Lecture 7

Applications of PATH COUPLING

- Path Coupling Setup:

  given a Markov Chain $P(\cdot,\cdot)$ on some state space Ω, need to engineer a good coupling to ~~exhibit~~ bound the mixing time of the chain

- Idea: rather than defining coupling step for all pairs of states only define it for states that are adjacent in the pre-metric

<u>Path Coupling theorem</u>: Suppose have coupling $(X,Y) \to (X',Y')$ whenever (X,Y) are adjacent in pre-metric, that satisfies

$$\mathbb{E}[d(X',Y') \mid X,Y] \leq (1-\alpha)\cdot d(X,Y) \qquad (*)$$

for some $\alpha \in [0,1]$, for all X,Y adjacent in pre-metric.

Then can extend coupling to all pairs of states so that $(*)$ is satisfied.

Application to Coloring

<u>Theorem</u>: the natural MC given in last lecture has $\tau_{mix} = O(n \cdot \log n)$ whenever $q \geq 3\Delta + 1$.

<u>Proof</u>: **pre-metric on Ω**: edge of weight 1 between proper colorings x & y if can recolor single vertex in x to get y

metric d: shortest path distance in pre-metric.

diameter: $\forall\, x,y \in \Omega$: $d(x,y) \leq 2n$ (why?)

coupling in pre-metric: (suppose v_0 is disagreeing vertex in states x, y and it has color c_x in x and c_y in y)

- suppose $X_t = x$, $Y_t = y$
- pick the same vertex v in both chains
- if $v \notin N(v_0)$ (i.e. $v = v_0$ OR v not a neighbor of v_0)

  pick same color c in both chains & recolor if possible
- if $v \in N(v_0)$ then match up colors as follows:

chain X_t		chain Y_t	
c_x	$\longleftrightarrow$	c_y	
c_y	$\longleftrightarrow$	c_x	
c	$\longleftrightarrow$	c	, $\forall c \neq c_x, c_y$

good moves: choose disagreeing vertex v_0 and color c not present in $N(v_0)$

↳ decrease distance to 0

$\geq q - \Delta$ good moves

$\boxed{\text{* may be pessimistic, see page 5}}$

bad moves: choose $v \in N(v_0)$, & choose color c_y in X_t, c_x in Y_t

↳ increase distance to at most 3

(since move is allowed and vertex v will then have different colors in X_{t+1}, Y_{t+1}.)

$\leq \Delta$ bad moves.

hence for $d(x,y) = 1$ the above coupling gives:

$$\mathbb{E}\left[d(X_{t+1}, Y_{t+1}) \mid X_t = x, Y_t = y\right] \leq \underbrace{\left(1 - \frac{q-\Delta}{qn} + \frac{2\Delta}{qn}\right)}_{1 - \frac{q-3\Delta}{qn} =: \alpha} d(X_t, Y_t)$$

→ So by thm can extend coupling to ~~arbitr~~ full coupling so that

$$\mathbb{E}\left[d(X_{t+1}, Y_{t+1}) \mid X_t, Y_t\right] \leq (1-\alpha) \cdot d(X_t, Y_t)$$

$$\rightsquigarrow \mathbb{E}\left[d(X_t, Y_t) \mid X_0 = x_0, Y_0 = y_0\right] \leq (1-\alpha)^t d(x_0, y_0) \leq (1-\alpha)^t \cdot 2n \leq \varepsilon$$

→ hence: $\Delta(t) \leq \Pr[X_t \neq Y_t \mid X_0 \neq Y_0] = \Pr[d(X_t, Y_t) > 0 \mid X_0, Y_0] \leq \Pr[d(X_t, Y_t) \geq 1 \mid X_0, Y_0] \leq \mathbb{E}[d(X_t, Y_t) \mid X_0, Y_0] \leq (1-\alpha)^t \cdot 2n \leq \varepsilon$

↳ used integrality of distances

$t = \frac{q-3\Delta}{q} \cdot n \left(\log(2n) + \right.$ $\log$

so $\tau_{mix} = O(n \cdot \log n)$.

∎

- with a little effort we can strengthen the previous theorem to $q \geq 2\Delta+1$ which is a weaker condition on the number of colors

Theorem: The natural MC of lecture 6 has $T_{mix} = O(n \cdot \log n)$ whenever $q \geq 2\Delta+1$.

