Unbalanced Optimal Transport

Efficient solutions for outlier-robust machine learning

Laetitia Chapel
laetitia.chapel@irisa.fr

IRISA, Rennes, France

Institut Agro Rennes-Angers

Optimal Transport and Machine Learning @Neurips 2023

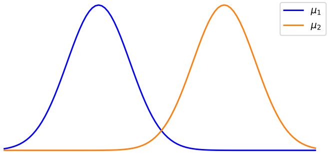
Table of Contents

Optimal transport
Balanced Optimal transport: Monge formulation

- **Balanced** optimal transport

$$\mathcal{OT}(\mu_1, \mu_2) \triangleq \inf \int c(x, \boldsymbol{t}(x))\, d\mu_1(x)$$

where $\boldsymbol{t}$ is a **transport map** and $\boldsymbol{t}_{\#}\mu_1 = \mu_2$

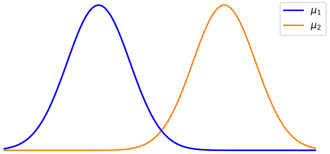


Optimal transport
Balanced Optimal transport: Monge formulation

- **Balanced** optimal transport

$$\mathcal{OT}(\mu_1, \mu_2) \triangleq \inf \int c(x, \boldsymbol{t}(x))\, d\mu_1(x)$$

where $\boldsymbol{t}$ is a **transport map** and $\boldsymbol{t}_{\#}\mu_1 = \mu_2$

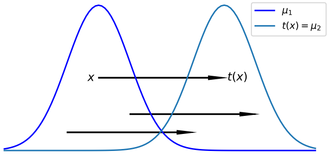


Optimal transport
Balanced Optimal transport: Monge formulation

- **Balanced** optimal transport

$$\mathcal{OT}(\mu_1, \mu_2) \triangleq \inf \int c(x, \boldsymbol{t}(x))\, d\mu_1(x)$$

where $\boldsymbol{t}$ is a **transport map** and $\boldsymbol{t}_{\#}\mu_1 = \mu_2$



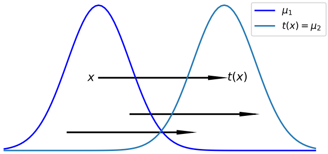
Defines for each particle located at x what is its destination $\boldsymbol{t}(x)$

Optimal transport
Balanced Optimal transport: Monge formulation

- **Balanced** optimal transport

$$\mathcal{OT}(\mu_1, \mu_2) \triangleq \inf \int c(x, \boldsymbol{t}(x))\, d\mu_1(x)$$

where $\boldsymbol{t}$ is a **transport map** and $\boldsymbol{t}_{\#}\mu_1 = \mu_2$



Defines for each particle located at x what is its destination $\boldsymbol{t}(x)$

- implies that μ_1 and μ_2 have the same masses (no mass creation nor destruction)

Optimal transport
Balanced Optimal transport: Kantorovich formulation

- **Balanced** optimal transport

Linear loss

$$\mathcal{OT}(\mu_1, \mu_2) \triangleq \inf_{\gamma \in \Gamma(\mu_1,\mu_2)} \int_{X \times Y} c(x,y) \, d\gamma(x,y)$$

where $\Gamma(\mu_1, \mu_2) \stackrel{\text{def}}{=} \{\gamma \in \mathcal{M}_+(X \times Y) \text{ s.t. } (\pi_x)_\# \gamma = \mu_1 \text{ and } (\pi_y)_\# \gamma = \mu_2 \}$ with $\pi_x : X \times Y \rightarrow X$.

Marginal constraints

Optimal transport

Balanced Optimal transport: Kantorovich formulation

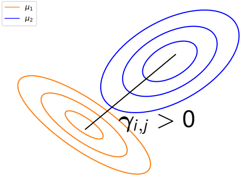
- **Balanced** optimal transport

Linear loss

$$\mathcal{OT}(\mu_1, \mu_2) \triangleq \inf_{\gamma \in \Gamma(\mu_1, \mu_2)} \int_{X \times Y} c(x,y) \, d\gamma(x,y)$$

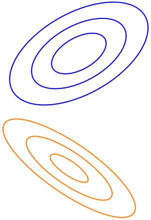
where $\Gamma(\mu_1, \mu_2) \stackrel{\text{def}}{=} \{\gamma \in \mathcal{M}_+(X \times Y) \text{ s.t. } (\pi_x)_\# \gamma = \mu_1 \text{ and } (\pi_y)_\# \gamma = \mu_2\}$ with $\pi_x : X \times Y \to X$.

