Linear optimal partial transport embedding

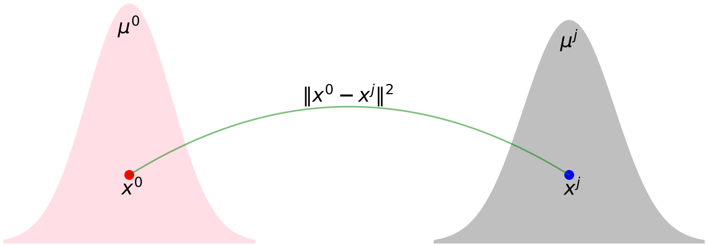
Yikun Bai[a] Ivan Medri[a] Rocio Diaz Martin[b]
Rana Muhammad Shahroz Khan[a] Soheil Kolouri[a]

[a]Department of Computer Science, Vanderbilt University

[b]Mathematics Department, Vanderbilt University

July 6, 2023

Introduction: optimal transport problem (OT)

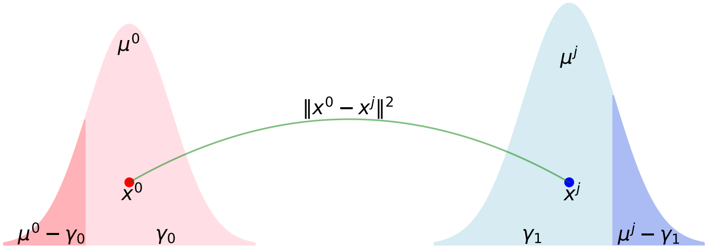


$$\text{OT}(\mu^0, \mu^j) := \inf_{\gamma \in \Gamma(\mu^0, \mu^j)} \int_{\Omega^2} \|x^0 - x^j\|^2 d\gamma(x, y)$$

where $\Omega \subset \mathbb{R}^d$, $\mu^0, \mu^j \in \mathcal{P}(\Omega), \Gamma(\mu^0, \mu^j) := \{\gamma \in \mathcal{P}(\Omega^2) : \gamma_0 = \mu^0, \gamma_1 = \mu^j\}$.

- Statistics: hypothesis test, statistical inference
- Machine learning: GAN, VAE, transfer learning
- Limitation: Requires equal total amount of mass between the two measures.

Introduction: optimal partial transport (OPT)



$$\mathrm{OPT}_\lambda(\mu^0, \mu^j) := \inf_{\gamma \in \Gamma_\leq(\mu^0, \mu^j)} \int_{\Omega^2} \|x^0 - x^j\| d\gamma(x^0, x^j) + \lambda(|\mu^0 - \gamma_0|_{TV} + |\mu^j - \gamma_1|_{TV}),$$

where $\mu^0, \mu^j \in \mathcal{M}_+(\Omega)$, $\Gamma_\leq(\mu^0, \mu^j) := \{\gamma \in \mathcal{M}_+(\Omega^2) : \gamma_0 \leq \mu^0, \gamma_1 \leq \mu^j\}$, and $\lambda \geq 0$.

- **Benefits**: Partial matching and comparison of measures with unequal mass.

Our work: LOPT embedding

Based on the above formulation, we define

$$\mu^j \mapsto (u^j, \bar{\mu}^j, \nu^j) := (T - \mathrm{id}, \gamma_0, \mu^j - \gamma_1).$$

where v_0, γ_0, ν^j are defined as in the last proposition.
Linear OPT discrepancy:

$$\begin{aligned} \mathrm{LOPT}_{\mu_0,\lambda} &= \|u^i - u^j\|^2_{\bar{\mu}^i \wedge \bar{\mu}^j, 2\lambda} + \lambda(|\bar{\mu}^i - \bar{\mu}^j| + |\nu^i - \nu^j|) \\ &\approx OPT_\lambda(\mu^i, \mu^j) \end{aligned}$$

where

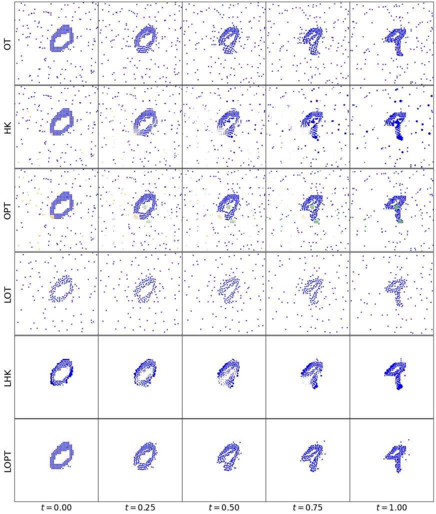
$$\|u^i - u^j\|^2_{\bar{\mu} \wedge \bar{\nu}, 2\lambda} := \int_{\Omega^2} \|u^i - u^j\|^2 \wedge 2\lambda d(\bar{\mu}^i \wedge \bar{\nu}^j)$$

When $\mu^i = \mu^0$, we proved that $LOPT$ can recover OPT distance, i.e.

$$\mathrm{LOPT}_{\mu_0,\lambda}(\mu^0, \mu^j) = OPT_\lambda(\mu^0, \mu^j).$$

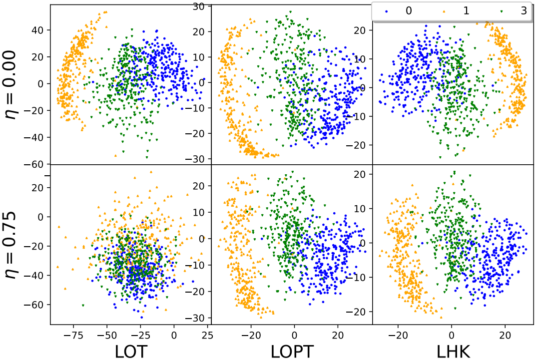
Application: Point cloud interpolation

Given two point clouds (e.g., digits 0 and 9) from MNIST, which are corrupted with uniformly distributed noise, how to find a reasonable interpolation between them?



PCA analysis

Given 900 digits from MNIST dataset corresponding to digits 0, 1, 3 in equal proportions. Each digit is a weighted point set with added $\eta * 100$ percentage of uniformly distributed noise.



Thank you