Learning Parametric Distributions from Samples and Preferences

Marc Jourdan, Gizem Yüce and Nicolas Flammarion
TML, School of Computer and Communication Sciences, EPFL, Switzerland

Motivation

Fact: Using preference data outperforms methods based on positive examples only, e.g., supervised fine-tuning vs. alignment phase.

How can preferences explain these empirical performance gains?

This paper: **Estimation for parametric distributions with preferences.**

Take-home message

- Preference-based M-estimators have **smaller asymptotic variance** than sample-only M-estimators.
- Deterministic preference-based MLE has an **accelerated estimation error rate of $\mathcal{O}(n^{-1})$**, improving upon the rate $\Theta(n^{-1/2})$ of M-estimators.
- **Matching minimax lower bound $\Omega(n^{-1})$**, up to constants.

Problem statement

Distribution $p_\theta \in \mathcal{P}(\mathcal{X})$ with $\theta \in \Theta \subseteq \mathbb{R}^k$ and $\mathcal{X} \subseteq \mathbb{R}^d$.

☞ Observe **samples** $(X_i, Y_i)_{i\in[n]} \sim p_{\theta^\star}^{\otimes[2n]}$ with $\theta^\star \in \Theta$.

Preference function $\ell_\theta : \mathcal{X}^2 \to \mathbb{R}$, e.g., $\ell_\theta(x,y) = r_\theta(x) - r_\theta(y)$.

- **Reward** function $r_\theta : \mathcal{X} \to \mathbb{R}$, e.g., $r_\theta = \log p_\theta$.

☞ Observe **preference** Z_i for each pair (X_i, Y_i)

- **Stochastic**: $Z_i = \begin{cases} 1 & \text{with prob. } \sigma(\ell_{\theta^\star}(X_i,Y_i)) \\ -1 & \text{otherwise} \end{cases}$ with $\sigma(x) \overset{\text{def}}{=} (1+e^{-x})^{-1}$.
- **Deterministic**: $Z_i = \operatorname{sign}(\ell_{\theta^\star}(X_i,Y_i))$.

Goal: **Estimate the unknown $\theta^\star$ based on n samples with preferences.**

Example: Gaussian natural parameter θ with known Σ and $r_\theta = \log p_\theta$, i.e., $\ell_\theta(X_i,Y_i) = \langle X_i - Y_i, \theta - \Sigma^{-1}(X_i+Y_i)/2\rangle$.

Sample-only M-estimator

The **sample-only maximum likelihood estimator** is

$$\widehat{\theta}_n^{\text{SO}} \in \arg\min_{\theta\in\Theta} L_n^{\text{SO}}(\theta) \quad \text{with} \quad L_n^{\text{SO}}(\theta) \overset{\text{def}}{=} -\sum_{i\in[n]} \log p_\theta^{\otimes 2}(X_i,Y_i) . \qquad \textbf{(SO MLE)}$$

Under enough regularity, **SO MLE** is asymptotically normal, i.e.,

$$\sqrt{n}(\widehat{\theta}_n^{\text{SO}} - \theta^\star) \rightsquigarrow_{n\to+\infty} \mathcal{N}(0_k, \mathcal{I}(p_{\theta^\star}^{\otimes 2})^{-1}) ,$$

where $\mathcal{I}(p_\theta) \overset{\text{def}}{=} \mathbb{E}_{p_\theta}[-\nabla_\theta^2 \log p_\theta]$ is the Fisher information matrix of p_θ.

Example: $L_n^{\text{SO}}(\theta) = \|\theta - \widehat{\theta}_n^{\text{SO}}\|_\Sigma^2$ with $\widehat{\theta}_n^{\text{SO}} = \frac{1}{2n}\sum_{i\in[n]} \Sigma^{-1}(X_i+Y_i)$.

Preference-based M-estimator

The **stochastic preferences MLE** is

$$\widehat{\theta}_n^{\text{SP}} \in \arg\min_{\theta\in\Theta}\{L_n^{\text{SO}}(\theta) - \frac{1}{n}\sum_{i\in[n]} \log \sigma(Z_i \ell_\theta(X_i,Y_i))\} . \qquad \textbf{(SP MLE)}$$

$\widehat{\theta}_n^{\text{SP}_{\text{det}}}$ defined similarly when preferences are deterministic.