Marginal constraints



with $(\pi_x)_\# \gamma = \mu_1$

and $(\pi_y)_\# \gamma = \mu_2$

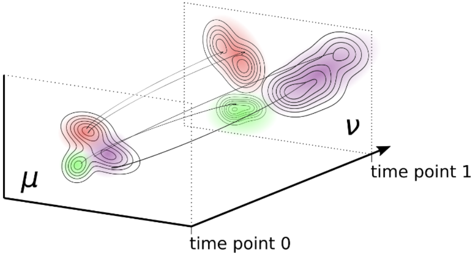


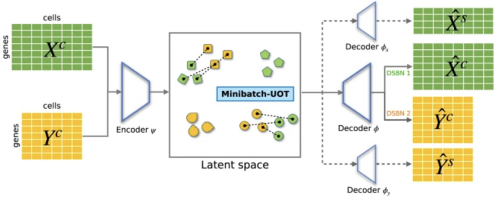
The **transport plan** $\gamma(x,y)$ specifies for each pair (x,y) how many particles go from x to y

- still implies that μ_1 and μ_2 have the same masses

Optimal transport
Balanced Optimal transport in action

- But, in many applications, we **cannot/do not want to have the same masses** and we may want to **discard some outliers or ood**
  - In biology, there are different cell proliferation or death in different sub-populations [8] or we may want to identify common genes [3].

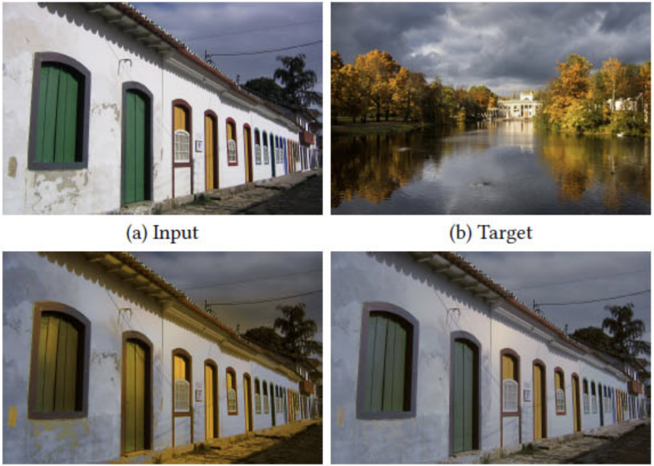




Optimal transport

Balanced Optimal transport in action

- But, in many applications, we **cannot/do not want to have the same masses** and we may want to **discard some outliers or ood**
  - In biology, there are different cell proliferation or death in different sub-populations [8] or we may want to identify common genes [3].
  - In color transfer, to account for different proportions of colors [1]

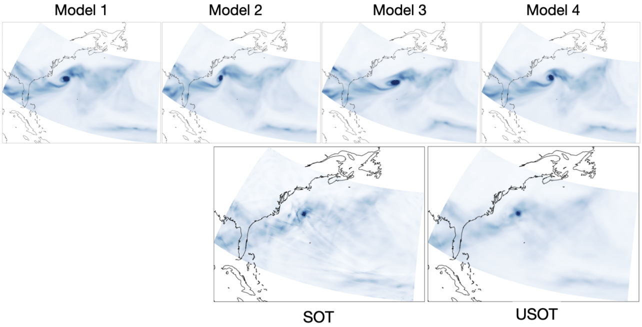


(a) Input (b) Target

(c) Full histogram matching (d) Partial histogram matching

Optimal transport
Balanced Optimal transport in action

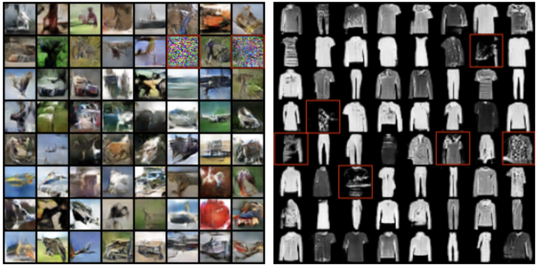
- But, in many applications, we **cannot/do not want to have the same masses** and we may want to **discard some outliers or ood**
  - In biology, there are different cell proliferation or death in different sub-populations [8] or we may want to identify common genes [3].
  - In color transfer, to account for different proportions of colors [1]
  - In geophysics, when averaging different models [6]



Optimal transport

Balanced Optimal transport in action

- But, in many applications, we **cannot/do not want to have the same masses** and we may want to **discard some outliers or ood**
  - In biology, there are different cell proliferation or death in different sub-populations [8] or we may want to identify common genes [3].
  - In color transfer, to account for different proportions of colors [1]
  - In geophysics, when averaging different models [6]
  - In machine learning, when some of the points are out of the distribution, for instance with WGAN [7]



