Rank Centrality: Ranking from Pair-wise Comparisons

Narutatsu (Edward) Ri

Overview

Problem Setup

Task: Rank Aggregation Problem

- n items: $1, \ldots, n$ with defined ordering $\sigma(n) = \{3, 10, 5, \ldots\}$
- Observe: Rankings $X_1, \ldots \overset{\text{i.i.d.}}{\sim} \text{Model}(\sigma(n))$
  - Rankings: $X_i = \sigma_i(n)$
  - Pairwise Comparisons: $X_i = \mathbb{1}[i \succ j]$

  $\ldots$
- Goal: Recover $\sigma(n)$ from observations $X_1, \ldots$

Problem Setup

Task: Rank Aggregation Problem

- n items: $1, \ldots, n$ with defined ordering $\sigma(n) = \{3, 10, 5, \ldots\}$
- Observe: Rankings $X_1, \ldots \overset{\text{i.i.d.}}{\sim} \text{Model}(\sigma(n))$
  - Rankings: $X_i = \sigma_i(n)$
  - Pairwise Comparisons: $X_i = \mathbb{1}[i \succ j]$

  $\ldots$
- Goal: Recover $\sigma(n)$ from observations $X_1, \ldots$

Why is this task interesting?

- Accurate estimate of rankings appears in **many** settings (Search Engines, Elections, Chess ELO, etc.)
- Also has great relevance in LLM scene (LLM Evaluation, RLHF, etc.)

Problem Setup

Model for Rankings

Natural setup: items $i \in [n]$ has a **score** $w_i \geq 0$ (Assume $\sum_{i=1}^{n} w_i = 1$)

Multinomial Logit (MNL)

Discrete choice model.

Choice probabilities proportional to systematic utility:

$$\Pr[i \text{ is chosen}] = \frac{\exp(\theta_i)}{\sum_{j=1}^{n} \exp(\theta_j)} = \frac{w_i}{\sum_{i=j}^{n} w_j}$$

where $\theta_i = \log\ w_i$.

How can sample rankings $\sigma_i(n)$?

→ Plackett–Luce Model

Problem Setup

Model for Rankings

Natural setup: items $i \in [n]$ has a **score** $w_i \geq 0$ (Assume $\sum_{i=1}^{n} w_i = 1$)

Plackett–Luce Model

Induces distribution over permutations of $[n]$

$$\Pr[\sigma] = \frac{w_{\sigma_1}}{\sum_{j=1}^{n} w_j} \cdot \frac{w_{\sigma_2}}{\sum_{j \neq \sigma_1} w_j} \cdot \dots \cdot \frac{w_{\sigma_n}}{w_{\sigma_n}}$$

$$= \prod_{j=1}^{n} \frac{w_{\sigma_j}}{\sum_{k=j}^{n} w_{\sigma_k}}$$

where σ_j is item ranked in j-th position.

When only comparing two items (pairwise marginal probabilities)
→ Bradley-Terry-Luce Model

Problem Setup

Model for Rankings

Natural setup: items $i \in [n]$ has a **score** $w_i \geq 0$ (Assume $\sum_{i=1}^{n} w_i = 1$)

Bradley-Terry-Luce (BTL) Model

$$\Pr[i \succ j] = \frac{w_i}{w_i + w_j}$$

Benefits

- Often easier to collect pairwise results
- Observing only pairwise outcomes is sufficient in recovering original parameters of MNL model ($\{w_i\}$)

Problem Setup

Addendum: Mallows Model

Assumes existence of correct ranking σ^*, induces distribution over rankings via Kendall-τ distance from σ^*:

$$\Pr[\sigma] \propto \exp\Big(-\theta \cdot D_\tau\big(\sigma, \sigma^*\big)\Big)$$

where $\theta > 0$ is a dispersion parameter (larger θ concentrates probability)

Problem Setup

Addendum: Mallows Model

Assumes existence of correct ranking σ^*, induces distribution over rankings via Kendall-τ distance from σ^*:

$$\Pr[\sigma] \propto \exp\left(-\theta \cdot D_\tau(\sigma, \sigma^*)\right)$$

where $\theta > 0$ is a dispersion parameter (larger θ concentrates probability)

Notes

- MLE is provably NP-hard
  (Intuition: Search space of all rankings, no easy way to narrow down)
- Kemeny Optimization: Finds ranking σ that minimizes

$$\sum_{\text{comparisons}} D_\tau(\sigma, \sigma(\text{comparisons}))$$

  …Also NP-Hard.
- Hence, not quite favored in practice

Problem Setup

Model for Rankings

Natural setup: items $i \in [n]$ has a **score** $w_i \geq 0$ (Assume $\sum_{i=1}^{n} w_i = 1$)