Theorem 1 (Smaller asymptotic variance). *Under regularity and **geometric assumptions** on p_θ and ℓ_θ:*

- *$\widehat{\theta}_n^{SP}$ and $\widehat{\theta}_n^{SP_{\det}}$ are asymptotically normal estimators,*
- *with asymptotic variance $V_{\theta^\star}^{SP_{\det}} \preceq V_{\theta^\star}^{SP} \preceq \mathcal{I}(p_{\theta^\star}^{\otimes 2})^{-1}$.*

Beyond M-estimators for deterministic preferences

Minimizers of the **empirical** 0-1 **loss** are

$$\mathcal{C}_n \overset{\text{def}}{=} \arg\min_{\theta\in\Theta} \sum_{i\in[n]} \mathbb{1}\left(Z_i\ell_\theta(X_i,Y_i) < 0\right) = \{\theta \mid \forall i \in [n],\ Z_i\ell_\theta(X_i,Y_i) \geq 0\} .$$

Any estimator $\widehat{\theta}_n^{\text{AE}} \in \mathcal{C}_n$. The **deterministic preferences MLE** is

$$\widehat{\theta}_n^{\text{DP}} \in \arg\min\left\{L_n^{\text{SO}}(\theta) \mid \theta \in \mathcal{C}_n\right\} . \qquad \textbf{(DP MLE)}$$

Theorem 2 (Fast estimation rate within $\mathcal{C}_n$). *For Gaussian distributions with known Σ and $r_\theta = \log p_\theta$, for all $n \geq \widetilde{\mathcal{O}}(\log(1/\delta))$, with probability $1-\delta$,*

$$\forall \widehat{\theta}_n \in \mathcal{C}_n, \quad \left\|\widehat{\theta}_n - \theta^\star\right\|_\Sigma \leq \mathcal{O}\left(\frac{A_d}{n}\log(1/\delta)\log n\right) \quad \textit{with} \quad A_d =_{+\infty} \mathcal{O}(\sqrt{d}) .$$

Theorem also holds under **geometric assumptions** on p_θ and ℓ_θ:

- **Identifiability** under preferences feedback,
- **Linearization validity** of the preferences constraints, i.e., $\mathcal{C}_n \subseteq \widetilde{\mathcal{C}}_n$,
- **Positive regular** p.d.f. of $V_{\theta^\star,u}(X_i,Y_i) \overset{\text{def}}{=} \frac{\ell_{\theta^\star}(X_i,Y_i)}{-\langle u, \nabla_{\theta^\star}\ell_{\theta^\star}(X_i,Y_i)\rangle}$ for all u near 0.

Proof sketch: let $R_{n,u} \overset{\text{def}}{=} \max\{\varepsilon \geq 0 \mid \theta^\star + \varepsilon u \in \widetilde{\mathcal{C}}_n\}$ for all $u \in \mathcal{S}_{k-1}$. Then,

$$R_{n,u} \leq \min_{i\in[n]}\{V_{\theta^\star,u}(X_i,Y_i) \mid \ell_{\theta^\star}(X_i,Y_i)\langle u, \nabla_{\theta^\star}\ell_{\theta^\star}(X_i,Y_i)\rangle < 0\} .$$

Restrictive geometric assumptions

✗ Monotonic p_θ, e.g., Exponential, Laplace and Pareto with known location.
✓ Other distributions, e.g., Laplace with known scale, Rayleigh.
✓ Monotonic rewards $r_\theta = g \circ p_\theta$, e.g., odds-ratio $g(x) = -\log(x^{-1}-1)$.

Variants: Gaussian for $r_\theta = \log p_\theta$ with
✗ Reference model $\ell_\theta - \ell_{\theta_0}$ with known $\theta_0 \in \Theta$.
✓ Margins $\ell_\theta + c$ with known $c \in \mathbb{R}$.

