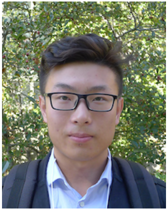
Optimal Ranking Recovery from Pairwise Comparisons

Anderson Ye Zhang

Department of Statistics
University of Pennsylvania



Pinhan Chen
UChicago Stat

Chao Gao
UChicago Stat

Ranking Examples

Sports and Gaming:

Image source: `www.psdcovers.com/wp-content/uploads/2012/07/NFL-vector-logos-1024x772.jpg`

Ranking Examples

Recommendation System and Web Search:

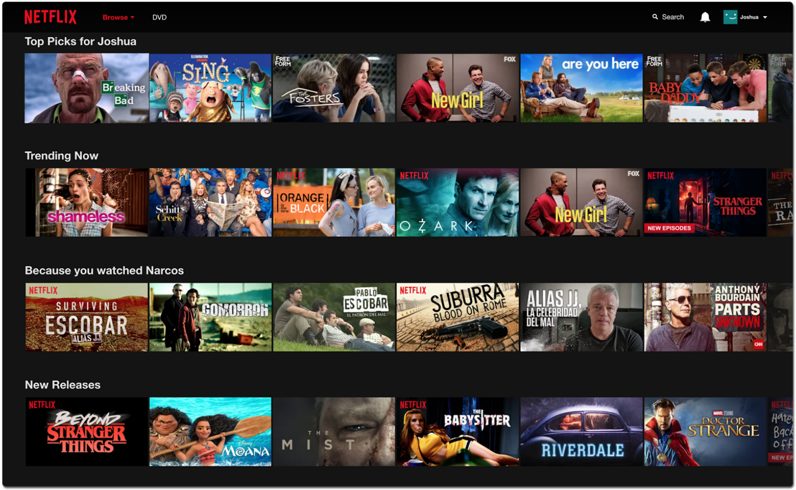
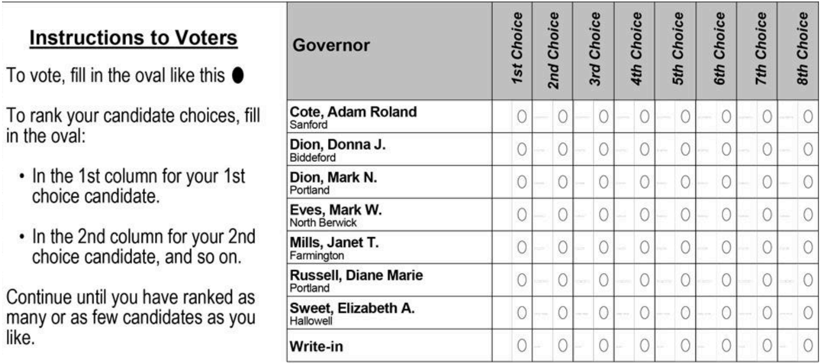


Image source: `https://miro.medium.com/max/2400/1*dMR3xmufnmKiw4crlisQUA.png`

Ranking Examples

Ranked Voting:



Instructions to Voters

To vote, fill in the oval like this ●

To rank your candidate choices, fill in the oval:

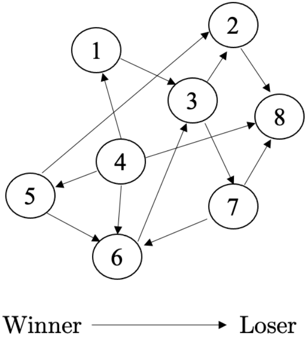
- In the 1st column for your 1st choice candidate.
- In the 2nd column for your 2nd choice candidate, and so on.

Continue until you have ranked as many or as few candidates as you like.

Governor	1st Choice	2nd Choice	3rd Choice	4th Choice	5th Choice	6th Choice	7th Choice	8th Choice
Cote, Adam Roland Sanford	O	O	O	O	O	O	O	O
Dion, Donna J. Biddeford	O	O	O	O	O	O	O	O
Dion, Mark N. Portland	O	O	O	O	O	O	O	O
Eves, Mark W. North Berwick	O	O	O	O	O	O	O	O
Mills, Janet T. Farmington	O	O	O	O	O	O	O	O
Russell, Diane Marie Portland	O	O	O	O	O	O	O	O
Sweet, Elizabeth A. Hallowell	O	O	O	O	O	O	O	O
Write-in	O	O	O	O	O	O	O	O

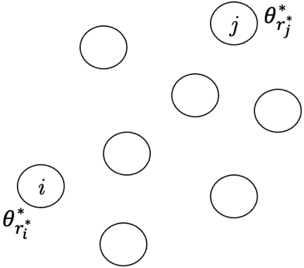
www.mainepublic.org/sites/mpbn/files/styles/x_large/public/201805/DEMRCV2Contests20180420.jpg

Ranking from Pairwise Comparisons



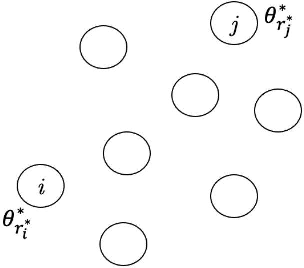
- ✓ Team 4 is the strongest
- ✓ Team 8 is the weakest
- ? Other Teams

Bradley-Terry-Luce (BTL) Model



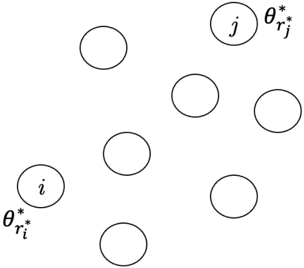
- n teams
- A sorted skill parameter θ^*: $\theta^*_1 \geq \ldots \geq \theta^*_n$

Bradley-Terry-Luce (BTL) Model



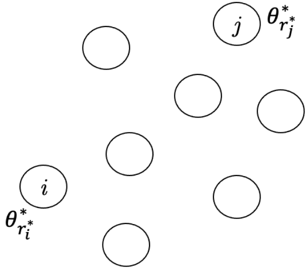
- n teams
- A sorted skill parameter θ^*: $\theta^*_1 \geq \ldots \geq \theta^*_n$
- Rank vector r^*: a permutation of $1, \ldots, n$

Bradley-Terry-Luce (BTL) Model

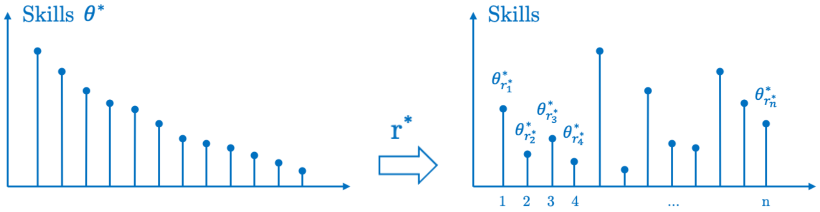


- n teams
- A sorted skill parameter θ^*: $\theta^*_1 \geq \ldots \geq \theta^*_n$
- Rank vector r^*: a permutation of $1, \ldots, n$
- For team i, its ranking among the n teams is r^*_i, and its skill parameter is $\theta^*_{r^*_i}$