Bradley-Terry-Luce (BTL) Model

$$\Pr[i \succ j] = \frac{w_i}{w_i + w_j}$$

Setup in the Paper

- For each of E item pairs (i, j), observe k comparison outcomes:

$$X_1, \ldots, X_k \overset{\text{i.i.d.}}{\sim} \text{Bernoulli}\left(\frac{w_i}{w_i + w_j}\right)$$

- Want to estimate $w_1, \ldots, w_n$

Methods

Borda Count

Estimate w_i via average winrate of item i

$$\text{Borda Score}(i) = \frac{\sum_{j \neq i} a_{ij}}{\sum_{j \neq i} \mathbb{1}\{i \text{ and } j \text{ are compared}\}},$$

where

$$a_{ij} = \begin{cases} \text{fraction of times } i \succ j & \text{if } i \text{ and } j \text{ compared} \\ 0, & \text{otherwise} \end{cases}$$

Intuition: Compute weights via **local** pairwise comparisons

Methods

Borda Count

Estimate w_i via average winrate of item i

$$\text{Borda Score}(i) = \frac{\sum_{j \neq i} a_{ij}}{\sum_{j \neq i} \mathbb{1}\{i \text{ and } j \text{ are compared}\}},$$

where

$$a_{ij} = \begin{cases} \text{fraction of times } i > j & \text{if } i \text{ and } j \text{ compared} \\ 0, & \text{otherwise} \end{cases}$$

Intuition: Compute weights via **local** pairwise comparisons

Shortcomings

- Biased and inconsistent estimator of w_i
- Constant error in adversarial case
  ($w_i = 1$ for $i \in [n/2]$, $w_i = \epsilon$ for $i \in [n] \setminus [n/2]$)

Methods

Maximum Likelihood Estimation (MLE)

Choose w_i's that maximize likelihood given observed comparisons

$$\mathcal{L}(w_1, \ldots, w_n) = \prod_{(i,j) \in E} \left(\frac{w_i}{w_i + w_j} \right)^{Y_{ij}} \left(\frac{w_j}{w_i + w_j} \right)^{1 - Y_{ij}}$$

where $Y_{ij} = \mathbb{1}[j \succ i]$.

But, not possible to solve for $\nabla \mathcal{L}(w_1, \ldots, w_n) = 0$ closed-form,
must use gradient/second-order optimization ($\mathbf{w}^{(t+1)} = \mathbf{w}^{(t)} + \eta \nabla \mathcal{L} \left(\mathbf{w}^{(t)} \right)$)

When is MLE slower than Rank Centrality?

Spectral Methods

MC1~MC4 (Dwork et al., 2001)

Random walk on Markov Chain(s) with transition kernel(s) populated with win ratios

$$\mathbf{P}_{ij}^{(\mathrm{MC2})} = \frac{a_{ij}}{\sum_{\ell \neq i} a_{i\ell}},$$

$$\mathbf{P}_{ij}^{(\mathrm{MC3})} = \begin{cases} \frac{a_{ij}}{\deg(i)}, & \text{if } i \neq j, \\ 1 - \frac{1}{\deg(i)} \sum_{\ell \neq i} a_{i\ell}, & \text{otherwise}, \end{cases}$$

$$\mathbf{P}_{ij}^{(\mathrm{MC4})} = \begin{cases} \frac{1}{n}, & \text{if } a_{ij} \geq a_{ji}, \\ 0, & \text{if } a_{ij} < a_{ji}, \\ 1 - (\text{appropriate sum}), & \text{if } i = j. \end{cases}$$

Notes

- Stationary distributions are biased, $\hat{w}_i \propto w_i$ not guaranteed
- Unclear theoretical guarantees
  (transition kernel is not constructed to have "nice" properties)

This Paper: Rank Centrality

Random walk on Markov Chain(s) with transition kernel(s) populated with win ratios

$$\mathbf{P}_{ij}^{(\mathrm{RC})} = \begin{cases} \frac{1}{d_{\max}} a_{ij}, & \text{if } i \neq j, \\ 1 - \frac{1}{d_{\max}} \sum_{\ell \neq i} a_{i\ell}, & \text{otherwise } (i = j), \end{cases}$$

where $d_{\max} = \max_{i \in [n]} \deg(i)$.

