sim \frac{c}{n}$

- Lower bound rate of coupling
- upper bound $\frac{1}{n}$ from probability coupling / monotonicity
  $\mathbb{E} X_t^+ - \mathbb{E} X_t^- \ge (1 - \frac{1}{n})^t$

Lower Bound

$$t_{mix}(\varepsilon) \ge (t_{rel} - 1) \log(1/2\varepsilon)$$

Let f be eigenvector for λ_*, $\langle f, 1 \rangle = \sum \pi(y) f(y) = 0$

$$|\lambda_*^t f(x)| = |(P^t f)(x)|$$

$$= \Big| \sum_y P^t(x,y) f(y) - \underbrace{\pi(y) f(y)}_{\text{sums to zero}} \Big|$$

$$\le \|f\|_\infty \, d_{TV}(P_x^t, \pi)$$

Take $|f(x)| = \|f\|_\infty$,

$$\lambda_*^t \le d_{TV}(P_x^t, \pi)$$

Poincare Inequality $\|f\|_2^2 \le C\|\nabla f\|_2^2$

$$\mathcal{E}(f) = \langle (I-P)f, f\rangle_\pi = \sum \pi(x) f(x)^2 - \sum_x \sum_y \pi(x) f(x) P(x,y) f(y)$$

$$= \frac{1}{2}\sum_{x,y} \pi(x) P(x,y) \big(f(x) - f(y)\big)^2$$

$$1 - \lambda_2 = \inf \frac{\mathcal{E}(f)}{\mathrm{Var}_\pi f}$$ Can take $\mathbb{E}_\pi f = 0$.

If $f = \sum a_i f_i$ $\mathrm{Var}_\pi f = \sum_{i=2}^n a_i^2$

$$\mathcal{E}(f) = \langle (I-P)\sum a_i f_i, \sum a_i f_i\rangle$$

$$= \langle \sum (1-\lambda_j) a_i f_i, \sum a_i f_i \rangle = \sum_{i=2}^n (1-\lambda_j) a_i^2$$

$$\frac{\mathcal{E}(f)}{\mathrm{Var}_\pi(f)} \le 1 - \lambda_2$$

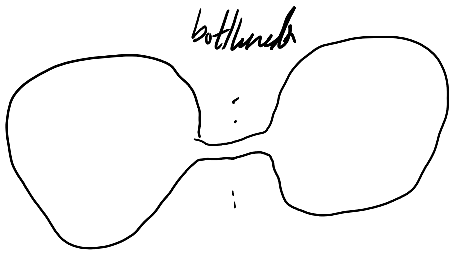
Comparison: If P, $\tilde{P}$ are two Markov chains with same S.D. π then

$$\frac{\tilde{\gamma}_2}{\gamma_2} \le \max_{x,y} \frac{\tilde{P}(x,y)}{P(x,y)}$$

E.G. Gap for Glauber & metropolis equivalent up to a constant.

Bottleneck Ratio / Conductance / Cheeger Constant

What is slow



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

$$Q(A,B) = \sum_{x \in A,\, y \in B} Q(x,y)$$

$$= \mathbb{P}_\pi[X_0 \in A, X_1 \in B]$$

$$\Phi(S) = \frac{Q(S,S^c)}{\pi(S)}$$

$$\Phi_* = \min_{S : \pi(S) \le 1/2} \Phi(S)$$

<u>Thm</u> $t_{mix}\left(\frac{1}{4}\right) \geq \frac{1}{4\Phi_*}$

<u>Proof:</u> Let $X_0 \sim \pi$

$$\mathbb{P}[X_0 \in S, X_t \in S^c]$$

$$\le \sum_{i=1}^{t} \mathbb{P}[X_{i-1} \in S, X_i \in S^c]$$

$$= t\, Q(S,S^c)$$

$$\mathbb{P}[X_t \in S^c \mid X_0 \in S] \le t\, \Phi_*$$

So $d_{TV}(\mathbb{P}[X_t \in \cdot \mid X_0 \in S], \pi)$

$$\geq |\mathbb{P}[X_t \in S^c \mid X_0 \in S] - \pi(S^c)|$$

$\geq \frac{1}{4}.$

<u>Thm</u> (Jerrum + Sinclair), (Lawler & Sokal))

$$\frac{1}{2}\Phi_*^2 \leq 1-\lambda_2 \leq 2\Phi_*$$

Upper bound, plug into

$$f(x) = \begin{cases} \pi(S^c) & x \in S \\ -\pi(S) & x \in S^c \end{cases}$$

<u>Lower Bound</u>:

$\pi\{f_2 > 0\} \leq \frac{1}{2}$ o.w. take $-f_2$.

Set $f = \max\{f_2, 0\}$.

<u>Claim</u> $(I-P)f(x) \leq \gamma f(x)$

<u>Case 1</u> $f(x)=0$, since $-Pf(x) \leq 0$.

<u>Case 2</u> $f(x)>0$,
- $If(x) = If_2(x)$
- $Pf(x) \geq Pf_2(x)$

So $(I-P)f(x) \leq (I-P)f_2(x) = \gamma f_2(x) = \gamma f(x)$

Hence since $f \geq 0$

$$\langle (I-P)f, f\rangle_\pi \leq \gamma \langle f, f\rangle_\pi$$

$$\frac{\mathcal{E}(f)}{\langle f, f\rangle_\pi} \geq \gamma$$ (note $\mathbb{E}_\pi f \neq 0$ so doesn't violate Dirchet form)