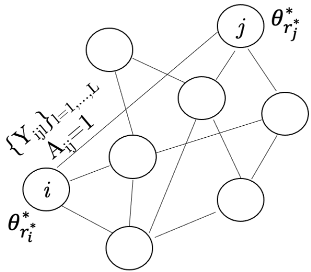
Bradley-Terry-Luce (BTL) Model



- n teams
- A sorted skill parameter θ^*: $\theta^*_1 \geq \ldots \geq \theta^*_n$
- Rank vector r^*: a permutation of $1, \ldots, n$
- For team i, its ranking among the n teams is r^*_i, and its skill parameter is $\theta^*_{r^*_i}$



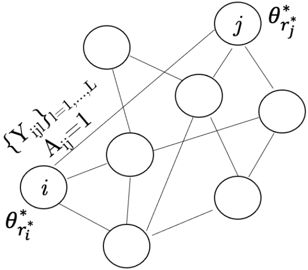
Bradley-Terry-Luce (BTL) Model



$$\mathbb{P}(i \text{ beats } j) \propto \exp\left(\theta^*_{r^*_i}\right)$$

$$\mathbb{P}(j \text{ beats } i) \propto \exp\left(\theta^*_{r^*_j}\right)$$

Bradley-Terry-Luce (BTL) Model



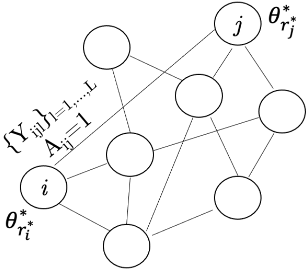
$$\mathbb{P}(i \text{ beats } j) \propto \exp\left(\theta^*_{r_i^*}\right)$$

$$\mathbb{P}(j \text{ beats } i) \propto \exp\left(\theta^*_{r_j^*}\right)$$

$$\mathbb{P}(i \text{ beats } j) = \frac{\exp\left(\theta^*_{r_i^*}\right)}{\exp\left(\theta^*_{r_i^*}\right) + \exp\left(\theta^*_{r_j^*}\right)} = \psi(\theta^*_{r_i^*} - \theta^*_{r_j^*})$$

where $\psi(x) = \frac{e^x}{e^x+1}$

Bradley-Terry-Luce (BTL) Model



$$\mathbb{P}(i \text{ beats } j) \propto \exp\left(\theta^*_{r^*_i}\right)$$

$$\mathbb{P}(j \text{ beats } i) \propto \exp\left(\theta^*_{r^*_j}\right)$$

$$\mathbb{P}(i \text{ beats } j) = \frac{\exp\left(\theta^*_{r^*_i}\right)}{\exp\left(\theta^*_{r^*_i}\right) + \exp\left(\theta^*_{r^*_j}\right)} = \psi(\theta^*_{r^*_i} - \theta^*_{r^*_j})$$

where $\psi(x) = \frac{e^x}{e^x+1}$

- Incomplete graph: $A_{ij} \overset{iid}{\sim} \text{Ber}(p)$
- L outcomes for each observed pair (i,j):

$$y_{ijl} \overset{ind}{\sim} \text{Ber}\left(\psi(\theta^*_{r^*_i} - \theta^*_{r^*_j})\right)$$

An Incomplete List of Prior Art

- Dwork, Kumar, Naor, and Sivakumar (2001)
- Hunter (2004)
- Jiang, Lim, Yao, and Ye (2011)
- Negahban and Wainwright (2012)
- Rajkumar and Agarwal (2014)
- Chen and Suh (2015)
- Jin, Zhang, Balakrishnan, Wainwright, and Jordan (2016)
- Shah, Balakrishnan, Guntuboyina, and Wainwright (2016)
- Shah and Wainwright (2017)
- Agarwal, Agarwal, Assadi, and Khanna (2017)
- Pananjady, Mao, Muthukumar, Wainwright, and Courtade (2017)
- Balakrishnan, Wainwright, and Yu (2017)
- Negahban, Oh, and Shah (2017)
- Chen, Fan, Ma, and Wang (2019)
- Shah, Balakrishnan, and Wainwright (2019)
- Heckel, Shah, Ramchandran, and Wainwright (2019)
- ...

An Incomplete List of Prior Art

- Dwork, Kumar, Naor, and Sivakumar (2001)
- Hunter (2004)
- Jiang, Lim, Yao, and Ye (2011)
- Negahban and Wainwright (2012)
- Rajkumar and Agarwal (2014)
- Chen and Suh (2015)
- Jin, Zhang, Balakrishnan, Wainwright, and Jordan (2016)
- Shah, Balakrishnan, Guntuboyina, and Wainwright (2016)
- Shah and Wainwright (2017)
- Agarwal, Agarwal, Assadi, and Khanna (2017)
- Pananjady, Mao, Muthukumar, Wainwright, and Courtade (2017)
- Balakrishnan, Wainwright, and Yu (2017)
- Negahban, Oh, and Shah (2017)
- Chen, Fan, Ma, and Wang (2019)
- Shah, Balakrishnan, and Wainwright (2019)
- Heckel, Shah, Ramchandran, and Wainwright (2019)
- ...

Most focus on θ^*

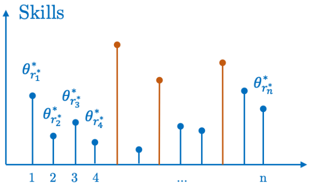
Recovery of r^*?

Two Tasks

- Top-k Ranking
- Full Ranking

Top-k Ranking

Problem Statement



- Top-k subset $S^* \subset \{1, 2, \dots, n\}$
  - $|S^*| = k$
  - For all $i \in S^*$, $\theta^*_{r_i^*} \geq \max_{j \notin S^*} \theta^*_{r_j^*}$

- How to estimate / recover S^*?

Image source: `https://en.wikipedia.org/wiki/Big_Three_(tennis)`

Problem Statement

- Top-k subset $S^* \subset \{1, 2, \ldots, n\}$
  - $|S^*| = k$
  - For all $i \in S^*$, $\theta^*_{r^*_i} \geq \max_{j \notin S^*} \theta^*_{r^*_j}$

- How to estimate / recover S^*?

