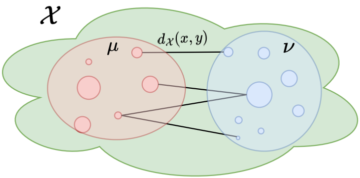
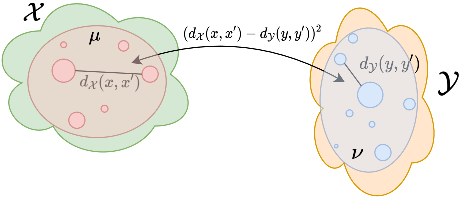
Entropic Gromov-Wasserstein between Gaussian Distributions

Khang Le

Joint work with Dung Le, Huy Nguyen, Dat Do, Tung Pham, Nhat Ho

July 21, 2022





$\mathrm{OT}(\mu, \nu)$

$\mathrm{GW}(\mu, \nu)$

$$\inf_{\pi \in \Pi(\mu,\nu)} \mathbb{E}_{\pi}\left[d_{\mathcal{X}}(X, Y)^2\right]$$

$$\inf_{\pi \in \Pi(\mu,\nu)} \mathbb{E}_{\pi \otimes \pi}\left[(d_{\mathcal{X}}(X, X') - d_{\mathcal{Y}}(Y, Y'))^2\right]$$

Problem

$$\mathrm{GW}_\varepsilon(\mu,\nu) := \inf_{\pi\in\Pi(\mu,\nu)} \mathbb{E}_{\pi\otimes\pi}\left[(d_{\mathcal{X}}(X,X') - d_{\mathcal{Y}}(Y,Y'))^2\right] + \varepsilon\mathrm{KL}(\pi\|\mu\otimes\nu),$$

We provide the closed-form of GW_ε for the setting:

- zero-mean Gaussian measures $\mu = \mathcal{N}(0,\Sigma_\mu), \nu = \mathcal{N}(0,\Sigma_\nu)$
- inner-product structures $d_{\mathcal{X}} = \langle\cdot,\cdot\rangle_{\mathcal{X}}, d_{\mathcal{Y}} = \langle\cdot,\cdot\rangle_{\mathcal{Y}}$

We denote this case as IGW_ε.

Invariant of IGW_ε

$$\underbrace{\mathrm{IGW}_\varepsilon(\mu, \nu)}_{\pi^* \text{ optimal}} = \underbrace{\mathrm{IGW}_\varepsilon(f^{\text{ortho}}_\sharp \mu, g^{\text{ortho}}_\sharp \nu)}_{(f^{\text{ortho}}, g^{\text{ortho}})\sharp \pi^* \text{ optimal}}$$

This means we only need to consider Gaussian measures with diagonal covariance matrices!

Main Result: Entropic IGW

$$\mu = \mathcal{N}(0, \mathrm{Diag}(\{\alpha_i\}_{i=1}^m)), \qquad \nu = \mathcal{N}\left(0, \mathrm{Diag}(\{\beta_i\}_{i=1}^n)\right)$$

$$\alpha_1 \geq \cdots \geq \alpha_m > 0, \qquad \beta_1 \geq \cdots \geq \beta_n > 0$$

$$m \geq n$$

Theorem

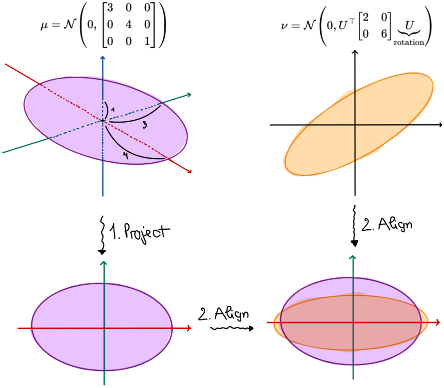
$$\begin{aligned}\mathrm{IGW}_\varepsilon(\mu, \nu) = \sum_{i=1}^m \alpha_i^2 + \sum_{i=1}^n \beta_i^2 - 2\sum_{i=1}^n (\alpha_i\beta_i - \varepsilon/4)^+ \\ + (\varepsilon/2)\sum_{i=1}^n (\log(\alpha_i\beta_i) - \log(\varepsilon/4))^+\end{aligned}$$

$$\pi^* = \mathcal{N}(0, \Sigma_{\pi^*}), \quad \Sigma_{\pi^*} = \begin{bmatrix} \Sigma_\mu & K \\ K^\top & \Sigma_\nu \end{bmatrix}, \quad K = \mathrm{Diag}\left(\left\{\sqrt{\alpha_i\beta_i\left(1 - \frac{\varepsilon}{4\alpha_i\beta_i}\right)^+}\right\}_{i=1}^n\right)$$

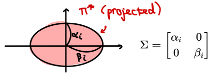
Intuition of Optimal Transportation

$$\mathrm{IGW}_{\varepsilon}(\mu,\nu) = \sum_{i=1}^{m} \alpha_i^2 + \sum_{i=1}^{n} \beta_i^2 - 2\sum_{i=1}^{n} (\alpha_i\beta_i - \varepsilon/4)^{+} + (\varepsilon/2)\sum_{i=1}^{n} (\log(\alpha_i\beta_i) - \log(\varepsilon/4))^{+}$$

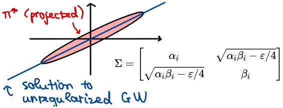
$$\pi^* = \mathcal{N}(0, \Sigma_{\pi^*}), \quad \Sigma_{\pi^*} = \begin{bmatrix} \Sigma_\mu & K \\ K^\top & \Sigma_\nu \end{bmatrix}, \quad K = \mathrm{Diag}\left(\left\{\sqrt{\alpha_i\beta_i\left(1 - \frac{\varepsilon}{4\alpha_i\beta_i}\right)^{+}}\right\}_{i=1}^{n}\right)$$



3. Check the effect of regularizer which forces π to look like $\mu \otimes \nu$:
   - Case $\alpha_i\beta_i \le \varepsilon/4$ (regularizer is stronger)
     π^* projected onto directions of α_i, β_i looks exactly like $\mu \otimes \nu$



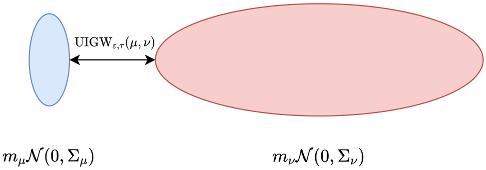
   - Case $\alpha_i\beta_i > \varepsilon/4$ (regularizer is weaker)
     π^* projected onto directions of α_i, β_i looks more like a solution of a GW problem, but it is still affected by regularizer.



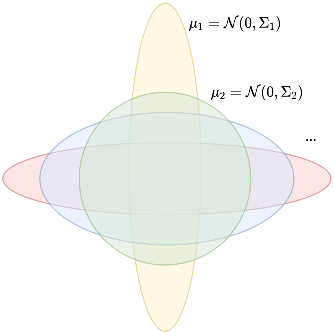
Variants

We also provide closed-form solutions for unbalanced IGW_ε and the (entropic-regularized) IGW barycenter problem (see the text for details).

Comparing scaled Gaussian measures



Averaging Gaussian Measures



$$\arg\min_{\mu} \sum_{k=1}^{K} \alpha_i \mathrm{IGW}_\varepsilon(\mu,\mu_i)$$

Thanks for listening!