Study Group - Lect #4

Mixtures of Permutations [Mao, Wu, 2020].

$\underline{\pi \in S_n}$

1) Perm. matrix
2) Clique (directed) $\rightsquigarrow$ Tournaments.
3) ...

statistical models on permutations.

<u>Motivation</u>

<u>Netflix</u> N movies $\to$ $[\underbrace{m1 > m2 > m3 > m5 > m4}]$

<u>Mallows model</u> (π_0, ϕ)

$\pi_0 \in S_n$ central ranking

distance

$$\Pr[\sigma \mid \pi_0, \phi] \propto \phi^{KT(\pi_0, \sigma)}$$

$0 < \phi < 1$

Physics
$\phi = e^{-\beta}$, $\beta > 0$
inverse temperature.

- π_0 : 1 2 3 4 5
- π_1 : 2 1 3 4 5
- π_2 : 5 4 3 2 1

$KT(\pi, \sigma) =$ # steps bubblesort should make starting from π ending at σ.

$$KT(\pi, \sigma) = \sum_{i<j} \mathbb{1}\{(\pi(i) - \pi(j))(\sigma(i) - \sigma(j)) < 0\}$$

if π is closer to π_0 than σ then $\Pr[\pi \mid \pi_0, \varphi] > \Pr[\sigma \mid \pi_0, \varphi]$.

$$\Pr[\pi \mid \pi_0, \varphi] \propto \phi^{KT(\pi_0, \pi)}$$

Mixtures of Mallows

$$\mathcal{M} = \sum w_i \, \mathcal{M}(\pi_0^{(i)}, \underline{\phi})$$

Retrieve $\underline{\{\pi_0^{(i)}\}}$. $\boxed{(w_i)}$

Identifiability

Zagier

$$A(\phi) \in \mathbb{R}^{n! \times n!} = \begin{array}{c} \pi \\ \left[\quad \phi^{KT(\pi, \sigma)} \quad \right]_{\pi, \sigma} \end{array}$$

PMF of $\mathcal{M}(\pi, \phi)$

symmetric ✓

$$\det(A(\phi)) \neq 0 \quad \forall \phi \in (0,1)$$

$$\mathcal{M} = \sum w_i \, \mathcal{M}(\pi_i, \phi)$$

Draw $\sigma_1, \sigma_2, \ldots, \sigma_N \longrightarrow$ Can you learn (π_i)?

Try something simpler:

$$\boxed{\delta_\pi}$$

Oracle

output

Input

(i, j)

(1, 2)

π hidden

12345

$\longrightarrow$ 1 $= \mathbb{1}\{\pi(i) < \pi(j)\}$

Algo: Quicksort.

Example 2: $\frac{1}{2}\left(\delta_{\pi_1} + \delta_{\pi_2}\right)$

Input

(i, j)

$\longrightarrow$ π_1 ?

π_2 ?

Pairwise Comp. Oracle is not sufficient for identifying M.

$\pi_1 = \pi_2^{-1} \implies \mathrm{Be}\left(\frac{1}{2}\right)$

Example 3 : You can compare two pairs simultaneously

Input $\quad \pi \sim M \quad$ Outputs $\qquad$ strengthen the oracle.

$$\begin{bmatrix} (\alpha, \beta) \\ (\gamma, \delta) \end{bmatrix} \longrightarrow \boxed{\pi} \longrightarrow \begin{bmatrix} \mathbb{1}\{\pi(\alpha) < \pi(\beta)\} \\ \mathbb{1}\{\pi(\gamma) < \pi(\delta)\} \end{bmatrix}$$

$\frac{1}{2}(\delta_{\pi_1} + \delta_{\pi_2})$ where $1 <_{\pi_1} 2$

$1 >_{\pi_2} 2$

$$\left[\begin{pmatrix} \boxed{(1,2)} \\ (a,b) \end{pmatrix} \right. \longrightarrow$$

certificate.

↓

you learn who answers

---

oracle

Group of pairwise comparisons, size m

$\pi \sim M$ for this round

$$m\begin{cases} (1,2) \\ (4,7) \\ \vdots \\ (17,42) \end{cases} \longrightarrow \begin{bmatrix} (i_1, j_1) \\ \vdots \\ (\end{bmatrix} \longrightarrow \boxed{\pi} \longrightarrow \left\{ \begin{bmatrix} 0,1 \\ 0,1 \\ \vdots \\ 0,1 \end{bmatrix} \right\} \quad \chi(\pi, I) \in \{0,1\}^m$$

I

$\pi \sim M$

$$\begin{bmatrix} (1,2) \\ (2,3) \\ (17,42) \end{bmatrix} \longrightarrow \boxed{\pi_{123}} \longrightarrow \begin{bmatrix} \ \end{bmatrix} \quad \chi(\pi_{123}, I)$$

This is the (strong) oracle.