- A natural idea:
  - Estimate $\{\theta^*_{r^*_1}, \ldots, \theta^*_{r^*_n}\}$ with $\{\hat{\theta}_1, \ldots, \hat{\theta}_n\}$
  - Find the top-k subset $\hat{S} \subset \{1, 2, \ldots, n\}$ such that
    - $|\hat{S}| = k$
    - For all $i \in \hat{S}$, $\hat{\theta}_i \geq \max_{j \notin \hat{S}} \hat{\theta}_j$

Algorithm 1: MLE

Step 1: Compute $\bar{y}_{ij} = \frac{1}{L}\sum_{l=1}^{L} y_{ijl}$

Step 2: Find the MLE $\hat{\theta}$ by minimizing

$$\ell_n(\theta) = \sum A_{ij}\left(\bar{y}_{ij}\,\text{log}\,\frac{1}{\psi(\theta_i - \theta_j)} + (1 - \bar{y}_{ij})\,\text{log}\,\frac{1}{1 - \psi(\theta_i - \theta_j)}\right)$$

Step 3: Find the top-k subset $\hat{S}$ from $\hat{\theta}$

Algorithm 2: Spectral Method

(Rank Centrality Algorithm)

Step 1: Construct the Markov transition matrix

$$P_{ij} = \begin{cases} \frac{1}{d} A_{ij} \bar{y}_{ji}, & i \neq j \\ 1 - \frac{1}{d} \sum_{l \neq i} A_{il} \bar{y}_{li}, & i = j \end{cases}$$

Step 2: Find the stationary distribution $\hat{\pi}$

Step 3: Find the top-k subset $\hat{S}$ from $\hat{\pi}$

Algorithm 2: Spectral Method

(Rank Centrality Algorithm)

Why spectral method works?

Population version:

$$M_{ij} = \mathbb{E}(P_{ij}|A) = \begin{cases} \frac{1}{d} A_{ij} \psi(\theta^*_{r^*_j} - \theta^*_{r^*_i}), & i \neq j \\ 1 - \frac{1}{d} \sum_{l \neq i} A_{il} \psi(\theta^*_{r^*_l} - \theta^*_{r^*_i}), & i = j \end{cases}$$

$$\pi^* = \left(\frac{\exp\left(\theta^*_{r^*_1}\right)}{\sum_l \exp\left(\theta^*_{r^*_l}\right)}, \ldots, \frac{\exp\left(\theta^*_{r^*_n}\right)}{\sum_l \exp\left(\theta^*_{r^*_l}\right)} \right)^T$$

Easy to check π^* is the stationary distribution of M

Our Result 1: Exact Recovery

Exact recovery: $\hat{S} = S^*$?

- ☺ MLE is optimal
- 😐 Spectral method is (in general) suboptimal, with a worse constant

Our Result 1: Exact Recovery

Exact recovery: $\hat{S} = S^*$?

- MLE is optimal
- Spectral method is (in general) suboptimal, with a worse constant

Our result complements Chen, Fan, Ma, Wang (2019):


The Annals of Statistics
2019, Vol. 47, No. 4, 2204–2235
https://doi.org/10.1214/18-AOS1745

SPECTRAL METHOD AND REGULARIZED MLE ARE BOTH OPTIMAL FOR TOP-K RANKING[1]

BY YUXIN CHEN[*,2], JIANQING FAN[†,*,3], CONG MA[*] AND KAIZHENG WANG[*]

Princeton University[*] *and Fudan University*[†]

Assumptions: $\theta_1^* \geq \theta_2^* \geq \ldots \geq \theta_n^*$

Separation: $\theta_k^* - \theta_{k+1}^* \geq \Delta$

Dynamic Range: $\theta_1^* - \theta_n^* \leq \kappa = O(1)$

Assumptions: $\theta_1^* \geq \theta_2^* \geq \ldots \geq \theta_n^*$

Separation: $\theta_k^* - \theta_{k+1}^* \geq \Delta$

Dynamic Range: $\theta_1^* - \theta_n^* \leq \kappa = O(1)$

Two variance functions:

MLE:

$$V(\kappa) = \max_{\substack{\kappa_1+\kappa_2 \leq \kappa \\ \kappa_1,\kappa_2 \geq 0}} \frac{n}{k\psi'(\kappa_1) + (n-k)\psi'(\kappa_2)}$$

Spectral method:

$$\overline{V}(\kappa) = \max_{\substack{\kappa_1+\kappa_2 \leq \kappa \\ \kappa_1,\kappa_2 \geq 0}} \frac{k\psi'(\kappa_1)(1+e^{\kappa_1})^2 + (n-k)\psi'(\kappa_2)(1+e^{-\kappa_2})^2}{(k\psi(\kappa_1) + (n-k)\psi(-\kappa_2))^2/n}$$

MLE

Theorem

Suppose

$$\Delta^2 > 2.001 V(\kappa) \frac{\left(\sqrt{\log k} + \sqrt{\log(n-k)}\right)^2}{npL}.$$

Then the MLE recovers the top-k subset S^* whp.
Suppose

$$\Delta^2 < 1.999 V(\kappa) \frac{\left(\sqrt{\log k} + \sqrt{\log(n-k)}\right)^2}{npL}.$$

Then no algorithm works.

☺ MLE is optimal

Spectral Method

Theorem

Suppose

$$\Delta^2 > 2.001\overline{V}(\kappa)\frac{\left(\sqrt{\log k} + \sqrt{\log(n-k)}\right)^2}{npL}.$$

Then the spectral method recovers the top-k subset S^* whp.
Suppose

$$\Delta^2 < 1.999\overline{V}(\kappa)\frac{\left(\sqrt{\log k} + \sqrt{\log(n-k)}\right)^2}{npL}.$$

Then the spectral method fails.

Spectral Method

Lemma

$\overline{V}(\kappa) \geq V(\kappa)$. *The equality holds if and only if* $\kappa = 0$.

Spectral Method

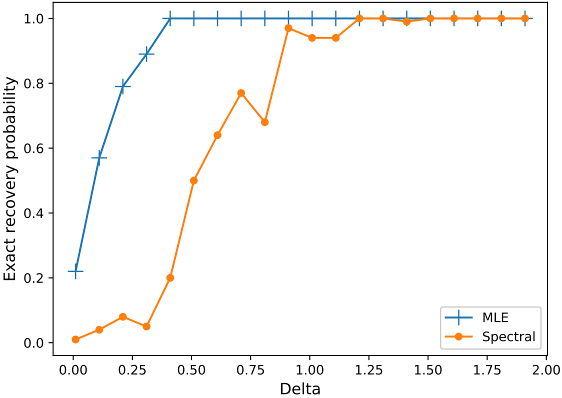
Lemma

$\overline{V}(\kappa) \geq V(\kappa)$. *The equality holds if and only if* $\kappa = 0$.