Proof:
- Consider the set $\tilde{\Omega} \supseteq \Omega$ of all (not-necessarily proper) colorings.
- Consider the MC that picks a vertex v, color c u.a.r and changes the color of vertex v to c, if c does not appear in $N(v)$.

Important Observation
- Clearly, if we start this MC from a point in Ω (i.e. a proper coloring) then we never leave Ω and we are running a copy of the natural MC of lecture 6.
- $\tilde{\Omega}$ is just introduced for convenience as we see below.
- pre-metric on $\tilde{\Omega}$: x, y are adjacent if they differ at exactly one vertex in which case $d(x,y) = 1$.

  metric: hamming metric $d(x,y) = \#$ vertices where x, y disagree.

path coupling for x,y adjacent in pre-metric: Same as in previous proof (see page 2)

good moves: decrease distance to 0

$\geq q - \Delta$ good moves

bad moves: increase distance to 2 (rather than 3 that we had earlier)

$\leq \Delta$ bad moves

hence: $\mathbb{E}[d(X_{t+1}, Y_{t+1}) \mid X_t = x, Y_t = y, d(x,y) = 1] \leq \left(1 - \frac{q-2\Delta}{qn}\right) \cdot d(X_t, Y_t)$

Did not use irreducibility of the MC

coupling thm

$\Rightarrow$ $\mathbb{E}[d(X_{t+1}, Y_{t+1}) \mid X_t, Y_t] \leq \left(1 - \frac{q-2\Delta}{qn}\right) d(X_t, Y_t)$. $(*)$

.

- Suppose now $X_0, Y_0 \in \Omega$ but otherwise arbitrary

$$(*) \Rightarrow \quad \mathbb{E}\left[d(X_t,Y_t) \mid X_0,Y_0\right] \le \left(1-\frac{q-2\Delta}{qn}\right)^t d(X_0,Y_0) \le 2\cdot n\cdot\left(1-\frac{q-2\Delta}{qn}\right)^t$$

$$\Rightarrow \quad \Pr\left[X_t \ne Y_t \mid X_0,Y_0\right] \le \Pr\left[d(X_t,Y_t)\ge 1 \mid X_0,Y_0\right] \le \mathbb{E}\left[d(X_t,Y_t) \mid X_0,Y_0\right]$$

$$\le 2n\cdot\left(1-\frac{q-2\Delta}{qn}\right)^t$$

in original chain of lecture 6

Coupling lemma in original chain; true for any coupling of the evolution of X_t, Y_t ⊛

Hence: $$\Delta(t) \le \max_{x,y\in\Omega} \|P_x^t - P_y^t\|_{TV} \le \max_{x,y\in\Omega} \Pr\left[X_t \ne Y_t \mid X_0 = x, Y_0 = y\right] \le 2n\cdot\left(1-\frac{q-2\Delta}{qn}\right)^t$$

~~in modified chain using the~~

$$\le \frac{1}{2e} \quad \text{for} \quad t = O(n\log n)$$

$$\Rightarrow T_{mix} = O(n\cdot\log n)$$

$\boxtimes$

⊛ our observation in page 3 implies that since $(X_t)_t, (Y_t)_t$ start inside Ω they never leave Ω and hence the coupling produced by the path coupling theorem in the modified chain is also a valid coupling in the original chain; so the inequality is still true for that coupling

Some Historical Remarks:

- The conjecture $T_{mix} = O(n\log n)$ whenever $\Delta \ge q+2$ is still outstanding.
- but was recently shown for Δ-regular trees by Martinelli, Sinclair, Weitz 2006
- [Dyer & Frieze 2003] shows $T_{mix} = O(n\log n)$ for $\Delta = \Omega(\log n)$, girth $g = \Omega(\log\log n)$ and $q \ge \alpha^*\cdot\Delta$, $\alpha^* \approx 1.76$.
- [Dyer, Frieze, Hayes, Vigoda '05] show $T_{mix} = O(n\log n)$ for
  - $q \ge \alpha^*\cdot\Delta$ ($\alpha^* \approx 1.76$), $g \ge 5$, Δ = large enough constant
  - $q \ge \beta^*\cdot\Delta$ ($\beta^* \approx 1.49$), $g \ge 7$, Δ = large enoug constant.