Optimal transport
Balanced Optimal transport in action

- But, in many applications, we **cannot/do not want to have the same masses** and we may want to **discard some outliers or ood**
  - In biology, there are different cell proliferation or death in different sub-populations [8] or we may want to identify common genes [3].
  - In color transfer, to account for different proportions of colors [1]
  - In geophysics, when averaging different models [6]

- How to define outlier-robust OT?
  - define robust variants of OT (e.g. medians of means OT)
  - *pick a dedicated ground cost* to avoid too much influence of samples that are too far away from the distributions
  - **allow for some mass variation**

Table of Contents

Unbalanced Optimal Transport
Definition

- **key idea**: relax the mass conservation constraint

NUMERICAL RESOLUTION OF AN "UNBALANCED" MASS TRANSPORT PROBLEM

JEAN-DAVID BENAMOU[1]

Abstract. We introduce a modification of the Monge–Kantorovitch problem of exponent 2 which accommodates non balanced initial and final densities. The augmented Lagrangian numerical method introduced in [6] is adapted to this "unbalanced" problem. We illustrate the usability of this method on an idealized error estimation problem in meteorology.

Mathematics Subject Classification. 35J60, 65K10, 78A05, 90B99.

Received: April 1st, 2003.

2.4. The mixed distance

In this paper we propose to work on unbalanced data by considering the mixed Wasserstein/L^2-distance in the following sense: given two possibly unbalanced densities ρ_0 and ρ_1, find $\tilde{\rho}_1$ – the closest density to ρ_1 in the L^2-sense – which minimizes the Wasserstein distance $d_{\text{wass}}(\rho_0, \tilde{\rho}_1)$. It can be formulated as

$$\inf_{\tilde{\rho}_1} \left\{ d_{\text{wass}}(\rho_0, \tilde{\rho}_1)^2 + \frac{\gamma}{2} d_{L^2}(\tilde{\rho}_1, \rho_1)^2 \right\} \tag{16}$$

Unbalanced Optimal Transport
Definition

- Regularizing the **balanced** optimal transport, by replacing the hard constraints with some divergences

$$\mathcal{UOT}(\mu_1,\mu_2) \triangleq \inf_{\gamma\geq 0} \int_{\mathbb{R}^d\times\mathbb{R}^d} \underbrace{c(x,y)}_{\text{Linear loss}} d\gamma(x,y) + \underbrace{\lambda}_{\text{reg}} \Big(\underbrace{\mathcal{D}_\varphi((\pi^1)_\#\gamma|\mu_1) + \mathcal{D}_\varphi((\pi^2)_\#\gamma|\mu_2)}_{\text{Marginal constraints}} \Big)$$

with $\lambda \geq 0$: relaxing the constraints.
When $\lambda \rightarrow \infty$ we recover the balanced OT problem.

Unbalanced Optimal Transport
Definition

- Regularizing the **balanced** optimal transport, by replacing the hard constraints with some divergences

Linear loss

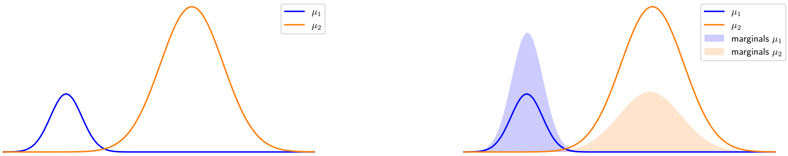
$$\mathcal{UOT}(\mu_1, \mu_2) \triangleq \inf_{\gamma \geq 0} \int_{\mathbb{R}^d \times \mathbb{R}^d} c(x,y)\, d\gamma(x,y)$$

reg

$$+ \lambda \left(\mathcal{D}_\varphi((\pi^1)_\# \gamma | \mu_1) + \mathcal{D}_\varphi((\pi^2)_\# \gamma | \mu_2) \right)$$

Marginal constraints

with $\lambda \geq 0$: relaxing the constraints.
When $\lambda \to \infty$ we recover the balanced OT problem.
When the masses are different



Unbalanced Optimal Transport
Definition

- Regularizing the **balanced** optimal transport, by replacing the hard constraints with some divergences