- ☺ When $\kappa = o(1)$ the spectral method is optimal.
- ⚠ Otherwise the spectral method is suboptimal with a worse constant.

Simulation

$n = 200$, $k = 50$, $p = 0.25$, $L = 20$
$\theta_1^*, \ldots, \theta_{50}^* \sim \text{Uniform}[6, 10], \quad \theta_{51}^*, \ldots, \theta_{200}^* \sim \text{Uniform}[0, 6 - \Delta]$
$\Rightarrow \kappa = 10$



Our Result 2: Partial Recovery

Partial recovery: Distance between $\hat{S}$ and S^*?

$$H(\hat{S}, S^*) = \frac{1}{2k} \left(|\hat{S} \cap S^{*C}| + |\hat{S}^C \cap S^*| \right)$$

Our Result 2: Partial Recovery

Partial recovery: Distance between $\hat{S}$ and S^*?

$$H(\hat{S}, S^*) = \frac{1}{2k}\left(|\hat{S} \cap S^{*C}| + |\hat{S}^C \cap S^*|\right)$$

- ☺ MLE is optimal
- 😐 Spectral method is (in general) rate-suboptimal

Minimax Rates

Theorem

The minimax rate of top-k ranking w.r.t. the loss $H(\hat{S}, S^*)$ is

$$\exp\left(-\frac{1}{2}\left(\frac{\sqrt{\mathsf{SNR}}}{2} - \frac{1}{\sqrt{\mathsf{SNR}}}\log\frac{n-k}{k}\right)_+^2\right)$$

where

$$\mathsf{SNR} = (1+o(1))\frac{npL\Delta^2}{V(\kappa)}.$$

Minimax Rates

Theorem

The minimax rate of top-k ranking w.r.t. the loss $H(\hat{S}, S^*)$ is

$$\exp\left(-\frac{1}{2}\left(\frac{\sqrt{\mathit{SNR}}}{2} - \frac{1}{\sqrt{\mathit{SNR}}}\log\frac{n-k}{k}\right)_+^2\right)$$

where

$$\mathit{SNR} = (1+o(1))\frac{npL\Delta^2}{V(\kappa)}.$$

Similar to support recovery problem for variable selection in the high-dimensional regression.

[Butucea, Ndaoud, Steppanova, and Tsybakov 2018]

[Ndaoud and Tsybakov 2020]

MLE

Theorem

The minimax rate of top-k ranking w.r.t. the loss $H(\hat{S}, S^*)$ is

$$\exp\left(-\frac{1}{2}\left(\frac{\sqrt{\mathit{SNR}}}{2} - \frac{1}{\sqrt{\mathit{SNR}}}\log\frac{n-k}{k}\right)_+^2\right)$$

where

$$\mathit{SNR} = (1+o(1))\frac{npL\Delta^2}{V(\kappa)}.$$

Moreover, the MLE achieves the above rate.

☺ MLE is optimal

Spectral Method

Theorem

The error rate of the spectral method w.r.t. the loss $H(\hat{S}, S^*)$ is

$$\exp\left(-\frac{1}{2}\left(\frac{\sqrt{\overline{\mathit{SNR}}}}{2} - \frac{1}{\sqrt{\overline{\mathit{SNR}}}}\log\frac{n-k}{k}\right)_+^2\right)$$

where

$$\overline{\mathit{SNR}} = (1+o(1))\frac{npL\Delta^2}{\overline{V}(\kappa)}.$$

Spectral Method

Theorem

The error rate of the spectral method w.r.t. the loss $H(\hat{S}, S^*)$ is

$$\exp\left(-\frac{1}{2}\left(\frac{\sqrt{\overline{\mathsf{SNR}}}}{2} - \frac{1}{\sqrt{\overline{\mathsf{SNR}}}}\log\frac{n-k}{k}\right)_+^2\right)$$

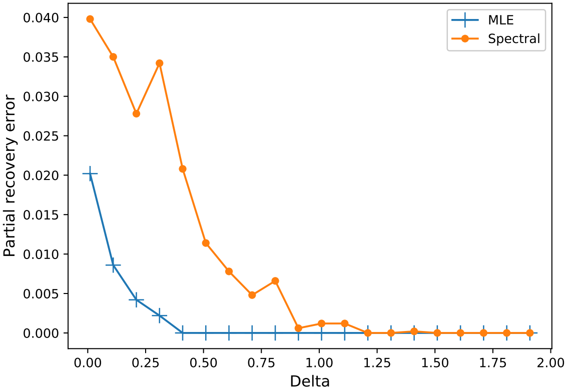
where

$$\overline{\mathsf{SNR}} = (1+o(1))\frac{npL\Delta^2}{\overline{V}(\kappa)}.$$

- ☺ When $\kappa = o(1)$ the spectral method is optimal.
- ⚠ Otherwise the spectral method is rate-suboptimal.

Simulation

$n = 200$, $k = 50$, $p = 0.25$, $L = 20$
$\theta_1^*, \ldots, \theta_{50}^* \sim \text{Uniform}[6, 10], \quad \theta_{51}^*, \ldots, \theta_{200}^* \sim \text{Uniform}[0, 6 - \Delta]$
$\Rightarrow \kappa = 10$



Summary for the Top-k Ranking Task

For both exact recovery and partial recovery:

- ☺ the MLE is optimal
- 😐 the spectral method is (in general) suboptimal

Full Ranking

Goal: to estimate / recover r^*

"Power Ranking" in sports: to rank all teams.

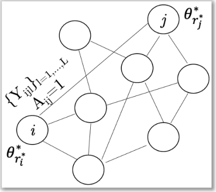
`sidelinespice.com/2018/11/29/week-12-2018-nfl-power-rankings/`

- Loss Function: Kendall's tau

$$\mathsf{K}(\hat{r}, r^*) = \frac{1}{n} \sum_{1 \le i < j \le n} \mathbb{I}\left\{\mathsf{sign}(\hat{r}_i - \hat{r}_j)\mathsf{sign}(r_i^* - r_j^*) < 0\right\}$$

- Loss Function: Kendall's tau

$$\mathsf{K}(\hat{r}, r^*) = \frac{1}{n} \sum_{1 \le i < j \le n} \mathbb{I}\left\{\text{sign}(\hat{r}_i - \hat{r}_j)\text{sign}(r_i^* - r_j^*) < 0\right\}$$



$$Y_{ijl} \overset{ind}{\sim} \text{Ber}\left(\psi(\theta^*_{r_i^*} - \theta^*_{r_j^*})\right)$$

