Today
- another example of coupling
- martingales

Last time:

approx counting $\Rightarrow$ approx sampling

Q: given Ω & distn $\pi = (\pi_1, \dots, \pi_{|\Omega|})$ over elts of Ω

how to approx sample from this distn?

Cool idea:

- design a MC whose stationary distn is π
- Show that it mixes in polynomial time

i.e. after $T = \text{poly}(n)$ steps $\Pr(X_T = j) \approx \pi_j \quad \forall j \in \Omega$

$\boxed{\text{Coupling}}$

simple & elegant approach to bounding mixing time

Given a MC P on Ω, a coupling is a MC on $\Omega \times \Omega$ defining stochastic process (X_t, Y_t) s.t.

① each of X_t & Y_t in isolation is faithful copy of MC

$Pr(X_{t+1} = z \mid X_t = x) = p_{xz}$

$Pr(Y_{t+1} = w \mid X_t = y) = p_{yw}$

② If $X_t = Y_t$ then $X_{t+1} = Y_{t+1}$

$\boxed{\text{Coupling Lemma}}$

Let (X_t, Y_t) be a coupling

Suppose $\exists\, T$ s.t. $\forall x, y$

$Pr(X_T \neq Y_T \mid X_0 = x, Y_0 = y) \leq \varepsilon$

Then $\tau(\varepsilon) \leq T$

$$\|D_1 - D_2\|_{TV} = \frac{1}{2} \sum_{x \in \Omega} |D_1(x) - D_2(x)| = \max_{A \subseteq \Omega} |D_1(A) - D_2(A)|$$

Graph coloring

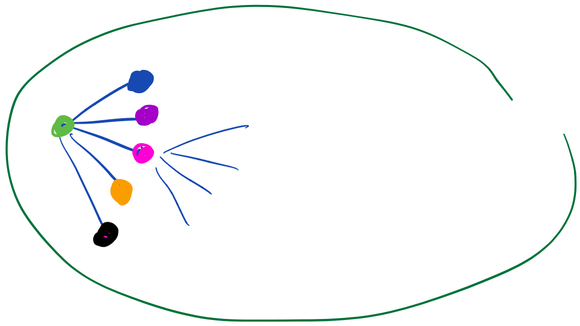
Input: $G=(V,E)$ undirected; max degree Δ. k colors

Recall if $k \geq \Delta+1$,
k-colorable

Markov chain:

① pick vertex v & color c u.a.r.
② recolor v with c if legal

- aperiodic
- irreducible:



- π uniform

Conjecture: $\forall \; k \geq \Delta+2$, this MC has poly mixing time
even $O(n \log n)$

Thm If $k \geq 4\Delta+1$, then MC has mixing time $O(n\log n)$

Coupling: X_t & Y_t choose same v & c each step

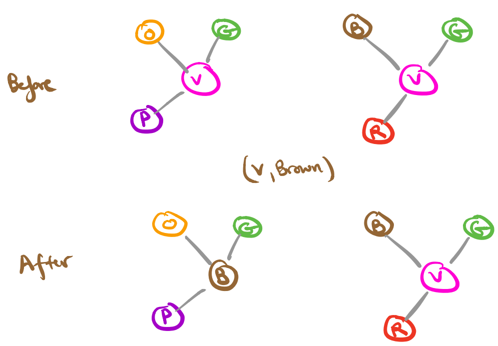
$D_t = \{ v \mid X_t \text{ \& } Y_t \text{ disagree on color of } v \}$ $\quad |D_t| = d_t$

$A_t = V - D_t$

here d_t may $\uparrow$

$\boxed{\text{Good moves } d_t \downarrow}$ $\quad g_t = \#\text{ good moves} \geq d_t(k-2\Delta$

$\boxed{\text{Bad moves } d_t \uparrow}$ $\quad b_t = \#\text{ bad moves} \leq 2d_t\Delta$