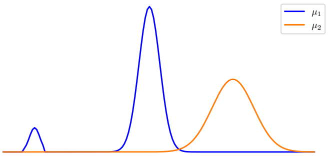
Linear loss

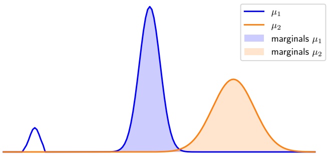
$$\mathcal{UOT}(\mu_1, \mu_2) \triangleq \inf_{\gamma \geq 0} \int_{\mathbb{R}^d \times \mathbb{R}^d} c(x, y) \, d\gamma(x, y) + \lambda \left(\mathcal{D}_\varphi((\pi^1)_\# \gamma | \mu_1) + \mathcal{D}_\varphi((\pi^2)_\# \gamma | \mu_2) \right)$$

reg

Marginal constraints

with $\lambda \geq 0$: relaxing the constraints.
When $\lambda \to \infty$ we recover the balanced OT problem.
When there are some outliers





Unbalanced Optimal Transport
Definition

- Regularizing the **balanced** optimal transport, by replacing the hard constraints with some divergences

$$\mathcal{UOT}(\mu_1, \mu_2) \triangleq \inf_{\gamma \geq 0} \int_{\mathbb{R}^d \times \mathbb{R}^d} \underbrace{c(x,y)}_{\text{Linear loss}} d\gamma(x,y) + \underbrace{\lambda}_{\text{reg}} \Big(\underbrace{\mathcal{D}_\varphi((\pi^1)_\# \gamma | \mu_1) + \mathcal{D}_\varphi((\pi^2)_\# \gamma | \mu_2)}_{\text{Marginal constraints}} \Big)$$

  with $\lambda \geq 0$: relaxing the constraints.
  When $\lambda \to \infty$ we recover the balanced OT problem.
- has similar properties as OT (is a distance, weak convergence etc.)

Unbalanced Optimal Transport
Definition

- Regularizing the **balanced** optimal transport, by replacing the hard constraints with some divergences

$$\mathcal{UOT}(\mu_1, \mu_2) \triangleq \inf_{\gamma \geq 0} \int_{\mathbb{R}^d \times \mathbb{R}^d} \underbrace{c(x,y)}_{\text{Linear loss}} d\gamma(x,y) + \underbrace{\lambda}_{\text{reg}} \Big(\underbrace{\mathcal{D}_\varphi((\pi^1)_\# \gamma | \mu_1) + \mathcal{D}_\varphi((\pi^2)_\# \gamma | \mu_2)}_{\text{Marginal constraints}} \Big)$$

  with $\lambda \geq 0$: relaxing the constraints.
  When $\lambda \to \infty$ we recover the balanced OT problem.
- has similar properties as OT (is a distance, weak convergence etc.)
- questions:
  - Which D_φ?
  - how to solve the problem?

Unbalanced Optimal Transport
Discrete UOT

- When the distributions are discrete $\mu_1 = \sum_{i=1}^{n} h_i \delta_{x_i}$ and $\mu_2 = \sum_{j=1}^{m} g_j \delta_{y_j}$, we usually solve

Linear loss (on positions $c(x_i, y_j)$)

$$\mathcal{UOT}(\mu_1, \mu_2) \triangleq \min_{\gamma \geq 0} \sum_{i,j} C_{i,j}\gamma_{i,j} + \lambda \left(\mathcal{D}_\varphi(\gamma \mathbb{1}_m | h) + \mathcal{D}_\varphi(\gamma^\top \mathbb{1}_n | g) \right)$$

Marginal constraints (on the masses only)

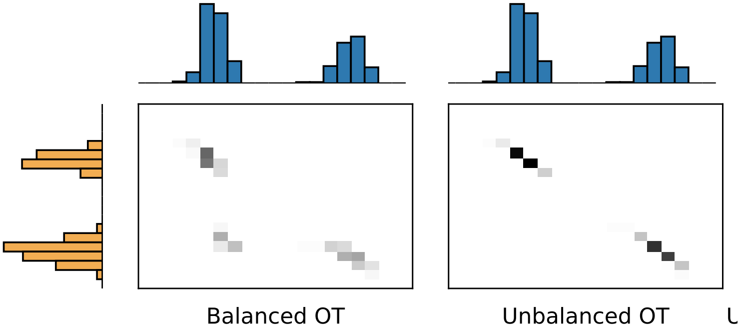
Unbalanced Optimal Transport
Discrete UOT

- When the distributions are discrete $\mu_1 = \sum_{i=1}^n h_i \delta_{x_i}$ and $\mu_2 = \sum_{j=1}^m g_j \delta_{y_j}$, we usually solve

Linear loss (on positions $c(x_i, y_j)$)

$$\mathcal{UOT}(\mu_1, \mu_2) \triangleq \min_{\gamma \geq 0} \sum_{i,j} C_{i,j}\gamma_{i,j} + \lambda \left(\mathcal{D}_\varphi(\gamma \mathbf{1}_m | h) + \mathcal{D}_\varphi(\gamma^\top \mathbf{1}_n | g) \right)$$

