What can we do with the large scale optimal transport problem?

@ SAMBa conference 2022

Fengpei Wang

Supervisors: Dr Clarice Poon and Dr Tony Shardlow

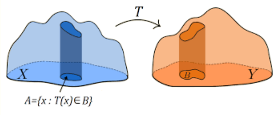
July 5, 2022

Introduction

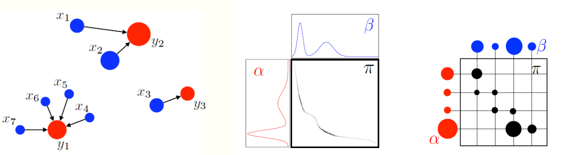
What is optimal transport?

Optimal transport (OT) seeks the best way of transforming one probability distribution into another

$$\min_{\mathbf{P}\in\Pi(a,b)} \sum_{i,j} \mathbf{C}_{i,j}\mathbf{P}_{i,j},$$



where $\mathbf{C}$ is the cost between two distributions, and the constraint set $\Pi(\mathbf{a},\mathbf{b}) = \left\{\mathbf{P} \in \mathbb{R}_+^{n\times m} : \mathbf{P}\mathbb{1} = \mathbf{a}, \mathbf{P}\mathbb{1} = \mathbf{b}\right\}$.

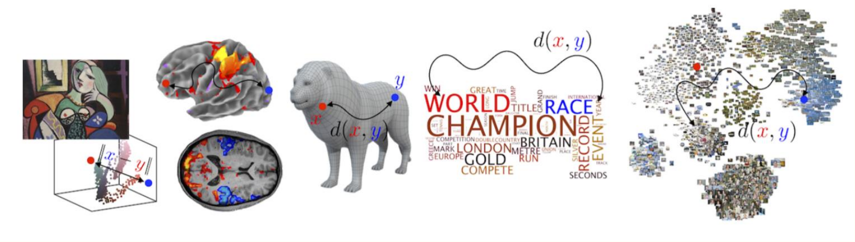


[1]

---

[1] Gabriel Peyré, Marco Cuturi, et al. "Computational optimal transport: With applications to data science". In: *Foundations and Trends® in Machine Learning* 11.5-6 (2019), pp. 355–607.

Applications



 [2]

[2] Peyré, Cuturi, et al., "Computational optimal transport: With applications to data science".

How do we solve a OT problem?

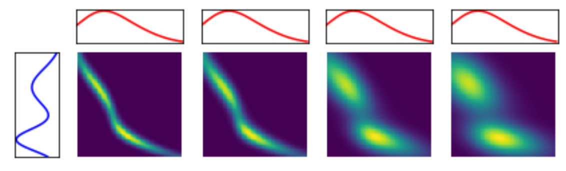
Entropic regularisation:

$$\min_{\mathbf{P}\in\Pi(\mathbf{a},\mathbf{b})} \sum_{i,j} \mathbf{c}_{i,j}\mathbf{P}_{i,j} + \epsilon \mathbf{P}_{i,j} \log\left(\mathbf{P}_{i,j} - 1\right).$$

Proposition: $\mathbf{P}_{i,j} = \mathbf{u}_i \mathbf{K} \mathbf{v}_j \quad \mathbf{K}_{i,j} = \exp\left(-\mathbf{C}_{i,k}/\epsilon\right)$

Marginal constraints: $\mathbf{u} \odot (\mathbf{K}\mathbf{v}) = \mathbf{a} \quad \mathbf{v} \odot \left(\mathbf{K}^T \mathbf{u}\right) = \mathbf{b}$

Sinkhorn iterations: $\mathbf{u} \leftarrow \frac{\mathbf{a}}{\mathbf{K}\mathbf{v}} \quad \mathbf{v} \leftarrow \frac{\mathbf{b}}{\mathbf{K}^T \mathbf{u}}$



Online Sinkhorn

Continuous OT

For general distributions α, β,

$$OT_\epsilon(\alpha, \beta) = \min_{\pi \in \Pi(\alpha,\beta)} \int c(x, y) \mathrm{d}\pi(x, y) + \epsilon \mathrm{KL}(\pi, \alpha \otimes \beta),$$

Continuous OT

For general distributions α, β,

$$OT_\epsilon(\alpha, \beta) = \min_{\pi \in \Pi(\alpha,\beta)} \int c(x, y) \mathrm{d}\pi(x, y) + \epsilon \mathrm{KL}(\pi, \alpha \otimes \beta),$$

Sinkhorn iterations:

$$u_{t+1} = \int \frac{1}{v_t} \exp\left(c/\epsilon\right) \mathrm{d}\beta$$

$$v_{t+1} = \int \frac{1}{u_t} \exp\left(c/\epsilon\right) \mathrm{d}\alpha,$$

Continuous OT

For general distributions α, β,

$$OT_\epsilon(\alpha, \beta) = \min_{\pi \in \Pi(\alpha,\beta)} \int c(x,y) \mathrm{d}\pi(x,y) + \epsilon \mathrm{KL}(\pi, \alpha \otimes \beta),$$

Sinkhorn iterations:

$$u_{t+1} = \int \frac{1}{v_t} \exp(c/\epsilon) \mathrm{d}\beta$$
$$v_{t+1} = \int \frac{1}{u_t} \exp(c/\epsilon) \mathrm{d}\alpha,$$

which can be rewritten as

$$(u_{t+1}, v_{t+1}) = (u_t, v_t) - \mathscr{F}_{\alpha,\beta}(u_t, v_t),$$

where $\mathscr{F}_{\alpha,\beta}(u,v) = \left(u - \int \frac{1}{v} \exp(c/\epsilon) \mathrm{d}\beta, v - \int \frac{1}{u} \exp(c/\epsilon) \mathrm{d}\alpha\right)$.

