NEURAL INFORMATION PROCESSING SYSTEMS

Preferential Normalizing Flows

NeurIPS 2024

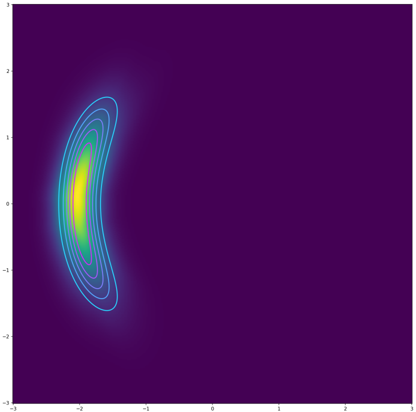
Petrus Mikkola Luigi Acerbi Arto Klami

University of Helsinki

UNIVERSITY OF HELSINKI

FCAI

Goal: Learn a probability density from preference data (comparisons or rankings).



Examples:

- Prior elicitation: **joint** prior distribution from expert comparisons
- AI alignment: probabilistic reward model from human preferences
- Density estimation: when the target density cannot be sampled or evaluated
  - E.g. LLM's belief density from rankings

Learn the probability density $p(\mathbf{x})$ from...

- Density estimation: Samples from density, $\mathbf{x} \sim p(\mathbf{x})$
- VI, MCMC: (unnormalized) Density evaluations, $p(\mathbf{x})$

Our setting:

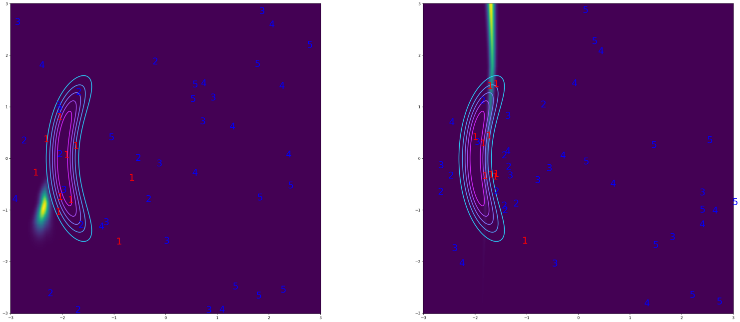
- Preferences (comparisons or rankings), "$\mathbf{x}$ *is preferred to* $\mathbf{x}'$" or "$\mathbf{x}$ *is more probable than* $\mathbf{x}'$", which means in a noise-free setting $p(\mathbf{x}) > p(\mathbf{x}')$

- We model the log target density $\log p(\mathbf{x})$ as the log density of a normalizing flow
  - Works with any flow as longs as the density is fast to compute; we use RealNVP and Neural Spline Flow
- Choices are generated by a random utility model (RUM) with exponentially distributed noise and utility function $\log p(\mathbf{x})$

Density estimation using Real NVP [Dinh *et al.*, ICLR17]

Neural Spline Flows [Durkan *et al.*, NeurIPS19]

What is the training objective? How about maximizing the likelihood of the preference data? Failure modes:



(a) Collapsing probability (b) Diverging probability

Our objective is the function-space maximum a posteriori estimate,

$$p(f \mid \mathcal{D}) \propto p(\mathcal{D} \mid f)p(f),$$

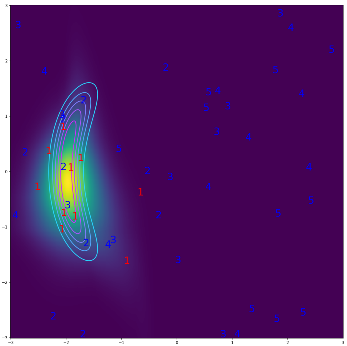
where f is the flow log-density.

- We introduce the empirical functional prior $p(f)$ that allocates probability mass to high-density points
  - Building on a limit theorem, the prior can be made consistent with the underlying decision-model (RUM)

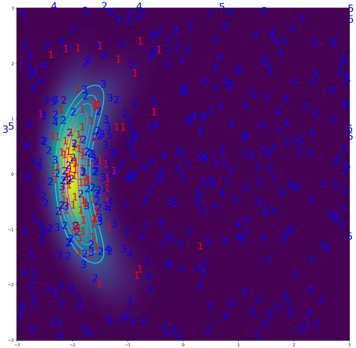
Should we learn most likely functions or parameters? [Qiu *et al.*, NeurIPS23]

Probabilistic choice with an infinite set of options [Malmberg&Hössjer, 2014]

With the prior:



(c) 10 rankings, $k = 5$



(d) 100 rankings, $k = 5$