This Paper: Rank Centrality

Random walk on Markov Chain(s) with transition kernel(s) populated with win ratios

$$\mathbf{P}_{ij}^{(\mathrm{RC})} = \begin{cases} \frac{1}{d_{\max}} a_{ij}, & \text{if } i \neq j, \\ 1 - \frac{1}{d_{\max}} \sum_{\ell \neq i} a_{i\ell}, & \text{otherwise } (i = j), \end{cases}$$

where $d_{\max} = \max_{i \in [n]} \deg(i)$. Denote $\widetilde{\mathbf{P}}^{(\mathrm{RC})}$ as when $k \to \infty$.

Rank Centrality

This Paper: Rank Centrality

Random walk on Markov Chain(s) with transition kernel(s) populated with win ratios

$$\mathbf{P}_{ij}^{(\mathrm{RC})} = \begin{cases} \frac{1}{d_{\max}} a_{ij}, & \text{if } i \neq j, \\ 1 - \frac{1}{d_{\max}} \sum_{\ell \neq i} a_{i\ell}, & \text{otherwise } (i = j), \end{cases}$$

where $d_{\max} = \max_{i \in [n]} \deg(i)$. Denote $\widetilde{\mathbf{P}}^{(\mathrm{RC})}$ as when $k \to \infty$.

Key point lies in the construction (we will see soon)

- $\widetilde{\mathbf{P}}^{(\mathrm{RC})}$ is reversible and satisfies **detailed balance condition**:

$$\widetilde{\pi}_i \widetilde{\mathbf{P}}_{ij}^{(\mathrm{RC})} = \widetilde{\pi}_j \widetilde{\mathbf{P}}_{ji}^{(\mathrm{RC})},$$

  where $\widetilde{\pi}_i = w_i / \sum_{j=1}^{n} w_j$ is a stationary distribution.
- $\mathbf{P}^{(\mathrm{RC})}$ (and hence $\widetilde{\mathbf{P}}^{(\mathrm{RC})}$) is ergodic (under easy assumptions), so $\widetilde{\pi}$ is unique

Rank Centrality

This Paper: Rank Centrality

Random walk on Markov Chain(s) with transition kernel(s) populated with win ratios

$$\mathbf{P}_{ij}^{(\mathrm{RC})} = \begin{cases} \frac{1}{d_{\max}} a_{ij}, & \text{if } i \neq j, \\ 1 - \frac{1}{d_{\max}} \sum_{\ell \neq i} a_{i\ell}, & \text{otherwise } (i = j), \end{cases}$$

where $d_{\max} = \max_{i \in [n]} \deg(i)$.

Motivating Questions

1. Conditions on k (num. comparisons per pair) to get low error with π?
   $\rightarrow$ Should depend on how "deviated" $\mathbf{P}^{(\mathrm{RC})}$ is from $\widetilde{\mathbf{P}}^{(\mathrm{RC})}$
2. How optimal is the sample efficiency of Rank Centrality?
   $\rightarrow$ Compare to best possible error (adversarial, non-adversarial)

Rank Centrality

This Paper: Rank Centrality

Random walk on Markov Chain(s) with transition kernel(s) populated with win ratios

$$\mathbf{P}^{(RC)}_{ij} = \begin{cases} \frac{1}{d_{\max}} a_{ij}, & \text{if } i \neq j, \\ 1 - \frac{1}{d_{\max}} \sum_{\ell \neq i} a_{i\ell}, & \text{otherwise } (i = j), \end{cases}$$

where $d_{\max} = \max_{i \in [n]} \deg(i)$.

Motivating Questions

1. Conditions on k (num. comparisons per pair) to get low error with π?

   → Should depend on how "deviated" $\mathbf{P}^{(RC)}$ is from $\widetilde{\mathbf{P}}^{(RC)}$
2. How optimal is the sample efficiency of Rank Centrality?

   → Compare to best possible error (adversarial, non-adversarial)

Good to Ask

- Does how E (observed pairwise comparisons) is sampled, or range of w_i's matter?

Rank Centrality

Sample Complexity Bounds

Conditions on k (num. comparisons per pair) to get low error with π?

$\rightarrow$ Should depend on how "deviated" $\mathbf{P}^{(RC)}$ is from $\widetilde{\mathbf{P}}^{(RC)}$

How can approach question?

Sample Complexity Bounds

Conditions on k (num. comparisons per pair) to get low error with π?