Marginal constraints (on the masses only)



Unbalanced Optimal Transport
Discrete UOT

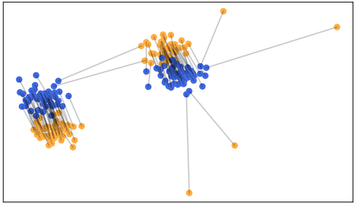
- When the distributions are discrete $\mu_1 = \sum_{i=1}^{n} h_i \delta_{x_i}$ and $\mu_2 = \sum_{j=1}^{m} g_j \delta_{y_j}$, we usually solve

Linear loss (on positions $c(x_i, y_j)$)

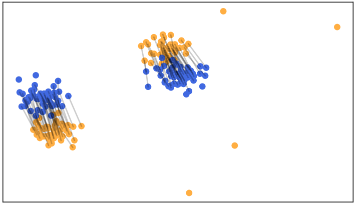
$$\mathcal{UOT}(\mu_1, \mu_2) \triangleq \min_{\gamma \geq 0} \sum_{i,j} C_{i,j}\gamma_{i,j} + \lambda \left(\mathcal{D}_\varphi(\gamma \mathbb{1}_m | h) + \mathcal{D}_\varphi(\gamma^\top \mathbb{1}_n | g) \right)$$

Marginal constraints (on the masses only)

OT, cost = 0.20



Unbalanced OT, cost = 0.14



Unbalanced Optimal Transport
Partial Optimal Transport

- **Unbalanced OT with L_1 penalty**

$$\mathcal{UOT}(\mu_1, \mu_2) \triangleq \min_{\gamma \geq 0} \sum_{i,j} C_{i,j}\gamma_{i,j} + \lambda \left(\|\gamma 1_m - h\|_1 + \|\gamma^\top 1_n - g\|_1 \right)$$

is equivalent to writing

$$\mathcal{UOT}(\mu_1, \mu_2) \triangleq \inf_{\gamma \in \Gamma_{\leq(\mu_1,\mu_2)}} \sum_{i,j} C_{i,j}\gamma_{i,j}$$

where $\Gamma_{\leq(\mu_1,\mu_2)} = \{\gamma \geq 0,\ \gamma 1_m \leq h \text{ and } \gamma^\top 1_n \leq g \text{ and } 1_n^\top \gamma 1_m = s\}$

amount of mass to be transported

Unbalanced Optimal Transport
Partial Optimal Transport

- **Unbalanced OT with L_1 penalty**

$$\mathcal{UOT}(\mu_1, \mu_2) \triangleq \inf_{\gamma \in \Gamma_{\leq(\mu_1,\mu_2)}} \sum_{i,j} C_{i,j}\gamma_{i,j}$$

where $\Gamma_{\leq(\mu_1,\mu_2)} = \{\gamma \geq 0,\ \gamma 1_m \leq h \text{ and } \gamma^\top 1_n \leq g$ and $1_n^\top \gamma 1_m = s\ \}$

- Can be solved easily by adding dummy points $h_{n+1} = \|\boldsymbol{g}\|_1 - s$ and $g_{m+1} = \|\boldsymbol{h}\|_1 - s$ with null cost and solve the extended OT problem [4, 2]

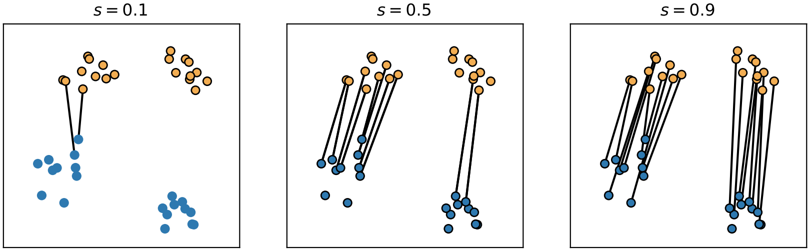
Unbalanced Optimal Transport
Partial Optimal Transport

- **Unbalanced OT with L_1 penalty**

$$\mathcal{UOT}(\mu_1, \mu_2) \triangleq \inf_{\gamma \in \Gamma_{\leq}(\mu_1,\mu_2)} \sum_{i,j} C_{i,j}\gamma_{i,j}$$

where $\Gamma_{\leq(\mu_1,\mu_2)} = \{\gamma \geq 0,\ \gamma 1_m \leq h \text{ and } \gamma^\top 1_n \leq g$ and $1_n^\top \gamma 1_m = s\ \}$