- Regularity of Parameter:

$$\beta \le \theta_i^* - \theta_{i+1}^* \le C_0 \beta \text{ for all } i$$

Fundamental Limits

Theorem

Assume $p/\beta \gg \log n$. *Then*

$$\inf_{\hat{r}} \sup_{r^*} \mathbb{E} K(\hat{r}, r^*) \asymp \begin{cases} \frac{1}{n-1} \sum_{i=1}^{n-1} \exp\left(-\frac{(1+\delta)npL(\theta_i^* - \theta_{i+1}^*)^2}{4V_i(\theta^*)}\right), & \text{if } \frac{Lp\beta^2}{\beta \vee n^{-1}} > 1 \\ n \wedge \sqrt{\frac{\beta \vee n^{-1}}{Lp\beta^2}}, & \text{if } \frac{Lp\beta^2}{\beta \vee n^{-1}} \leq 1 \end{cases}$$

where

$$V_i(\theta^*) = \frac{n}{\sum_{j \in [n] \setminus \{i\}} \psi'(\theta_i^* - \theta_j^*)}.$$

Fundamental Limits

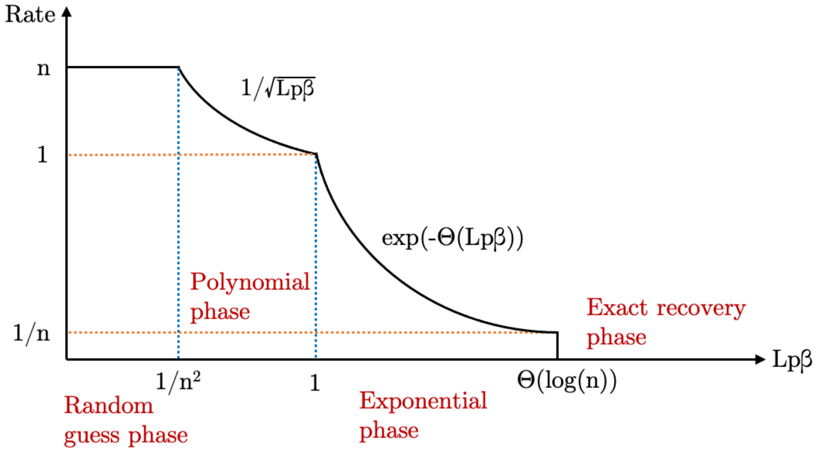
Special Case: $\beta \gtrsim n^{-1}$

The minimax rate becomes

$$\inf_{\hat{r}} \sup_{r^*} \mathbb{E}\mathsf{K}(\hat{r}, r^*) \asymp \begin{cases} \exp\left(-\Theta(Lp\beta)\right), & Lp\beta > 1, \\ n \wedge \sqrt{\frac{1}{Lp\beta}}, & Lp\beta \leq 1. \end{cases}$$

Phase Transition

$$\inf_{\hat{r}} \sup_{r^*} \mathbb{E}\mathsf{K}(\hat{r}, r^*) \asymp \begin{cases} \exp\left(-\Theta(Lp\beta)\right), & Lp\beta > 1, \\ n \wedge \sqrt{\frac{1}{Lp\beta}}, & Lp\beta \leq 1. \end{cases}$$

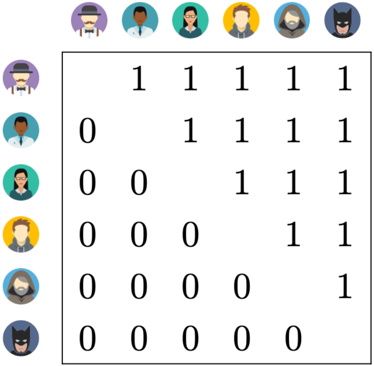


Pairwise Relation Matrix

Estimation of r^* $\Leftrightarrow$ Estimation of pairwise relation matrix R^*

$$R^*_{ij} = \mathbb{I}\left\{r^*_i < r^*_j\right\}$$

R^* with rows and columns properly rearranged:



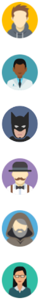
Strongest		1	1	1	1	1
	0		1	1	1	1
	0	0		1	1	1
	0	0	0		1	1
	0	0	0	0		1
Weakest	0	0	0	0	0	

Pairwise Relation Matrix

Estimation of r^* $\Leftrightarrow$ Estimation of pairwise relation matrix R^*

$$R^*_{ij} = \mathbb{I}\left\{r^*_i < r^*_j\right\}$$

R^*:



		0	1	0	1	0
	1		1	0	1	1
Weakest	0	0		0	0	0
Strongest	1	1	1		1	1
	0	0	1	0		0
	1	0	1	0	1	

Pairwise Relation Matrix

Estimation of r^* $\Leftrightarrow$ Estimation of pairwise relation matrix R^*

$$R^*_{ij} = \mathbb{I}\left\{r^*_i < r^*_j\right\}$$

Kendall's tau

$$\begin{aligned} \mathsf{K}(\hat{r}, r^*) &= \frac{1}{n} \sum_{1 \le i < j \le n} \mathbb{I}\left\{\text{sign}(\hat{r}_i - \hat{r}_j)\text{sign}(r^*_i - r^*_j) < 0\right\} \\ &= \frac{1}{n} \sum_{1 \le i < j \le n} \mathbb{I}\left\{\hat{R}_{ij} \neq R^*_{ij}\right\} \end{aligned}$$

$\hat{R}_{ij} = \mathbb{I}\left\{\hat{r}_i < \hat{r}_j\right\}$

Pairwise Relation Matrix

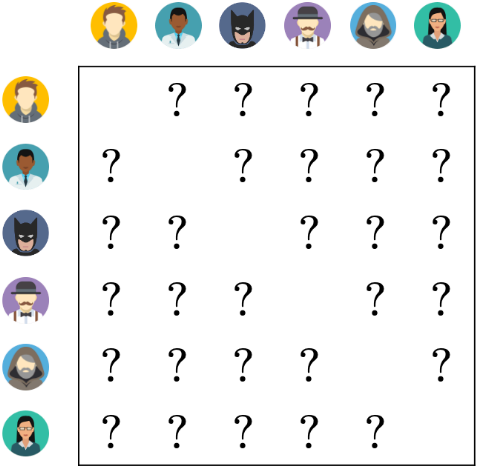
From $\hat{R}$ to $\hat{r}$:

Lemma

For any $\hat{R} \in \{0,1\}^{n \times n}$, let $\hat{r}$ be the rank obtained by sorting $\{\sum_{j \neq i} \hat{R}_{i,j}\}_{i=1,\dots,n}$. Then

$$K(\hat{r}, r^*) \leq \frac{4}{n} \sum_{1 \leq i \neq j \leq n} \mathbb{I}\{\hat{R}_{ij} \neq R^*_{ij}\}.$$

How to estimate R^*?