- [Hayes, Vigoda '03] show $\tau_{mix} = O_{1/\varepsilon}(\log n)$ whenever $q \geq (1+\varepsilon)\cdot\Delta$, $\Delta = \Omega(\log n)$

- Best unconstraind result: a slightly more complicated chain has $\tau_{mix} = O(n\cdot\log n)$ as long as $q \geq \frac{11}{6}\cdot\Delta$.

Where does the constant $\alpha^* \approx 1.76$ come from in above results?

- Recall our path coupling proof for $q \geq 2\Delta+1$.
- used crude bound of $\geq q-\Delta$ for # good moves.
- ~~let~~ instead, let $A(X, v_0)$ = # available colors at v_0 in coloring X
  - if $A(X, v_0) > \Delta$, previous analysis goes through.
  - Gedanken Experiment: suppose nodes in $N(v_0)$ have random colors

    then $\mathbb{E}[A(X, v_0)] = q\cdot(1-\frac{1}{q})^{\Delta} \approx q\cdot e^{-\frac{\Delta}{q}}$

    hence $\mathbb{E}[A(X,v_0)] > \Delta$ provided $q\cdot e^{-\Delta/q} > \Delta$, which is true if $q > 1.76\Delta$...

- this idea is exploited in [DF '03] & [HV 05] to give $\tau_{mix} = O(n\cdot\log n)$, when $\Delta = \Omega(\log n)$, $g = \Omega(\log\log n)$, $q \geq 1.76\cdot\Delta$

Mixing Time from Path Coupling

→ Suppose have coupling $(X,Y) \to (X',Y')$ whenever (X,Y) are adjacent ~~in premetric~~ in premetric

s.t. $\mathbb{E}\left[d(X',Y') \mid X,Y\right] \le (1-\alpha)\cdot d(X,Y)$ $(*)$, $\alpha \in [0,1]$

→ Can extend couping to all ~~pair~~ (X,Y) so that $(*)$ is satisfied

→ If $\alpha > 0$ and d is integer-valued then

$$\tau_{mix} = O\left(\tfrac{1}{\alpha}\log D\right), \text{ where } D = \max_{x,y} d(x,y)$$

is maximum distance in the induced metric d

→ How about $\alpha = 0$?

ex (1pt): Show $\tau_{mix} = O\left(\frac{1}{\beta} D^2\right)$, where

$D = \max_{x,y} d(x,y)$ as before

d : is assumed integer-valued

and $\beta = \min_{X,Y\in\Omega} \mathbb{E}\left[\left(d(X',Y') - d(X,Y)\right)^2 \mid X,Y\right]$

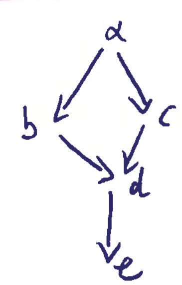
(clearly $\beta \ge \min_{X,Y} \Pr\left[\,|d(X',Y') - d(X,Y)| \ge 1 \mid X,Y\right]$)

Application 2: Linear Extensions of Partial Order

Input: partial order $\preceq$ on V (think of it as a DAG)

Linear extension: a total order $\sqsubseteq$ on V s.t. $x \preceq y \Rightarrow x \sqsubseteq y$, $\forall x, y$

e.g.



poset

$a \to b \to c \to d \to e$

linear extension

(in particular, easy to generate a linear extension via topological sorting)

Goal: Sample linear extension u.a.r.

Applications: combinatorics, sorting, rankings, decision theory...

[BW '91]: counting linear extensions is #P-complete
Brightwell & Winkler.

However, if can sample efficiently, can use this to count approximately

Markov Chain:
- w/prob $1/2$ do nothing
- else pick random position $p \in \{1, 2, \dots, n-1\}$ u.a.r.
- exchange elements at positions p & $p+1$ if resulting total order is a linear extension of $\preceq$

Aperiodic, Symmetric (easy)

Ex 1pt: Show irreducibility; in particular, every linear extension can be reached from another in $\binom{n}{2}$ steps

$\Rightarrow$ Stationary distn' is uniform over linear extensions.

Analysis of MC using path coupling:

pre-metric:
- $x, y \in \Omega$ are adjacent iff they differ at exactly two positions i, j $\quad 1 \le i < j \le n$
- distance is $j - i$

ex. 1pt: check that this is a pre-metric.

Coupling for adjacent pairs: let X, Y be adjacent in premetric and let i, j be the positions where they differ $1 \le i < j \le n$.