- Can be solved easily by adding dummy points $h_{n+1} = \|\boldsymbol{g}\|_1 - s$ and $g_{m+1} = \|\boldsymbol{h}\|_1 - s$ with null cost and solve the extended OT problem [4, 2]



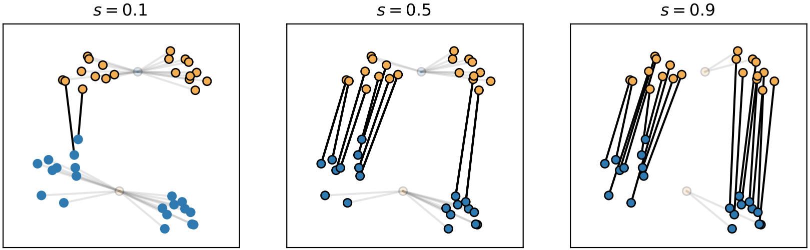
Unbalanced Optimal Transport
Partial Optimal Transport

- **Unbalanced OT with L_1 penalty**

$$\mathcal{UOT}(\mu_1, \mu_2) \triangleq \inf_{\gamma \in \Gamma_{\leq}(\mu_1,\mu_2)} \sum_{i,j} C_{i,j}\gamma_{i,j}$$

where $\Gamma_{\leq(\mu_1,\mu_2)} = \{\gamma \geq 0,\ \gamma 1_m \leq h \text{ and } \gamma^\top 1_n \leq g \text{ and } 1_n^\top \gamma 1_m = s\ \}$

- Can be solved easily by adding dummy points $h_{n+1} = \|\boldsymbol{g}\|_1 - s$ and $g_{m+1} = \|\boldsymbol{h}\|_1 - s$ with null cost and solve the extended OT problem [4, 2]



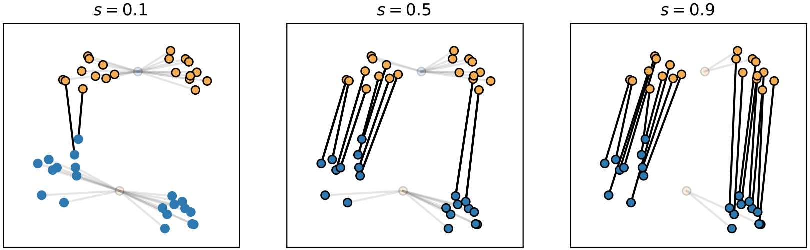
Unbalanced Optimal Transport
Partial Optimal Transport

- **Unbalanced OT with L_1 penalty**

$$\mathcal{UOT}(\mu_1, \mu_2) \triangleq \inf_{\gamma \in \Gamma_{\leq(\mu_1,\mu_2)}} \sum_{i,j} C_{i,j}\gamma_{i,j}$$

where $\Gamma_{\leq(\mu_1,\mu_2)} = \{\gamma \geq 0, \; \gamma 1_m \leq h \text{ and } \gamma^\top 1_n \leq g \; \text{ and } \; 1_n^\top \gamma 1_m = s \; \}$

- Can be solved easily by adding dummy points $h_{n+1} = \|g\|_1 - s$ and $g_{m+1} = \|h\|_1 - s$ with null cost and solve the extended OT problem [4, 2]



- Any OT solver can be used!

Unbalanced Optimal Transport
Unbalanced Optimal Transport with KL

- **Unbalanced OT with *KL* penalty**

$$\mathcal{UOT}(\mu_1, \mu_2) \triangleq \min_{\gamma \geq 0} \sum_{i,j} C_{i,j}\gamma_{i,j} + \lambda \left(\mathrm{KL}(\gamma 1_m, \boldsymbol{h}) + \mathrm{KL}(\gamma^\top 1_n, \boldsymbol{g}) \right)$$

Unbalanced Optimal Transport
Unbalanced Optimal Transport with KL

- **Unbalanced OT with *KL* penalty**

$$\mathcal{UOT}(\mu_1, \mu_2) \triangleq \min_{\gamma \geq 0} \sum_{i,j} C_{i,j}\gamma_{i,j} + \lambda \left(\mathrm{KL}(\gamma \mathbb{1}_m, \boldsymbol{h}) + \mathrm{KL}(\gamma^\top \mathbb{1}_n, \boldsymbol{g}) \right)$$

- Use a Majorize-Minimization algorithm to solve the problem [5]
  - Deterministic updates
  - Resembles the Sinkhorn algorithm, allows for GPU computation

$$\gamma^{(k+1)} = \mathrm{diag}\left(\frac{\boldsymbol{g}}{\gamma^{(k)}\mathbb{1}_m}\right)^{\frac{1}{2}} \left(\gamma^{(k)} \odot \exp\left(-\frac{C}{2\lambda}\right)\right) \mathrm{diag}\left(\frac{\boldsymbol{h}}{\gamma^{(k)\top}\mathbb{1}_n}\right)^{\frac{1}{2}}$$