Lower bound on the estimation error

Theorem 3 (Fast estimation rate is minimax optimal). *For Gaussian distributions with known Σ and $r_\theta = \log p_\theta$, for all n,*

$$\inf_{\widehat{\theta}_n}\sup_{\theta^\star\in\Theta} \mathbb{E}_{q_{\theta^\star,h_{det}}^{\otimes[n]}}\left[\left\|\widehat{\theta}_n - \theta^\star\right\|_\Sigma\right] \geq \Omega\left(\min\left\{\frac{A_d\sqrt{d}}{n}, \sqrt{\frac{d}{n}}\right\}\right) .$$

Theorem also holds under **geometric assumptions** on p_θ and ℓ_θ:

- **Squared Hellinger distance** $\mathrm{H}^2(\mathbb{P},\mathbb{Q})$ is bounded by a **quadratic**,
- The **Bhattacharyya coefficient** $\mathrm{BC}(\mathbb{P},\mathbb{Q}) \overset{\text{def}}{=} \left\|\sqrt{\mathbb{PQ}}\right\|_1$ restricted to the set of paired observations with disagreeing preference $\widetilde{\mathrm{BC}}$ is **Lipschitz**.

Open problem: Close the $\sqrt{d}$ gap between the upper and lower bounds.

Proof sketch: Assouad's lemma, $\mathrm{TV}(\mathbb{P}^{\otimes n},\mathbb{Q}^{\otimes n}) \leq \sqrt{2n\mathrm{H}^2(\mathbb{P},\mathbb{Q})}$ and $\mathrm{H}^2(q_{\tilde{\theta}}, q_\theta) = \widetilde{\mathrm{BC}}(\tilde{\theta},\theta) + \mathrm{H}^2(p_{\tilde{\theta}}^{\otimes 2}, p_\theta^{\otimes 2})$ for all $\theta, \tilde{\theta} \in \Theta$

Experiments on Gaussian with log-likelihood reward

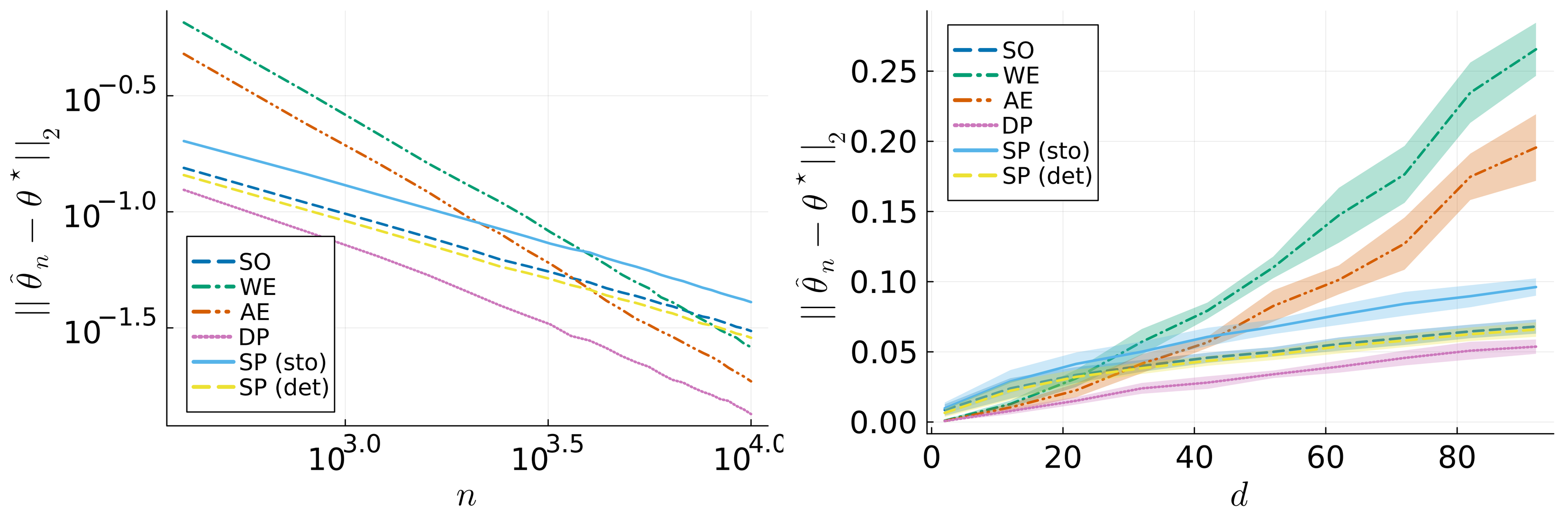

Figure 1: Estimation error with $\mathcal{N}(\theta^\star, I_d)$ where $\theta^\star \sim \mathcal{U}([1,2]^d)$, as a function of (a) the **sample size** n for $d = 20$ and (b) the **dimension** d for $n = 10^4$.