$E(X|Y=y)$ w.p. $Pr(Y=y)$

$E(E(X|Y))$

$= \sum_y E(X|Y=y) Pr(Y=y)$

$= \sum_y \sum_x x \, Pr(X=x|Y=y) Pr(Y=y)$

$= \sum_x x \sum_y Pr(X=x \cap Y=y)$

$= \sum_x x \, Pr(X=x)$

All other moves cause no change in d_t

Prob of each move $= \frac{1}{kn}$

$$\Rightarrow E[d_{t+1} \mid d_t] = d_t + \frac{b_t - g_t}{kn} \leq d_t + \frac{2d_t\Delta - d_t(k-2\Delta)}{kn}$$

$$= d_t\left(1 + \frac{4\Delta - k}{kn}\right) \leq d_t\left(1 - \frac{1}{kn}\right)$$

since $k \geq 4\Delta+1$

$$\Rightarrow E[d_t] = E[E[d_t \mid d_{t-1}]]$$

$\Rightarrow$ for $T = kn\ln\left(\frac{n}{\varepsilon}\right)$

$$\Pr\left(X_T \neq Y_T \mid X_0 = x, Y_0 = y\right) \leq \varepsilon \qquad \Rightarrow \tau(\varepsilon) \leq kn\ln\left(\frac{n}{\varepsilon}\right)$$

bound can be improved to $k \geq 2\Delta + 1$ with more clever coupling
see notes.

Martingales

A sequence of r.v.s $X_0, X_1, \ldots,$ is a martingale with respect to the sequence $Y_0, Y_1, \ldots$ if $\forall n \geq 0$ the following conditions hold:

- X_n is a fn of $Y_0, Y_1, \ldots, Y_n$
- $E(|X_n|) < \infty$
- $E(X_{n+1} \mid Y_0, \ldots, Y_n) = X_n$

Sequence of r.v.s $X_0, X_1, X_2, \ldots$ called a martingale when it is a martingale wrt itself

- $E(|X_n|) < \infty$
- $E(X_{n+1} \mid X_0, X_1, \ldots, X_n) = X_n$

Example: gambler plays sequence of fair games

Y_i: winnings on i^{th} game $\quad E(Y_i) = 0$

X_i = gamblers total winnings at end of i^{th} game

$$E(X_{i+1} \mid Y_1, \ldots, Y_n) = X_i + E(Y_{n+1}) = X_i$$

martingale regardless of amt bet each game, even if these amts are dependent on previous results

Examples

① Sums of indep random variables

$Y_0 = 0 \qquad Y_1, \ldots, Y_n$ iid w/ $E(Y_k) = 0 \quad \forall k$

Define $\qquad X_n = Y_0 + Y_1 + Y_2 + \cdots + Y_n$

$\{X_t\}$ is a martingale w.r.t. $\{Y_t\}$

$$E(X_{n+1} \mid Y_0, \ldots, Y_n) = E(X_n + Y_{n+1} \mid Y_0, \ldots, Y_n)$$

$$= E(X_n \mid Y_0, \ldots, Y_n) + E(Y_{n+1} \mid Y_0, \ldots, Y_n)$$

$$= X_n + E(Y_{n+1})$$

$$= X_n$$

② Variance of a sum

$Y_1, \ldots, Y_n$ iid w/ $E(Y_k) = 0 \quad \forall k \qquad E(Y_k^2) = \sigma^2$

Define $\quad X_0 = 0 \qquad X_n = \left(\sum_{k=1}^{n} Y_k\right)^2 - n\sigma^2$

$\{X_t\}$ is a martingale wrt $\{Y_t\}$

$$E(Y_{n+1} \mid Y_0, \ldots, Y_n) = E\left[\left(\sum_{k=1}^{n+1} Y_k\right)^2 - (n+1)\sigma^2 \,\Big|\, Y_0, \ldots, Y_n\right]$$

$$= E\left[Y_{n+1}^2 + 2Y_{n+1}\left(\sum_{k=1}^{n} Y_k\right) + \left(\sum_{k=1}^{n} Y_k\right)^2 - (n+1)\sigma^2 \,\Big|\, Y_0, \ldots, Y_n\right]$$