How to estimate $\mathbb{I}\{r_i^* < r_j^*\}$?

Algorithm: Divide-and-Conquer

Big Picture:

STEP 1 — League Partition: Partition the teams into several leagues. In each league, teams' skills are similar.

STEP 2 — Pairwise Relation Matrix Estimation: Estimate each $R^*_{ij} = \mathbb{I}\{r^*_i < r^*_j\}$ by local MLEs and other methods.

STEP 3 — Obtain $\hat{r}$ from $\hat{R}$

League Partition

...

National Football League (NFL)

Minor Football League

College Football

High School Football

League Partition

For each team i, count how many teams "dominate" it:

$$w_i = \sum_j A_{ij} \mathbb{I}\left\{j : \bar{y}_{ij} \leq t\right\}$$

League Partition

For each team i, count how many teams "dominate" it:

$$w_i = \sum_j A_{ij} \mathbb{I}\left\{j : \bar{y}_{ij} \leq t\right\}$$

Find the top league S_1 to include all teams that are dominated by at most h opponents:

$$S_1 = \{i : w_i \leq h\}$$

League Partition

For each team i, count how many teams "dominate" it:

$$w_i = \sum_j A_{ij} \mathbb{I}\{j : \bar{y}_{ij} \leq t\}$$

Find the top league S_1 to include all teams that are dominated by at most h opponents:

$$S_1 = \{i : w_i \leq h\}$$

League Partition

For each team i, count how many teams "dominate" it:

$$w_i = \sum_j A_{ij} \mathbb{I}\left\{ j : \bar{y}_{ij} \leq t \right\}$$

Find the top league S_1 to include all teams that are dominated by at most h opponents:

$$S_1 = \{ i : w_i \leq h \}$$

Remove all teams in S_1, and repeat the above procedure for the remaining teams.

League Partition

For each team i, count how many teams "dominate" it:

$$w_i = \sum_j A_{ij} \mathbb{I}\{j : \bar{y}_{ij} \leq t\}$$

Find the top league S_1 to include all teams that are dominated by at most h opponents:

$$S_1 = \{i : w_i \leq h\}$$

League Partition

For each team i, count how many teams "dominate" it:

$$w_i^{(2)} = \sum_{j \notin S_1} A_{ij} \mathbb{I} \{ j : \bar{y}_{ij} \leq t \}$$

Find the top league S_1 to include all teams that are dominated by at most h opponents:

$$S_1 = \{ i : w_i \leq h \}$$

League Partition

For each team i, count how many teams "dominate" it:

$$w_i^{(2)} = \sum_{j \notin S_1} A_{ij} \mathbb{I}\{j : \bar{y}_{ij} \leq t\}$$

Find the second top league S_2 to include all teams that are dominated by at most h opponents:

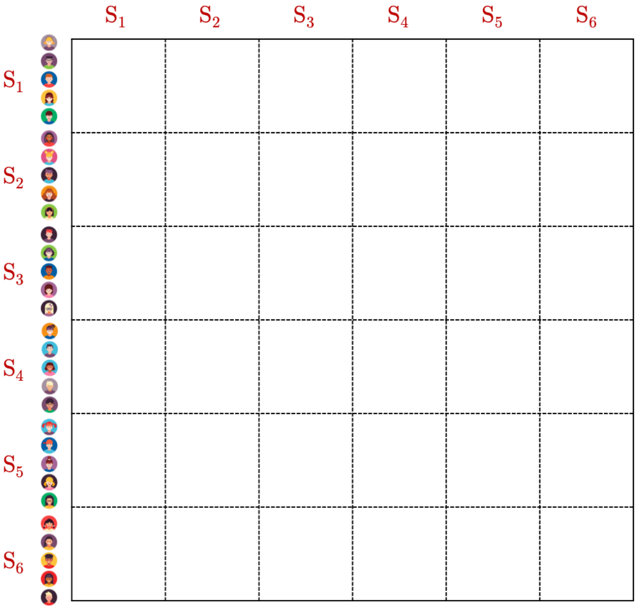
$$S_2 = \{i \notin S_1 : w_i^{(2)} \leq h\}$$

League Partition

We can show w.h.p.:

1. Teams have clear advantage against those who are at least two leagues below.
2. Teams in the same or neighboring leagues have close skills.
3. Teams having close skills are in the same or neighboring leagues.

Pairwise Relation Matrix Estimation



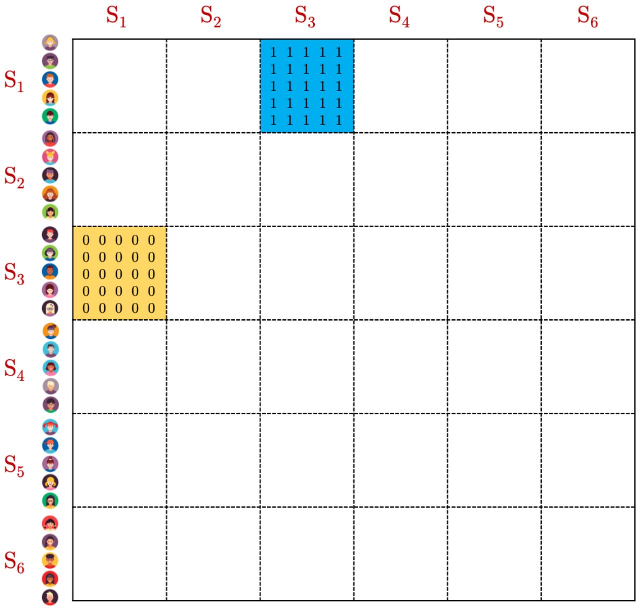
Pairwise Relation $R^*_{ij} = \mathbb{I}\{r^*_i < r^*_j\}$ Estimation

Scenario 1: j is at least two leagues below of i.

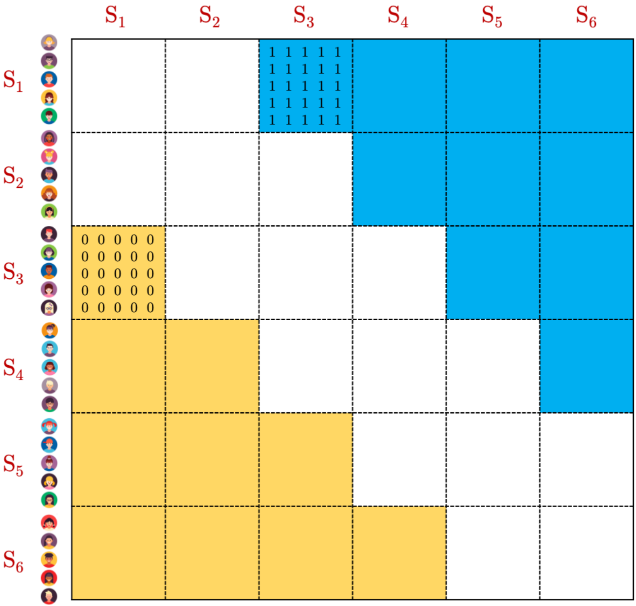
> 1 Teams have clear advantage against those who are at least two leagues below.