$\rightarrow$ Should depend on how "deviated" $\mathbf{P}^{(RC)}$ is from $\widetilde{\mathbf{P}}^{(RC)}$

How can approach question?

- Interested in $\|\pi - \widetilde{\pi}\|$ after sufficiently traversing with $\mathbf{P}^{(RC)}$
  Well... cannot directly compute closed-form π of $\mathbf{P}^{(RC)}$ 🙄

Sample Complexity Bounds

Conditions on k (num. comparisons per pair) to get low error with π?
$\rightarrow$ Should depend on how "deviated" $\mathbf{P}^{(RC)}$ is from $\widetilde{\mathbf{P}}^{(RC)}$

How can approach question?

- Interested in $\|\boldsymbol{\pi} - \widetilde{\boldsymbol{\pi}}\|$ after sufficiently traversing with $\mathbf{P}^{(RC)}$
  Well... cannot directly compute closed-form π of $\mathbf{P}^{(RC)}$ 🙄
- Intuition: $\|\boldsymbol{\pi} - \widetilde{\boldsymbol{\pi}}\|$ should depend on "deviation" $\Delta_k = \mathbf{P}^{(RC)} - \widetilde{\mathbf{P}}^{(RC)}$
  If Δ_k is "small enough," then perhaps could shave it off
  $\rightarrow$ Power iteration w/ $\mathbf{P}^{(RC)}$ and observe its deviation:

$$p_{t+1}^T = p_0^T \cdot \left(\mathbf{P}^{(RC)}\right)^t$$

  then take $t \rightarrow \infty$

Rank Centrality

Sample Complexity Bounds

Lemma (Lemma 2)

$\widetilde{\pi}_{\min} = \min_i \widetilde{\pi}(i)$, $\widetilde{\pi}_{\max} = \max_i \widetilde{\pi}(i)$, $\rho = \lambda_{\max}(\widetilde{\mathbf{P}}) + \|\Delta\|_2 \sqrt{\frac{\widetilde{\pi}_{\max}}{\widetilde{\pi}_{\min}}}$.

The following holds:

$$\frac{\|p_t - \widetilde{\pi}\|}{\|\widetilde{\pi}\|} \leq \underbrace{\rho^t \frac{\|p_0 - \widetilde{\pi}\|}{\|\widetilde{\pi}\|} \sqrt{\frac{\widetilde{\pi}_{\max}}{\widetilde{\pi}_{\min}}}}_{\text{mixing towards ideal chain}} + \underbrace{\frac{1}{1-\rho} \|\Delta\|_2 \sqrt{\frac{\widetilde{\pi}_{\max}}{\widetilde{\pi}_{\min}}}}_{\text{unavoidable fixed error floor}} .$$

Key Points

- Uses reversibility of $\widetilde{\mathbf{P}}$, define weighted space $L^2(\widetilde{\pi})$ where $\widetilde{\mathbf{P}}$ is self-adjoint, a bit of cleverness ($(p_\ell - \widetilde{\pi}) \cdot \mathbf{1} = 0$ for any ℓ)
- However, $\|\Delta\|_2$ and ρ still remain, so need to bound these $\rightarrow$ Lemma 3 and 4

Rank Centrality

Sample Complexity Bounds

Lemma (Lemma 3)

The error matrix Δ *satisfies*

$$\|\Delta\|_2 \leq C\sqrt{\frac{\log n}{k\, d_{\max}}}$$

with probability at least $1 - 4/n$.

Key Points

- Mostly grinding out math
- Disentangling $\Delta = \mathrm{diag}(\Delta) + \overline{\Delta}$
- Bounding $\|D\|$: Rewriting D_{ii} (which is Δ_{ii}) as martingale
- Bounding $\|\overline{\Delta}\|$: Rewriting $\overline{\Delta}$ into a form such that standard matrix martingale bounds can be applied

Sample Complexity Bounds

Lemma (Lemma 4)

$1 - \rho = 1 - \lambda_{\max}(\widetilde{\mathbf{P}}) - \|\Delta\|_2 \sqrt{\frac{\widetilde{\pi}_{\max}}{\widetilde{\pi}_{\min}}}$.

