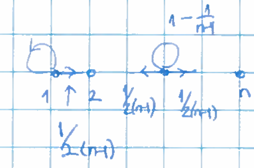
TÉCNICO LISBOA

Consider random adjacent transpositions:

$$S = \{(a\ a{+}1),\ a = 1, 2, \dots, n-1\} \cup \{id\}$$

$$P(x,y) = \begin{cases} \frac{1}{2} & y = x \\ \frac{1}{2(n-1)}, & y = x\,(a, a{+}1) \\ 0 & \text{otherwise} \end{cases}$$

Can project this to a much smaller Markov chain by looking at what happens to the first card. This has only n states rather than $n!$



The eigenvectors and eigenvalues are similar to the example about the circle:

$$\varphi(k) = \cos\frac{\pi(2k-1)}{2n} \quad \text{w/ eigenvalue} \quad \frac{n-2}{n-1} + \frac{1}{n}\cos\left(\frac{\pi}{n}\right)$$

$$\varphi_k(\sigma) = \cos\frac{\pi(2\sigma(k)-1)}{n} \qquad \sigma(k) = \text{card at position } k.$$

We could have considered any other card than the first. Let

$$\Phi(\sigma) = \sum_{k=1}^{n/2} \varphi(k)\,\varphi_k(\sigma)$$

$$\Phi(id) = \sum_{k=1}^{n/2} (\varphi(k))^2 = \sum_{k=1}^{n/2} \cos^2\frac{\pi(2k-1)}{2n} \geq \frac{n}{4}$$

If we perform $(a, a{+}1)$ ~~$\Phi(\sigma(a,a+1))$~~

(all other terms cancel)

$$\left|\Phi(\sigma \circ (a, a{+}1)) - \Phi(\sigma)\right| = \varphi(a)\varphi(\sigma(a{+}1)) + \varphi(a{+}1)\varphi(\sigma(a)) - \varphi(a)\varphi(\sigma(a)) - \varphi(a{+}1)\varphi(\sigma(a{+}1))$$

$$= |\varphi(a) - \varphi(a{+}1)|\ \underbrace{|\varphi(\sigma(a)) - \varphi(\sigma(a{+}1))|}_{\leq 2 \text{ (difference of cosines)}}$$

$$|\varphi(a) - \varphi(a{+}1)| = |\varphi'(\xi)| \leq \frac{\pi}{n}$$

Hence $\mathbb{E}\left((\Phi(\sigma_1) - \Phi(\sigma_0))^2 / \sigma_0\right) \leq \frac{1}{2(n-1)} \frac{2\pi}{n} \sim \frac{1}{n^2} = R$ in Wilson's Theorem.

Using next term in Taylor series of cosine we see that $\lambda \approx 1 - \frac{d}{n^3}$ for $\frac{1}{n}\cos\frac{\pi}{n} = 1 - \frac{\pi}{n^2}$

$$t \geq -\frac{1}{2\log\left(1 - \frac{d}{n^3}\right)}\left[\log\frac{\frac{1}{n^3}n^2}{n^2} + \log\varepsilon\right] \approx \frac{n^3}{2}\log n + n^3\log\frac{\varepsilon}{\pi}$$

<u>Lecture 5</u>: Recall the random walk on the cube $(\mathbb{Z}/2\mathbb{Z})^n$ $\quad P(x,y) = \begin{cases} \frac{1}{2}, & y = x \\ \frac{1}{2n}, & y = x + e_i \\ 0 & \text{otherwise} \end{cases}$

We struggled to produce a lower bound for a random walk. Today we want to show:
If $t = n\log n - cn$ then $s(t) \geq 1 - \frac{1}{e^c}$

Start from $\vec{0}$. It makes sense that $\vec{1}$ is the hardest vertex to hit. Let T = first time that all coordinates have been selected

This is a strong stationary time. — stationary distribution = uniform.

Recall this means $\mathbb{P}_0(X_t = x \mid T \le t) = \frac{1}{2^n}$

$$\mathbb{P}_0(X_t=\vec{1}) = \mathbb{P}_0(X_t=\vec{1} \mid T\le t)\,\mathbb{P}(T\le t) + \underbrace{\mathbb{P}_{\vec{0}}(X_t=\vec{1}\mid T>t)}_{0}\,\mathbb{P}(T>t)$$

$$= \frac{1}{2^n} P(T\le t)$$

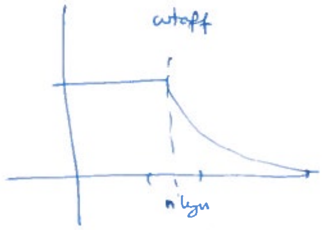
$$S(t) = \max_x \max_y \left\{1 - 2^n P_x^t(y)\right\} \underset{y=\vec{1}}{\overset{x=\vec{0}}{\geq}} 1-2^n P_0^t(\vec{1}) = 1-P(T\le t) = P(T>t)$$

$$S(t) \ge P(T > t) \ge 1 - \frac{1}{e^c}$$

if $t = n\log n - cn$ (from the first lecture)

Now the fact that T is a ~~SS~~ strong stationary time also means $S(t) \le P(T>t)$

So this is a miracle where $S(t)$ exactly equals the coupon collecting problem.



Theorem: a) If $t = n\log n + cn$ then $s(t) \le e^{-c}$

b) If $t = n\log n - cn$ then $s(t) \ge 1 - e^{-c}$

If we made the same thing for total variation distance there would be a factor of 2

Theorem: $d(t)$ = TV distance

a) If $t = \frac{1}{2} n\log n + cn$ $\quad d(t) < e^{-c}$

b) If $t = \frac{1}{2} n\log n - cn$ $\quad d(t) \sim 1 - \frac{1}{e^{2c}}$

Coupling techniques to estimate the total variation distance

Random-to-top model: n cards. Pick card at random. Move to top.