$\hat{R}_{ij} = 1$

Pairwise Relation Matrix Estimation



Pairwise Relation Matrix Estimation



Pairwise Relation $R^*_{ij} = \mathbb{I}\{r^*_i < r^*_j\}$ Estimation

Scenario 2: i, j are in the same or neighboring leagues.

Local MLE

- Find all teams with comparable skills to i or j

Pairwise Relation $R^*_{ij} = \mathbb{I}\{r^*_i < r^*_j\}$ Estimation

Scenario 2: i, j are in the same or neighboring leagues.

Local MLE

- Find all teams with comparable skills to i or j

Pairwise Relation $R^*_{ij} = \mathbb{I}\{r^*_i < r^*_j\}$ Estimation

Scenario 2: i, j are in the same or neighboring leagues.

Local MLE

- Find all teams with comparable skills to i or j

- Perform MLE on these teams only. Estimate their skills
- $\hat{R}_{ij} = \mathbb{I}\{\hat{\theta}_i > \hat{\theta}_j\}$

Pairwise Relation $R^*_{ij} = \mathbb{I}\{r^*_i < r^*_j\}$ Estimation

Scenario 2: i, j are in the same or neighboring leagues.

Local MLE

- Find all teams with comparable skills to i or j

> 2. Teams in the same or neighboring leagues have close skills.
> 3. Teams having close skills are in the same or neighboring leagues.

- Perform MLE on these teams only. Estimate their skills
- $\hat{R}_{ij} = \mathbb{I}\{\hat{\theta}_i > \hat{\theta}_j\}$

Pairwise Relation $R^*_{ij} = \mathbb{I}\{r^*_i < r^*_j\}$ Estimation

Scenario 2: i, j are in the same or neighboring leagues.

Local MLE

- Find all teams in the same / neighboring leagues of i or j

  2. Teams in the same or neighboring leagues have close skills.
  3. Teams having close skills are in the same or neighboring leagues.

- Perform MLE on these teams only. Estimate their skills
- $\hat{R}_{ij} = \mathbb{I}\{\hat{\theta}_i > \hat{\theta}_j\}$

Pairwise Relation $R^*_{ij} = \mathbb{I}\{r^*_i < r^*_j\}$ Estimation

Scenario 2: i, j are in the same or neighboring leagues.

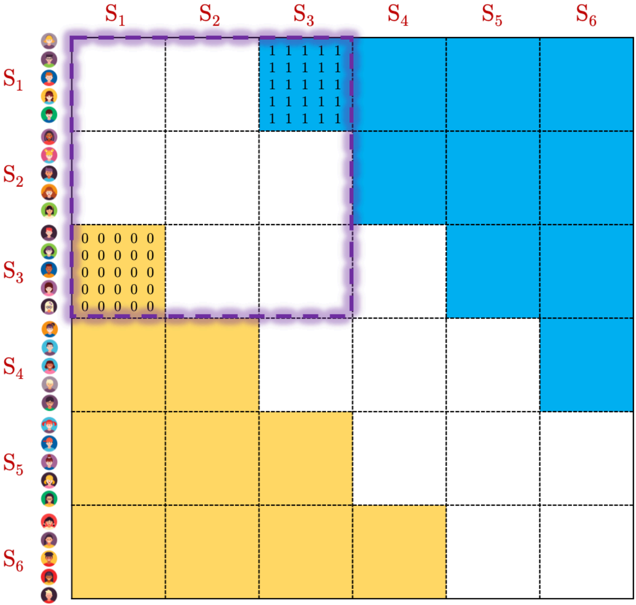
Local MLE

- Find all teams in the same / neighboring leagues of i or j
  - Ex. If $i, j \in S_1 \Rightarrow S_1 \cup S_2$
  - Ex. If $i \in S_2, j \in S_3 \Rightarrow S_1 \cup S_2 \cup S_3 \cup S_4$

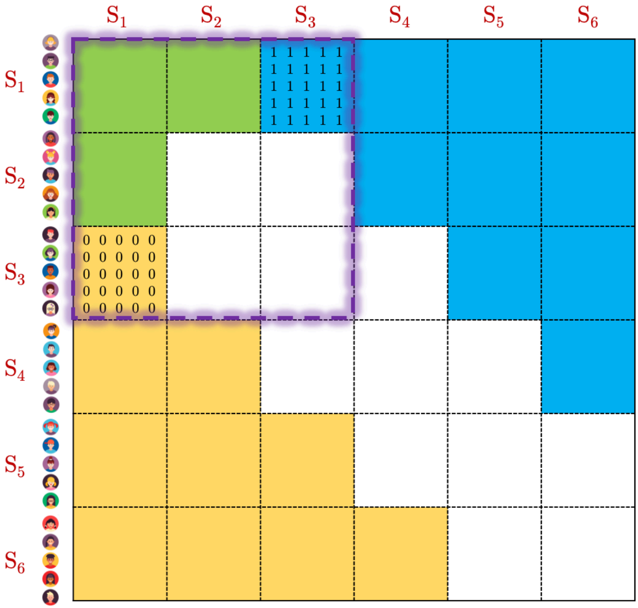
> 2 Teams in the same or neighboring leagues have close skills.
>
> 3 Teams having close skills are in the same or neighboring leagues.