For $k \geq \left(\frac{256\, b^5 \kappa^2}{d_{\max}\xi^2}\right) \log n$, $\|\Delta\|_2 \leq C\sqrt{\frac{\log n}{k\, d_{\max}}}$,

$$1 - \rho \geq \frac{\xi d_{\min}}{b^2 d_{\max}}$$

where $\xi = 1 - \lambda_{\max}(\widetilde{\mathbf{P}}^{(RC)})$ *holds.*

Key Points

- Applying **Comparison lemma**: for reversible MCs $\mathbf{Q}$ and $\widetilde{\mathbf{P}}$,

$$\frac{1 - \lambda_{\max}(\widetilde{\mathbf{P}})}{1 - \lambda_{\max}(\mathbf{Q})} \geq \frac{\alpha}{\beta}$$

- Consider MC with transition kernel $\mathbf{Q}$ where weights are uniform ($1/\deg(i)$ for item i), for which $\lambda_{\max}$ is known

Rank Centrality

Sample Complexity Bounds

Theorem

Denote $b = \max_{i,j} \frac{w_i}{w_j}$, *and* $\kappa = \frac{d_{\max}}{d_{\min}}$.
With probability at least $1 - 4/n$, *the following holds:*

$$\frac{\|\pi - \tilde{\pi}\|}{\|\tilde{\pi}\|} \leq \frac{8\, b^{5/2} \kappa}{\xi} \cdot \sqrt{\frac{\log n}{k\, d_{\max}}}.$$

Key Points

- $\sqrt{1/k}$ dependency: as $k \to \infty$, relative error $\to 0$
- Influence of graph structure
  - ξ: connectedness of comparisons (spectral gap)
  - κ: comparison "inequality"
  - b: spread of weights

Sample Complexity Bounds

Theorem

Denote $b = \max_{i,j} \frac{w_i}{w_j}$, *and* $\kappa = \frac{d_{\max}}{d_{\min}}$.
With probability at least $1 - 4/n$, *the following holds:*

$$\frac{\|\pi - \tilde{\pi}\|}{\|\tilde{\pi}\|} \leq \frac{8\, b^{5/2} \kappa}{\xi} \cdot \sqrt{\frac{\log n}{k\, d_{\max}}}.$$

For Erdos-Renyi graphs, with probability at least $1 - 10/n$:

$$\frac{\|\pi - \tilde{\pi}\|}{\|\tilde{\pi}\|} \leq 64\, b^{5/2} \sqrt{\frac{\log n}{k\, d}}\ .$$

Optimality of Performance

Minimax framework: "Adversarial" setting; Best worst-case performance of any algorithm $\mathcal{A}$

Theorem

Let $\mathcal{S}_b$ denote the space of all BTL score vectors with dynamic range at most b. Then,

$$\inf_{\mathcal{A}} \sup_{\tilde{\pi} \in \mathcal{S}_b} \frac{\mathbb{E}\left[\|\pi^{\mathcal{A}} - \tilde{\pi}\| \right]}{\|\tilde{\pi}\|} \geq \frac{1}{240\sqrt{10}} \frac{b-1}{b+1} \frac{1}{\sqrt{kd}} ,$$

under the Erdos-Renyi model.

Rank Centrality:

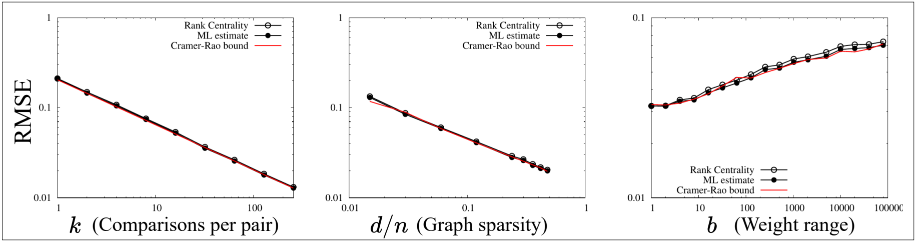
$$\frac{\|\pi - \tilde{\pi}\|}{\|\tilde{\pi}\|} \leq 64\, b^{5/2} \sqrt{\frac{\log n}{k\, d}} .$$

Optimal up to poly$(\log n)$ factor.