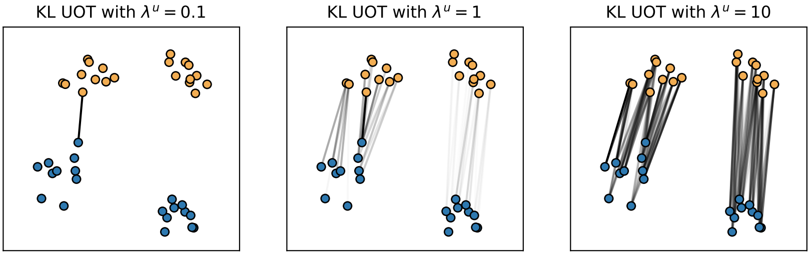
Unbalanced Optimal Transport
Unbalanced Optimal Transport with KL

- **Unbalanced OT with *KL* penalty**

$$\mathcal{UOT}(\mu_1, \mu_2) \triangleq \min_{\gamma \geq 0} \sum_{i,j} C_{i,j}\gamma_{i,j} + \lambda \left(\mathrm{KL}(\gamma 1_m, \boldsymbol{h}) + \mathrm{KL}(\gamma^\top 1_n, \boldsymbol{g}) \right)$$

- Use a Majorize-Minimization algorithm to solve the problem [5]
  - Deterministic updates
  - Resembles the Sinkhorn algorithm, allows for GPU computation

$$\gamma^{(k+1)} = \mathrm{diag}\left(\frac{\boldsymbol{g}}{\gamma^{(k)} 1_m}\right)^{\frac{1}{2}} \left(\gamma^{(k)} \odot \exp\left(-\frac{C}{2\lambda}\right)\right) \mathrm{diag}\left(\frac{\boldsymbol{h}}{\gamma^{(k)\top} 1_n}\right)^{\frac{1}{2}}$$



Unbalanced Optimal Transport

Unbalanced Optimal Transport with L2

- **Unbalanced OT with $L2$ penalty**

$$\mathcal{UOT}(\mu_1,\mu_2) \triangleq \min_{\gamma\geq 0} \sum_{i,j} C_{i,j}\gamma_{i,j} + \lambda \left(\|\boldsymbol{\gamma}\mathbf{1}_m - h\|_2^2 + \|\boldsymbol{\gamma}^\top\mathbf{1}_n - g\|_2^2 \right)$$

Unbalanced Optimal Transport
Unbalanced Optimal Transport with L2

- **Unbalanced OT with $L2$ penalty**

$$\mathcal{UOT}(\mu_1, \mu_2) \triangleq \min_{\gamma \geq 0} \sum_{i,j} C_{i,j}\gamma_{i,j} + \lambda \left(\|\boldsymbol{\gamma} \mathbf{1}_m - h\|_2^2 + \|\boldsymbol{\gamma}^\top \mathbf{1}_n - g\|_2^2 \right)$$

- When rewritten in a vectorial form:

$$\mathcal{UOT}(\mu_1, \mu_2) \triangleq \min_{\gamma \geq 0} \|\boldsymbol{H}\boldsymbol{\gamma}_v - \boldsymbol{y}\|_2^2 + \frac{1}{\lambda}\boldsymbol{c}^\top \|\gamma_v\|_1$$

where $\boldsymbol{c} = \text{vec}(\boldsymbol{C})$, $\gamma_v = \text{vec}(\gamma)$, $\boldsymbol{y}^\top = [\boldsymbol{h}^\top, \boldsymbol{g}^\top]$ and $\boldsymbol{H}$ is a design matrix.

Unbalanced Optimal Transport
Unbalanced Optimal Transport with L2

- **Unbalanced OT with $L2$ penalty**

$$\mathcal{UOT}(\mu_1, \mu_2) \triangleq \min_{\gamma \geq 0} \sum_{i,j} C_{i,j}\gamma_{i,j} + \lambda \left(\|\boldsymbol{\gamma} 1_m - h\|_2^2 + \|\boldsymbol{\gamma}^\top 1_n - g\|_2^2 \right)$$