Noiseless case ($\phi = 0$)

$$M = \sum_{i=1}^{k} w_i \, \delta_{\pi_i} \qquad \pi_i \in S_n$$

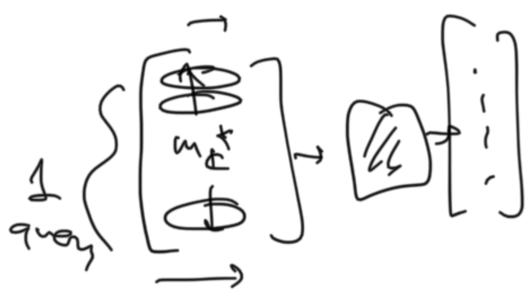
Q: Is there a poly-time algo that learns the k-mixt M from a group of poly-sized m (pairwise) comparisons for any large n where m depends only on k

Thm: $m = ???$

$$m = f(k) = \lfloor \log_2 k \rfloor + 1$$

Let $m_k^* = \lfloor \log_2 k \rfloor + 1$.

For any $M = \sum_{i=1}^{k} w_i \, \delta_{\pi_i}$, $\exists$ poly(n,k)-time algo that recovers M from gps of m_k^* pairwise comps with



$O\left(\frac{k}{2} (n-2)(n+1) \right)$ adaptive queries.

(to the weak oracle)

Key lemma:

Key lemma:

Give me k perms $\Sigma = \{\pi_1, \dots, \pi_k\}$ in S_n

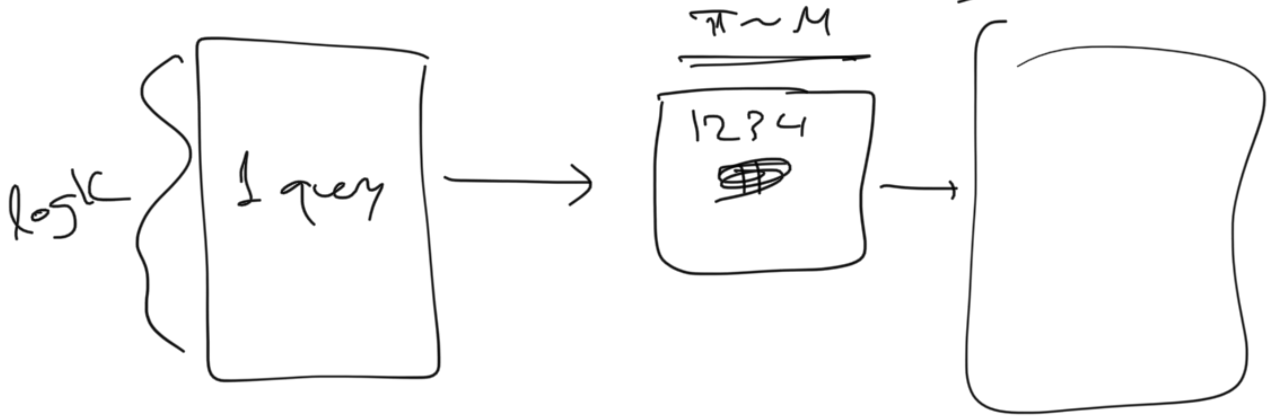
$\exists\ \pi^* \in \Sigma$ and I of ℓ pairs
$\ell = O(\log k)$

s.t.

$$\underbrace{\chi(\pi^*, I)} \neq \chi(\pi, I) \quad \forall\, \pi \in \Sigma \setminus \{\pi^*\}$$

$\pi \sim M \leftarrow \frac{1}{2}(1234) + \frac{1}{3}(2134) + \frac{1}{\cdot}(\ )$

k

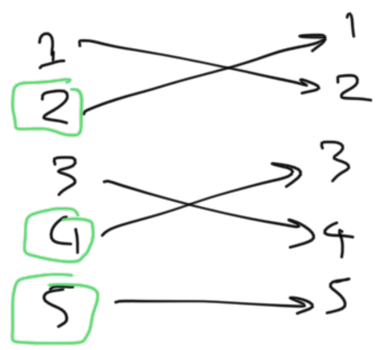


Decomposing Mallows

$J \subset [n]$

$\pi|_J$ $\qquad$ π $\qquad$ J

$\pi\|_J$



$\pi|_J$: injection (1-1)