- Perform MLE on these teams only. Estimate their skills
- $\hat{R}_{ij} = \mathbb{I}\{\hat{\theta}_i > \hat{\theta}_j\}$

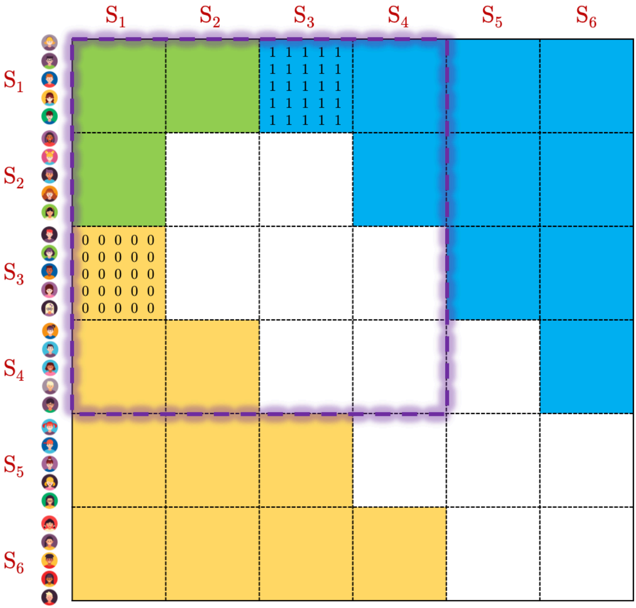
Pairwise Relation Matrix Estimation



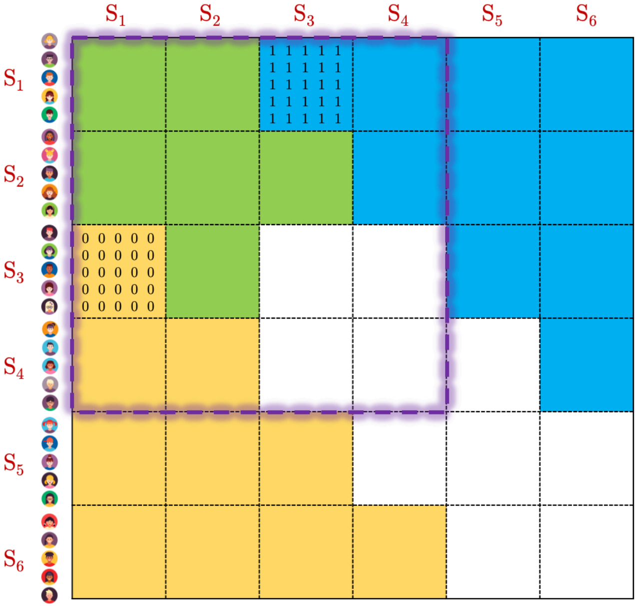
Pairwise Relation Matrix Estimation



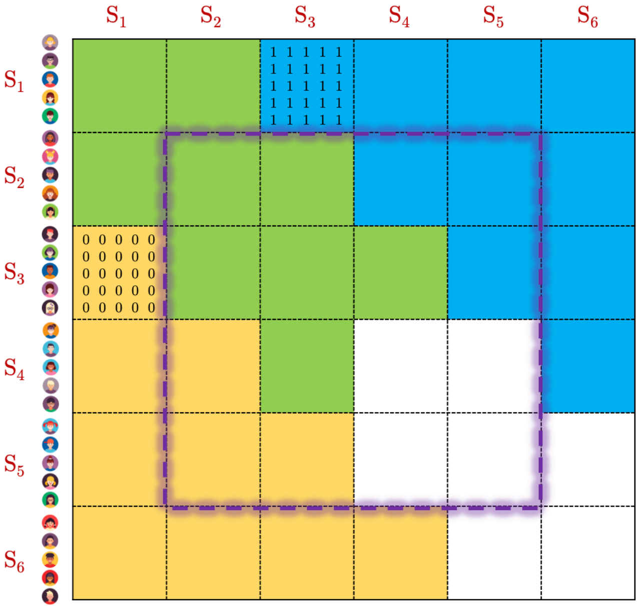
Pairwise Relation Matrix Estimation



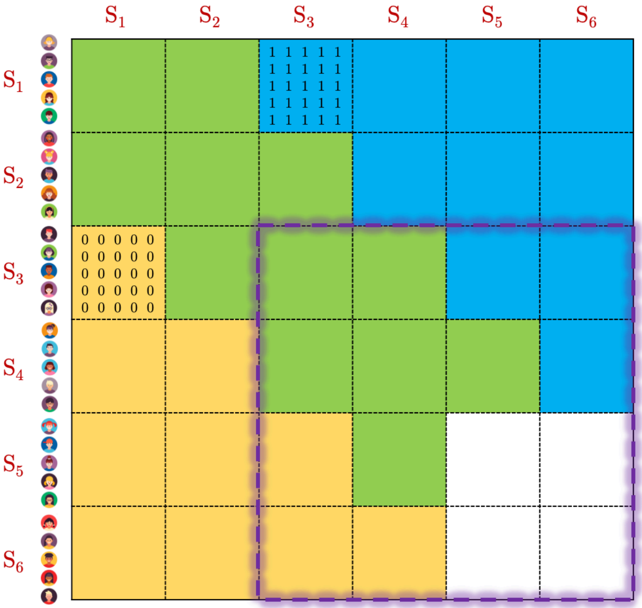
Pairwise Relation Matrix Estimation



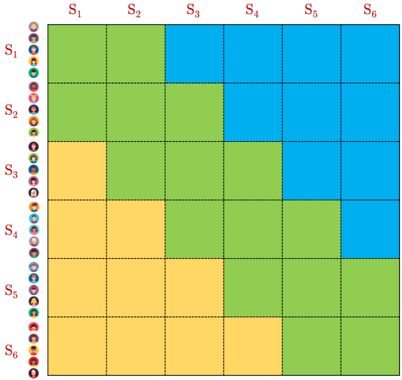
Pairwise Relation Matrix Estimation



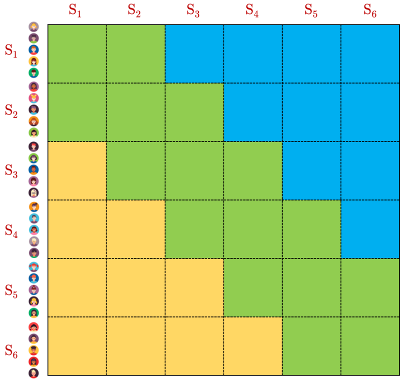
Pairwise Relation Matrix Estimation



Algorithm: Divide-and-Conquer



Algorithm: Divide-and-Conquer



Statistically Efficient

Computationally Efficient

Summary for the Full Ranking Task

- ☺ Minimax Rate: polynomial phase and exponential phase
- ☺ Divide-and-conquer Algorithm

Summary

[Top-k Ranking]

For both exact recovery and partial recovery:

- ☺ the MLE is optimal
- 😐 the spectral method is (in general) suboptimal

[Full Ranking]

- ☺ Minimax Rate: polynomial phase and exponential phaser
- ☺ Divide-and-conquer Algorithm

References

Pinhan Chen, Chao Gao, and Anderson Y Zhang. Partial recovery for top-k ranking: Optimality of mle and sub-optimality of spectral method. *arXiv preprint arXiv:2006.16485*, 2020

Pinhan Chen, Chao Gao, and Anderson Y Zhang. Optimal full ranking from pairwise comparisons. *arXiv preprint arXiv:2101.08421*, 2021