Development of Online Sinkhorn

$$(u_{t+1}, v_{t+1}) = (u_t, v_t) - \mathscr{F}_{\alpha,\beta}(u_t, v_t)$$

---

[3] Arthur Mensch and Gabriel Peyré. "Online Sinkhorn: Optimal Transport distances from sample streams". In: *arXiv e-prints* (2020), arXiv–2003.

Development of Online Sinkhorn

$$(u_{t+1}, v_{t+1}) = (u_t, v_t) - \mathscr{F}_{\alpha,\beta}(u_t, v_t)$$

The Randomised Sinkhorn[3] with samples newly drawn i.i.d. from α and β:

$$(u_{t+1}, v_{t+1}) = (u_t, v_t) - \mathscr{F}_{\hat{\alpha}_t,\hat{\beta}_t}(u_t, v_t),$$

where $\hat{\alpha}_t = \frac{1}{n_t}\sum_{i=1}^{n_t} \delta_{x_i}$, $\hat{\beta}_t = \frac{1}{n_t}\sum_{i=1}^{n_t} \delta_{y_i}$.

[3] Mensch and Peyré, "Online Sinkhorn: Optimal Transport distances from sample streams".

Development of Online Sinkhorn

$$(u_{t+1}, v_{t+1}) = (u_t, v_t) - \mathscr{F}_{\alpha,\beta}(u_t, v_t)$$

The Randomised Sinkhorn[3] with samples newly drawn i.i.d. from α and β:

$$(u_{t+1}, v_{t+1}) = (u_t, v_t) - \mathscr{F}_{\hat{\alpha}_t,\hat{\beta}_t}(u_t, v_t),$$

where $\hat{\alpha}_t = \frac{1}{n_t}\sum_{i=1}^{n_t}\delta_{x_i}$, $\hat{\beta}_t = \frac{1}{n_t}\sum_{i=1}^{n_t}\delta_{y_i}$.

$\Rightarrow$ Introduce the learning rate η_t:

$$(u_{t+1}, v_{t+1}) = (u_t, v_t) - \eta_t\mathscr{F}_{\hat{\alpha}_t,\hat{\beta}_t}(u_{t+1}, v_t).$$

[3] Mensch and Peyré, “Online Sinkhorn: Optimal Transport distances from sample streams”.

Low Rank OT

Problem formulation

Consider the low rank OT (LR-OT) problem in the following form

$$\min_{U,V\in\mathbb{R}_+^{n\times r}} \langle C, UV^T\rangle \text{ s.t. } \left(UV^T\right)\mathbb{1} = a \quad \text{and} \quad \left(VU^T\right)\mathbb{1} = b.$$

Problem formulation

Consider the low rank OT (LR-OT) problem in the following form

$$\min_{U,V \in \mathbb{R}_+^{n\times r}} \langle C, UV^T \rangle \text{ s.t. } \left(UV^T\right)\mathbb{1} = a \quad \text{and} \quad \left(VU^T\right)\mathbb{1} = b.$$

Further consider the regularised LR-OT problem:

$$\min_{U,V \in \mathbb{R}_+^{n\times r}} \langle C, UV^T \rangle + \epsilon\,(H(U) + H(V)) \text{ s.t. } \left(UV^T\right)\mathbb{1} = a \quad \text{and} \quad \left(VU^T\right)\mathbb{1} = b,$$

where $H(x) = x\log(x-1)$ is the entropy regularisation.

Results

We propose to Alternating Minimisation to solve the problem above:

for $i = 1$ to N **do**

$$U^{k+1} \leftarrow \min_{U} \langle C, U V^{k\,T} \rangle \text{ s.t. } \left(U V^{k\,T}\right) \mathbb{1} = a, \left(V^{k} U^{T}\right) \mathbb{1} = b$$

$$V^{k+1} \leftarrow \min_{V} \langle C, U^{k+1} V^{T} \rangle \text{ s.t. } \left(U^{k+1} V^{T}\right) \mathbb{1} = a, \left(V U^{k+1\,T}\right) \mathbb{1} = b$$

end for

Results

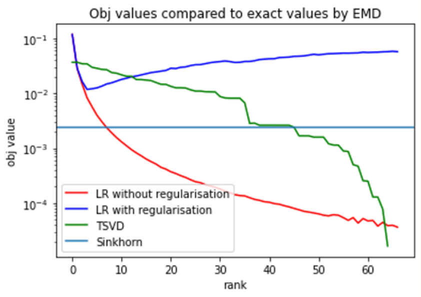
We propose to Alternating Minimisation to solve the problem above:

for $i = 1$ to N **do**

$$U^{k+1} \leftarrow \min_{U} \langle C, U V^{k\,T} \rangle \text{ s.t. } \left(U V^{k\,T}\right)\mathbb{1} = a, \left(V^k U^T\right)\mathbb{1} = b$$

$$V^{k+1} \leftarrow \min_{V} \langle C, U^{k+1} V^{T} \rangle \text{ s.t. } \left(U^{k+1} V^{T}\right)\mathbb{1} = a, \left(V U^{k+1\,T}\right)\mathbb{1} = b$$

end for



Main Takeaways

1. Introduction to optimal transport
2. Online Sinkhorn – for continuous OT
3. Low rank OT – to reduce the complexity

Thank you!