$\pi|_J(2) = 1$
$\pi|_J(4) = 3$
$\pi|_J(5) = 5$ ✓

$|J|$

$\pi\|_J \in S_{|J|}$

$\pi\|_J(2) = 1$
$\pi\|_J(4) = 2$
$\pi\|_J(5) = 3$

$\in S_3$

Subroutine (J) $J \subset [n]$

Given $\sigma_1, \dots, \sigma_N \sim M = \sum w_i M(\pi_i)$

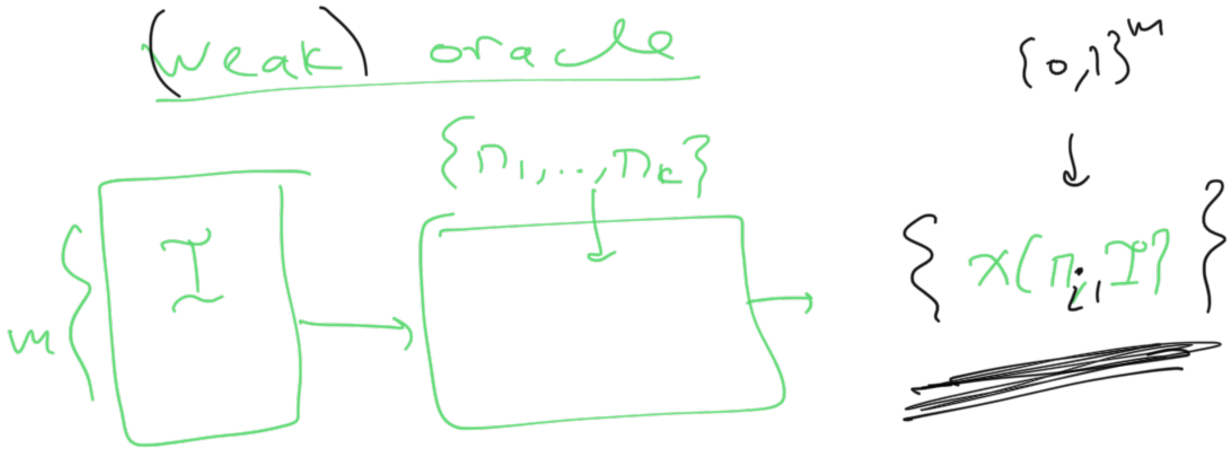
the subroutine (J) learns

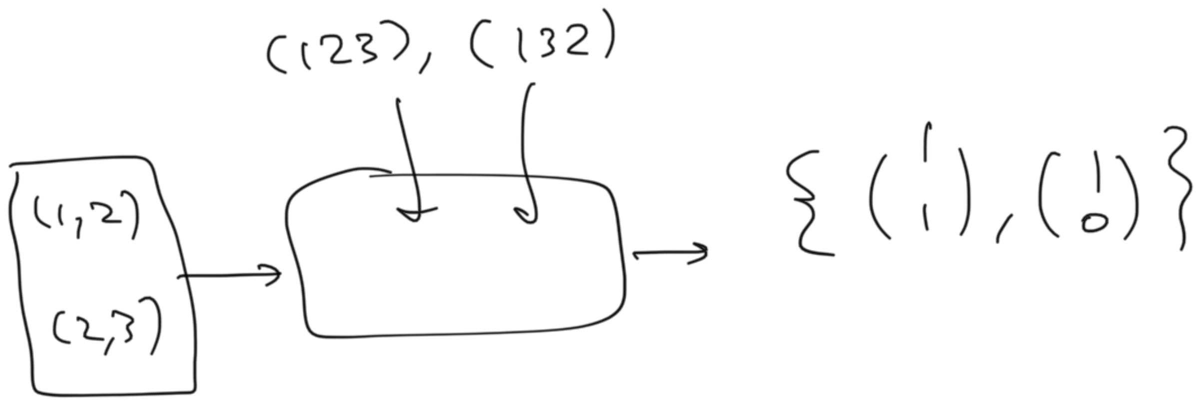
the set

$\{ \pi_1 \|_J , \dots , \pi_K \|_J \}$ → whp output of the weak oracle

Assume that we have this algo

weak oracle





$(i,j) \in J \times J$ Smart idea

$$\boxed{\begin{array}{c} \mathbb{1}\{ \pi\|_J(i) < \pi\|_J(j) \} \\ ||| \\ \mathbb{1}\{ \pi(i) < \pi(j) \} \end{array}}$$