③ "Doob's" martingale process

$Y_1, Y_2, \ldots$ arbitrary seq of random vars

X r.v. with finite expectation

$X_n = E(X \mid Y_1, \ldots, Y_n)$ forms martingale wrt $\{Y_n\}$

$X_0 = E(X)$

$$E(V \mid W) = E\left[E(V \mid U, W) \mid W\right]$$

prove this ↑

$$E(X_{n+1} \mid Y_1, \ldots, Y_n)$$
$$= E\left(E(X \mid Y_1, \ldots, Y_{n+1}) \mid Y_1, \ldots, Y_n\right)$$
$$= E(X \mid Y_1, \ldots, Y_n) = X_n$$

Example: Edge exposure martingale

$G(n,p)$ random graph

label $m = \binom{n}{2}$ potential edges $e_1, e_2, \ldots, e_m$

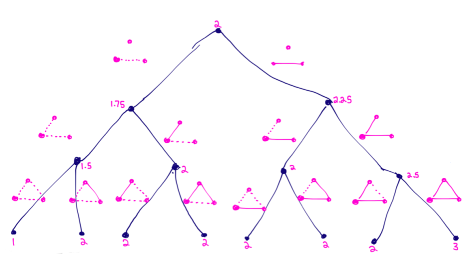
Let $f(G)$ be size of largest clique in G

$$Y_j = \begin{cases} 1 & \text{if edge } e_j \text{ present} \\ 0 & \text{otherwise} \end{cases} \qquad Pr(Y_j = 1) = p$$

$X_k = E[f(G) \mid Y_1, \ldots, Y_k]$ $\qquad X_0 = E[f(G)]$

$X_m = f(G)$

$G(n, \frac{1}{2})$



$X_0 = E(\text{max clique})$

$X_1 = E(\text{max clique} \mid Y_1)$

$X_2 = E(\text{max clique} \mid Y_1, Y_2)$

$X_3 = E(\text{max clique} \mid Y_1, Y_2, Y_3)$

Some useful facts about martingales:

1. $E(X_n) = E(X_0)$

by induction $\quad E(X_{n+1} \mid Y_0, \ldots, Y_n) = X_n$

$$\underbrace{E\left[E(X_{n+1} \mid Y_0, \ldots, Y_n)\right]}_{= E(X_{n+1})} = E(X_n)$$

$E(X \mid Y=y)$ w.p. $Pr(Y=y)$

$$E(\overbrace{E(X \mid Y)})$$
$$= \sum_y E(X \mid Y=y) Pr(Y=y)$$
$$= E(X)$$

2. Definition

A r.v. T is called a "stopping time" wrt $\{Y_t\}$ if

T takes values in $\{0,1,2,\ldots\}$

and if $\forall n \geq 0$, the event $\{T=n\}$ is determined by $Y_0, \ldots, Y_n$

i.e. can determine if $T=n$ or $T \neq n$ from knowledge of values $Y_0, \ldots, Y_n$

"know it when you see it"

$\boxed{\text{Optional Stopping Thm}}$

$\{Z_t\}$ is a martingale wrt $\{X_t\}$

For T a stopping time "know it when you see it"

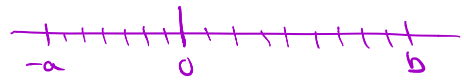
$$E(Z_T) = E(Z_0)$$

whenever any of the following hold

- Z_i's bounded $\quad (\exists c$ s.t. $\forall i \; |Z_i| \leq c)$
- T is bounded
- $E(T) < \infty$ and $\exists c$ s.t. $E\left(|Z_{i+1} - Z_i| \,\middle|\, X_1, \ldots, X_i\right) \leq c$

Applications of Optional Stopping Theorem

① unbiased r.w. on line starting at 0



$Y_i = \begin{cases} 1 & \text{with prob } \frac{1}{2} \\ -1 & \text{" " } \frac{1}{2} \end{cases}$ $\qquad X_n = \sum_{i=1}^{n} Y_i$ martingale

$T = \min \{ n \mid X_n = -a \text{ or } X_n = b \}$

T is a stopping time

Let $v_a = \Pr(X_n \text{ reaches } -a \text{ before reaching } b)$

By O.S.T. $\qquad E(X_T) = E(X_0) = 0$

② Same unbiased r.w. on line, same T

What is $E(T)$?