Optimality of Performance

"Non-adversarial" (structure-aware) bound on relative error via Cramer-Rao lower bound

Cannot compute closed-form; empirical validation:



- Both MLE and RC match Cramer-Rao bounds across different sweeps of k, d/n, and b

  (Default parameters: $k = 32$, $n = 400$, $d = 60$, $b = 10$)

Reinforcement Learning with Human Feedback

Preference-based Learning

- Comparing two options is easier than obtaining expert trajectory
- Turn qualitative pairwise preference data into quantitative reward signal $\rightarrow$ Can use BTL as underlying model

Reinforcement Learning with Human Feedback

Preference-based Learning

- Comparing two options is easier than obtaining expert trajectory
- Turn qualitative pairwise preference data into quantitative reward signal → Can use BTL as underlying model

Setup

Train reward model $r(x, y)$ with preferences $\mathcal{D} = \{(x^i, y_w^i, y_l^i)\}_{i \in [n]}$

Connect BTL to rewards:

$$\Pr[y_w \succ y_l \mid x] = \frac{\exp(r(x, y_w))}{\exp(r(x, y_w)) + \exp(r(x, y_l))} = \sigma(r(x, y_w) - r(x, y_l))$$

Train reward model by minimizing negative log likelihood:

$$\mathcal{L}(\mathcal{D}) = -\mathbb{E}_{(x, y_w, y_l) \sim \mathcal{D}} \left[\log \sigma(r(x, y_w) - r(x, y_l))\right]$$

Instead of estimating n weights $\{w_i\}_{i \in [n]}$, can think of estimating "infinite" number of weights via $r(x, y)$

Reinforcement Learning with Human Feedback

Preference-based Learning

Now, learn policy π_θ achieving high reward on $r(\cdot, \cdot)$

(typicall use LLM as backbone for π_θ)

$$\max_{\pi_\theta} \underbrace{\mathbb{E}_{x\sim\mathcal{D}, y\sim\pi_\theta(y|x)}\left[r(x, y)\right]}_{\text{maximize reward}} - \beta \underbrace{D_{\mathrm{KL}}\left(\pi_\theta(y \mid x) \parallel \pi_{\mathrm{ref}}(y \mid x)\right)}_{\text{without straying too far}}$$

Reinforcement learning in an online fashion: sample from π_{ref}, score with $r(\cdot, \cdot)$, train on result

Reinforcement Learning with Human Feedback

Preference-based Learning

Now, learn policy π_θ achieving high reward on $r(\cdot, \cdot)$

(typicall use LLM as backbone for π_θ)

$$\max_{\pi_\theta} \underbrace{\mathbb{E}_{x \sim \mathcal{D}, y \sim \pi_\theta(y|x)} \left[r(x, y) \right]}_{\text{maximize reward}} - \beta \underbrace{D_{\mathrm{KL}}(\pi_\theta(y \mid x) \,\|\, \pi_{\mathrm{ref}}(y \mid x))}_{\text{without straying too far}}$$

Reinforcement learning in an online fashion: sample from π_{ref}, score with $r(\cdot, \cdot)$, train on result

Two Mainstream Methods

- PPO (Proximal Policy Optimization):
  See Policy Gradient Theorem
- DPO (Direct Preference Optimization):
  Bypass learning $r(\cdot, \cdot)$ and directly trains π_θ with $\mathcal{D}$.

Reinforcement Learning with Human Feedback

Preference-based Learning

Now, learn policy π_θ achieving high reward on $r(\cdot,\cdot)$
(typicall use LLM as backbone for π_θ)

$$\max_{\pi_\theta} \underbrace{\mathbb{E}_{x\sim\mathcal{D}, y\sim\pi_\theta(y|x)}\left[r(x,y)\right]}_{\text{maximize reward}} - \beta \underbrace{D_{\mathrm{KL}}(\pi_\theta(y \mid x) \,||\, \pi_{\mathrm{ref}}(y \mid x))}_{\text{without straying too far}}$$

Reinforcement learning in an online fashion: sample from π_{ref}, score with $r(\cdot,\cdot)$, train on result

Two Mainstream Methods

- PPO (Proximal Policy Optimization):
  See Policy Gradient Theorem
- DPO (Direct Preference Optimization):
  Bypass learning $r(\cdot,\cdot)$ and directly trains π_θ with $\mathcal{D}$.

Can be thought of as doing a form of gradient "ascent." What are the exact effects on the backbone when doing this?

LLM Evaluation

Evaluation of LLMs is quite a hot topic

Trend: Steering away from static benchmarks (MMLU, HumanEval, etc.) and more open-ended evaluation by preference ranking (e.g., Chatbot Arena)

Questions

- How to collect annotations? (Dorner and Hardt, 2024)
- Automated evaluation (LLM-as-a-Judge (Dorner and Hardt, 2024; Guerdan et al., 2025; Gera et al., 2024))

  How can rank with weak evaluators?
  (Can think of as setting where observed preferences are ”noisy” not because of finite-samples but because of weak evaluator)