(1) Subroutine

(2) Moments (High-noise regime ($\varphi \to 1$)) (←)

(1) $\{\sigma_i\}_{i=1}^{N} \sim M = \sum w_i M(\pi_i)$

goal : $\pi_i ||_J$ $\qquad J$

ρ injection $\sigma|_J$

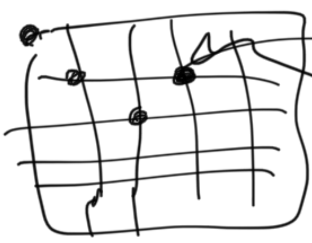
$$f(\rho) = \Pr_{\sigma\sim M}\left[\sigma|_J = \rho\right]$$

Design empirical

$$f_{emp}(\rho) = \frac{1}{N} \sum \mathbb{1}\{\sigma_m|_J = \rho\}$$

Idea: Collect poly # of Mallows mixtures & you do grid search

$$\mathcal{M} = \left\{ \sum_{i=1}^{k} \frac{r_i}{L} M(\pi_{\rho_i}) \right\}$$



N

emp.

Compare

$$Tv(\text{emp in the grid}, \text{emp of the true}) \le \eta$$

Return

$$\{\pi_{\rho_i} ||_J\}$$

feed this thing to the Thm for

noiseless.

---

(2) Moment

High-noise $\phi \to 1$ $\quad \varepsilon = 1 - \phi$

$$M = \sum_{i=1}^{k} \frac{1}{k} M(\pi_i)$$

$$M' = \sum^{k} \frac{1}{k} M(\pi_i')$$

$$TV(M, M') = ?$$

$$f(\sigma) = \frac{1}{Z} (1-\varepsilon)^{KT(\pi_i, \sigma)}$$

$$= \frac{1}{Z} \sum_{l=0}^{KT(\sigma,\pi_i)} \binom{KT(\sigma,\pi_i)}{l} (-\varepsilon)^{l}$$

Thm

$$TV(M, M') = O(\varepsilon^{\cancel{k}\, d+1})$$

$$\Updownarrow$$

$$\sum_{i=1}^{k} KT(\sigma, \pi_i)^{l} = \sum_{i=1}^{k} KT(\sigma, \pi_i')^{l}$$

$\forall\, l \in [n], \sigma \in S_n$

moment-like thing

informal

$$\boxed{\sum x^r}$$

$$KT(\pi, \sigma) = \sum_{\substack{(i,j) \\ \sigma(i) > \sigma(j)}} \mathbb{1}\{\pi(i) < \pi(j)\}$$

$\boxed{KT(\sigma,\pi)^r}$

$$KT(\pi,\sigma)^r = \sum_{\substack{(i_1,j_1)\\ \vdots \\ (i_\ell,j_\ell):\\ \sigma(i_1)>\sigma(j_1)\\ \vdots\\ \sigma(i_\ell)>\sigma(j_\ell)}} \underbrace{\prod_{k=1} X^{\pi}_{i_k j_k}}$$

$\pi \qquad \sigma$

$$\mathbb{1}\{\pi(i) < \pi(j)\}$$

$$\shortparallel$$

$$X^{\pi}_{i,j}$$

$$KT(\pi,\sigma) = \sum_{\substack{(i,j)\\ \sigma(i)>\sigma(j)}} X^{\pi}_{ij}$$

$$KT(\pi,\sigma)^2 = \left(\qquad \right)^2$$

$$= \Bigg(\underline{\underline{\sum_{\substack{i,j\\ \sigma(i)>\sigma(j)}}}} X^{\pi}_{ij}\Bigg)\Bigg(\underline{\underline{\sum_{\substack{i,j\\ \sigma(i)>\sigma(j)}}}} X^{\pi}_{ij}\Bigg)$$

$$= \sum_{\substack{(i,j)\ (a,\beta)\\ \sigma(i)>\sigma(j)\\ \sigma(a)>\sigma(\beta)}} \underbrace{X^{\pi}_{ij}\, X^{\pi}_{\alpha\beta}}$$

$\boxed{\sum x^r}$

$\pi \longrightarrow$ exp.
large
natural

$\boxed{(i,j)}$

Moment

Problems

⋮

Hamburger

⋮

Hausdorff

⋮

Stieltjes.

⋮

$\mu \longleftrightarrow$ moments

comp

powers of

π