- Define coupling step $(X, Y) \to (X', Y')$ as follows

case 1: if $j \neq i+1$

- w. prob. $1/2$ do nothing in both X, Y
- else choose $p \in \{1, 2, \dots, n-1\}$ u.a.r. and try to exchange elements at positions p and $p+1$ in each of X, Y

case 2: if $j = i+1$.

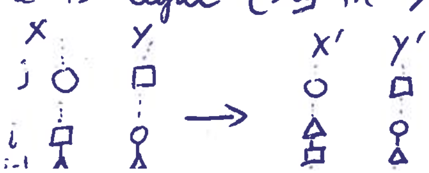
- w. prob. $\frac{1}{2(n-1)}$, do nothing in X, pick $p = i$ in Y and exchange elements at positi[ons] p & $p+1$ in Y
- w. prob. $\frac{1}{2(n-1)}$, do nothing in Y, pick $p = i$ in X and do the exchange. (note: this is allowed — // —)
- w pr. $\frac{n-2}{2(n-1)}$, do nothing in both
- else choose $p \in \{1, 2, \dots, n-1\} \setminus \{i\}$ uniformly at random and try to exchange elements at positions p & $p+1$ in each of X, Y

Analysis of coupling: in any move at most one position p is chosen.

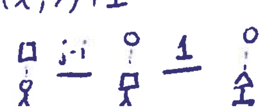
- **case 1**: $p \notin \{i-1, i, j-1, j\}$ $\rightsquigarrow$ w pr. 1: $d(X', Y') = j - i = d(X, Y)$
  (because either both exchanges legal or both — // — illegal & X', Y' differ at j, i)
- **case 2**: $p = i-1$ OR $p = j$

case 2 happens w. pr. $2 \times \frac{1}{2(n-1)}$ & distance satisfies $d(X', Y') \le d(X, Y) + 1$

- cases are symmetric so consider $p = i-1$.
- if exchanges both legal, then $d(X', Y') = j - (i-1) = d(X, Y) + 1$
- if — // — — // — illegal, then $d(X', Y') = d(X, Y)$
- if only one is legal (say in X) then



$d(X', Y') \le d(X, Y) + 1$

since in premetric 

- **Case 3**: $p = i$ OR $p = j-1$, distinguish subcases

in this subcase, prob. of case 3 is $\frac{1}{n-1}$

- $j-i=1$: always $d(X',Y') = 0 = d(X,Y) - 1$
- $j-i>1$: by symmetry consider only case $p=i$

  exchange always legal, hence $d(X',Y') = j-(i+1) = d(X,Y)-1$

in this subcase, prob. of case 3 is also $\frac{1}{n-1}$

Hence $d(X',Y') \leq d(X,Y)+1$ w. pr. $\frac{1}{n-1}$

$d(X,Y)-1$ w.pr. $\frac{1}{n-1}$

$d(X,Y)$ w.pr. $1 - \frac{2}{n-1}$.

$\Rightarrow \mathbb{E}[d(X',Y') \mid X,Y] \leq d(X,Y)$.

hence $\alpha = 0$

ex 1pt: $\beta \geq \max_{X,Y\in\Omega} \Pr[|d(X',Y') - d(X,Y)| \geq 1] \geq \frac{c}{n}$ for some $c>0$

$\Rightarrow$ from exercise of page 6 $\quad T_{mix} = O\left(\frac{1}{\beta} D^2\right) = O(n^5)$.

from above exercise and ex. of page 7

Remark: D. Wilson showed tight bound is $O(n^3 \log n)$

Improving T_{mix}: Change MC to the following:

- w pr. $1/2$ do nothing
- else pick $p \in \{1, 2, .., n-1\}$ wpr. $\frac{Q(p)}{Z}$ where $Q(p) = p(n-p)$
- exchange elements at positions p & $p+1$ if this is legal.

$$Z = \sum_{i=1}^{n-1} Q(p) = \frac{n^3-n}{6}$$

Same pre-metric, path coupling (with modified probabilities) gives:

$$\mathbb{E}[d(X',Y') - d(X,Y) \mid X,Y] \leq \frac{1}{2Z}[Q(i-1) + Q(j) - Q(i) - Q(j-1)]$$

$$\leq -\frac{6}{n^3} d(X,Y).$$

$\Rightarrow T_{mix} = O\left(\frac{1}{\alpha} \log D\right) = O(n^3 \log n)$