Given loss f, normalization $\beta > 0$ and regularization $\lambda \geq 0$,

$$\widehat{\theta}_n^{f,\lambda,\beta} \in \arg\min_{\theta\in\Theta}\{L_n^{\text{SO}}(\theta) + \lambda\sum_{i\in[n]} f(\beta Z_i\ell_\theta(X_i,Y_i))\} .$$

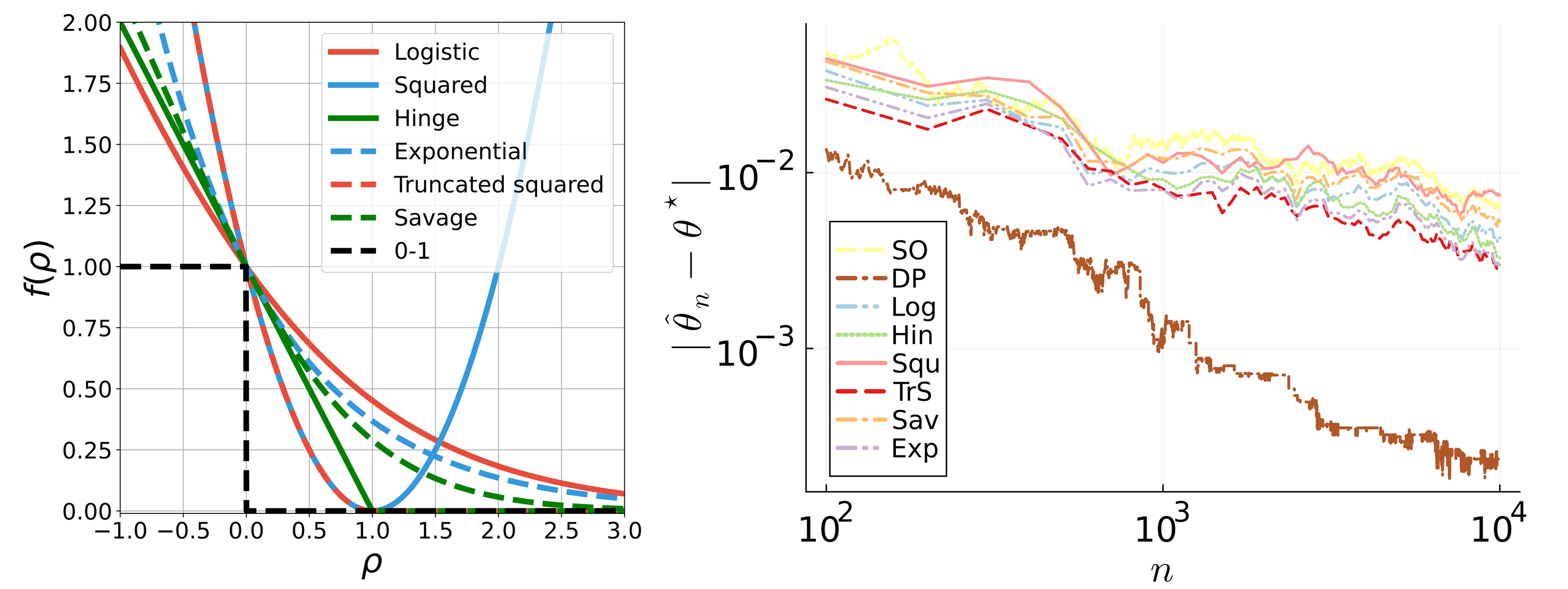

Figure 2: (a) Notable examples of binary classification loss functions. (b) **Estimation error** when minimizing the empirical losses for $\mathcal{N}(\theta^\star, 1)$ where $\theta^\star \sim \mathcal{U}([1,2])$ and $\beta = \lambda = 1$.