$Z_n = X_n^2 - n$ is a martingale [variance of a sum $E(Y_i^2) = 1$]

By O.S.T. $\qquad E(Z_T) = E(Z_0) = 0$

Same questions: biased r.w.

$$Y_i = \begin{cases} +1 & p \\ -1 & q \end{cases} \qquad p > q \; (=1-p)$$

$$\left.\begin{aligned} X_n &= \sum_{i=1}^{n} Y_i - n(p-q) \\ X_n' &= \left(\frac{q}{p}\right)^{\sum_{i=1}^{n} Y_i} \qquad X_0' = 1 \end{aligned}\right\} \text{ martingales wrt } \{Y_n\}$$

check it!

$$T = \min \left\{ n \;\middle|\; \sum_{i=1}^{n} Y_i = -a \text{ or } = b \right\}$$

$$v_a = \Pr\left(\sum_{i=1}^{} Y_i \text{ reaches } -a \text{ before } b\right)$$

$$E(X_T') = E(X_0') = 1$$

$$E(X_T') = v_a \left(\frac{q}{p}\right)^{-a} + (1-v_a)\left(\frac{q}{p}\right)^{b} = 1$$

$$\Rightarrow v_a = \frac{1-\left(\frac{q}{p}\right)^{b}}{\left(\frac{q}{p}\right)^{-a} - \left(\frac{q}{p}\right)^{b}}$$

$$E(X_T) = v_a(-a) + (1-v_a)b - E(T)(p-q) = 0$$

③ Tail inequalities

$E(X_n) = E(X_0)$ how far can it be from its expectation

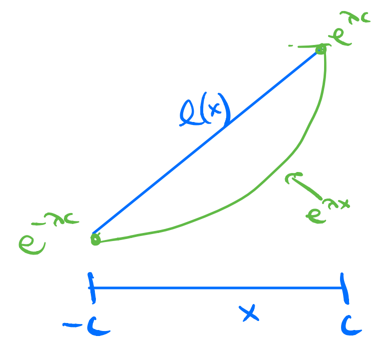
Azuma-Hoeffding Inequality

$X_0, \dots, X_m$ martingale s.t. $\forall k$ $|X_k - X_{k-1}| \le c_k$

c_k may depend on k

Then $\forall t \ge 0$, any $R > 0$

$$\Pr\left(|X_t - X_0| > R\right) \le 2 e^{-\left[\frac{R^2}{2\sum_{k=1}^{t} c_k^2}\right]}$$



$$\Rightarrow E\left[e^{\lambda X_{t+1}} \middle| H_t\right] = e^{\lambda X_t} E\left[e^{\lambda(X_{t+1}-X_t)} \middle| H_t\right]$$

$$\leq e^{\lambda X_t} e^{(\lambda c_t)^2/2}$$

take expectation on both sides

$$E\left[e^{\lambda X_{t+1}}\right] \leq e^{(\lambda c_t)^2/2} E\left[e^{\lambda X_t}\right] \underset{\uparrow}{\leq} e^{\lambda^2 \sum_{i=1}^{t} c_i^2/2}$$

by induction on t

Finally, $\Pr(X_t \geq R) = \Pr(e^{\lambda X_t} \geq e^{\lambda R}) \leq e^{-\lambda R} E\left[e^{\lambda X_t}\right]$

$$= e^{-\lambda R} e^{\lambda^2 \sum_{i=1}^{t} c_i^2/2}$$

Optimizing, we choose $\lambda = \frac{R}{\sum_{i=1}^{t} c_i^2}$

$$\Rightarrow \Pr(X_t \geq R) \leq e^{-R^2/2\sum_{i=1}^{t} c_i^2}$$

$$-\left(\lambda R - \frac{\lambda^2 \sum c_i^2}{2}\right) = -\left(\frac{R^2}{\sum c_i^2} - \frac{R^2}{2\sum c_i^2}\right)$$

Factor of 2 comes from $\Pr(X_t < -\lambda)$