X_t on X Markov chain with transition matrix P. π stationary measure

Let Y_t be another copy of the Markov chain.

X_t starts at x, Y_t starts at y $\quad$ X_t, Y_t are not necessarily independent

X_0: 1, 2, 3, 4, …, n $\qquad$ Y_0: 3, 2, 1, 4

Card 3, move to top, can do the same for Y

3, 1, 2, 4 $\qquad$ 3, 2, 1, 4 $\qquad$ so these are correlated (3 is on top) but the elements in the graph are different.

Question: When will $X_t = Y_t$? The first time we touch all the cards, the two decks will be the same. in fact need only n-1 and that is the sharpest estimate

Goal: $\mathbb{P}(X_t = Y_t) \to 1$ as $t \to \infty$ and once X_t and Y_t meet, we force them to stay together (this will be true for the rule above but will also be true for ~~[illegible]~~ other rules) let $T_{couple} = \min\{t : X_t = Y_t\}$ (for any given coupling)

We want to show that $\boxed{d_{TV}(t) \leq \mathbb{P}(T_{couple} > t)}$

$$d_{TV}(t) = \max_{x \in X} \max_{A \subseteq X} |P^t_x(A) - \pi(A)|$$

$$P^t_x(A) = \mathbb{P}_x(X_t \in A)$$

$Y_t \sim \pi$ uniform distribution ($\pi' = \pi$)

$$\pi(A) = \mathbb{P}(Y_t \in A)$$

$$\mathbb{P}(X_t \in A) - \mathbb{P}(Y_t \in A) \leq \mathbb{P}(X_t \in A, Y_t \notin A) \leq \mathbb{P}(X_t \neq Y_t) = \mathbb{P}(T_{couple} > t)$$

Taking max over A we see that we have the required inequality.

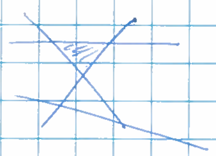
For random-to-top if T = first time that $(n-1)$ different cards have been picked. $d(t) \leq P(T > t)$.

<u>Hyperplane arrangements - random walks</u>: What we saw for the hypercube applies in this setting.

A hyperplane arrangement is a finite collection of hyperplanes in $\mathbb{R}^n$. let $m = |A|$ be the number of hyperplanes.

<u>Chambers</u> ~~are~~ of the arrangement are the connected components

faces = lower dimensional factors



~~Assign~~ X_t will be a process on the chambers. Assign weights $w_i \geq 0$ to the faces $\sum_i w_i = 1$

~~move~~ pick a face and move to adjacent chamber

Given $H \in \mathcal{A}$, it cuts $\mathbb{R}^n$ into 2 half spaces. The + and the −. A chamber is therefore determined by a vector on $\{+,-\}^n$ (some condition on the arrangement may be necessary).

Not all combinations are possible, of course.

Faces are also vectors in $\{0,+,-\}^m$ (0 means that the face lies on the hyperplane)

$\mathcal{F}$ = set of faces

$\mathcal{C}$ = set of chambers

$$\mathcal{F} \times \mathcal{C} \longrightarrow \mathcal{C}$$

$$(F \cdot C)_H = \begin{cases} H\text{-coordinate of } F & \text{if that is nonzero} \\ H\text{-coordinate of } C & \text{otherwise} \end{cases}$$

H-coordinate of the product chamber

one can check this is well defined and gives a nice associative structure.

Can also define $(F_1 \cdot F_2)_H = \begin{cases} H\text{-coordinate of } F & \text{if} \neq 0 \\ H\text{-coordinate of } C & \text{otherwise.} \end{cases}$

Start at C_0. Pick a face F with probability W_F. Move to $F \cdot C_0$.

The eigenvalues of this process have been described

<u>Coupling</u>:

Pick the same F for both processes X_t, Y_t

$$\begin{array}{ccc} C_0 & & \tilde{C}_0 \\ \downarrow & & \downarrow \\ F_1 C_0 & & F_1 \tilde{C} \\ \downarrow & & \downarrow \\ F_2 F_1 C_0 & & F_2 F_1 C_0 \\ \vdots & & \vdots \end{array}$$

(they must agree on the coordinates which are nonzero on F_1)

Once the product of the faces picked is a chamber we have coupled $X_t = Y_t$

Theorem (Athanasiadis, Diaconis)

a) For any $\mathcal{A}$, if T = first time that the product of the faces picked is a chamber then

$$d_{TV}(t) \leq \mathbb{P}(T>t)$$

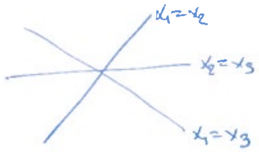
(this is a robust unifying what we have been talking about)

b) If $\mathcal{A}$ is central then $s(t) \geq \mathbb{P}(T>t)$ [Nestoridi]

Let's see why the random to top comes from such a hyperplane arrangement:

Braid arrangement:

$\{x_i = x_j\} \subset \mathbb{R}^n$



Chambers $\cong S_n$

$x_1<x_2<x_3$

$x_2<x_1<x_3$

$\vdots$

Faces - ordered block partitions

$\{123\}\{46\}\{57\}$ (order of the blocks matter)

$x_1=x_2=x_3 < x_4=x_6 < x_5=x_7$

Random-to-top assigns weights to the faces corresponding to $\{i\} \cup \{1,\dots,n\}\setminus\{i\}$

$x_i < x_1 = \dots = x_{i-1} = x_{i+1} = \dots = x_n$

Homework: what faces are used in the inverse riffle shuffle (this is also a random walk on a hyperplane arrangement)