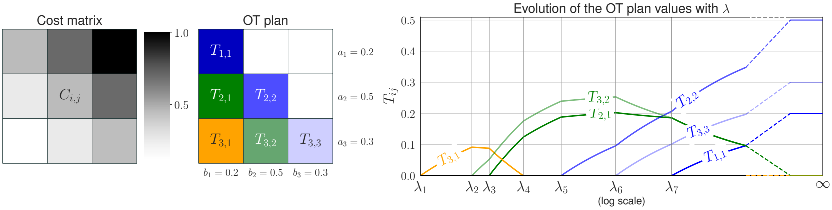
- When rewritten in a vectorial form:

$$\mathcal{UOT}(\mu_1, \mu_2) \triangleq \min_{\gamma \geq 0} \|\boldsymbol{H}\gamma_v - \boldsymbol{y}\|_2^2 + \frac{1}{\lambda}\boldsymbol{c}^\top \|\gamma_v\|_1$$

  where $\boldsymbol{c} = \mathrm{vec}(\boldsymbol{C})$, $\gamma_v = \mathrm{vec}(\gamma)$, $\boldsymbol{y}^\top = [\boldsymbol{h}^\top, \boldsymbol{g}^\top]$ and $\boldsymbol{H}$ is a design matrix.
- is a *classical* linear regression with positivity constraints, a sparse design matrix and a weighted L1 (Lasso) regularization
- we can borrow the tools from a large literature on solving those problems!

Unbalanced Optimal Transport
Unbalanced Optimal Transport with L2

- **Regularization path of UOT: a LARS-like algorithm**
- With quadratic divergence, solutions are piecewise linear with $\frac{1}{\lambda}$
- We can find the set of all solutions for all λ values
  1. start with $\lambda = 0$
  2. loop
  3. increase λ until there is a change on the support of γ_v
  4. update γ_v (incremental resolution of linear equations)
  5. repeat until $\lambda = \infty$



Unbalanced Optimal Transport
Unbalanced Optimal Transport with L2

- **Regularization path of UOT: a LARS-like algorithm**
- With quadratic divergence, solutions are piecewise linear with $\frac{1}{\lambda}$
- We can find the set of all solutions for all λ values
  1. start with $\lambda = 0$
  2. loop
  3. increase λ until there is a change on the support of γ_V
  4. update γ_V (incremental resolution of linear equations)
  5. repeat until $\lambda = \infty$

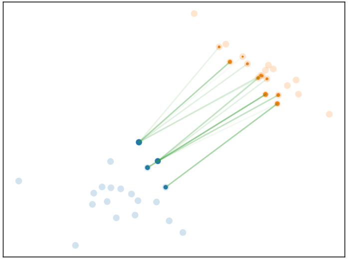


Table of Contents

Unbalanced Optimal Transport
Conclusion and pen challenges

- Conclusion
  - UOT is mandatory for many applications
  - (many) efficient solvers exist
  - implementation in POT python toolbox [1]
- Some open challenges
  - outlier removal?
  - which statistical guarantees?
  - enlarging the discrete *classical* formulation?

Mokthar Alaya Cédric Févotte Rémi Flamary Gilles Gasso

[1] figures have been generated with POT https://pythonot.github.io/, thanks @alex Tual for some layout

Unbalanced Optimal Transport

Efficient solutions for outlier-robust machine learning

Laetitia Chapel
laetitia.chapel@irisa.fr

IRISA, Rennes, France

Institut Agro Rennes-Angers

Optimal Transport and Machine Learning @Neurips 2023

Table of Contents

Bibliography I

[1] Nicolas Bonneel and David Coeurjolly. “Spot: sliced partial optimal transport”. In: *ACM Transactions on Graphics (TOG)* (2019).

[2] Luis A Caffarelli and Robert J McCann. “Free boundaries in optimal transport and Monge-Ampere obstacle problems”. In: *Annals of mathematics* (2010).

[3] Kai Cao et al. “A unified computational framework for single-cell data integration with optimal transport”. In: *Nature Communications* (2022).

[4] Laetitia Chapel, Mokhtar Z Alaya, and Gilles Gasso. “Partial optimal tranport with applications on positive-unlabeled learning”. In: *NeurIPS* (2020).

[5] Laetitia Chapel et al. “Unbalanced optimal transport through non-negative penalized linear regression”. In: *NeurIPS* (2021).

[6] Thibault Séjourné et al. “Unbalanced Optimal Transport meets Sliced-Wasserstein”. In: *arXiv preprint arXiv:2306.07176* (2023).

[7] G. Staerman et al. “When OT meets MoM: Robust estimation of Wasserstein Distance”. In: *AISTATS*. 2021.

[8] Karren D Yang and Caroline Uhler. “Scalable Unbalanced Optimal Transport using Generative Adversarial Networks”. In: *International Conference on Learning Representations*. 2018.