Data Driven Conditional Optimal Transport

Esteban G Tabak[1], Giulio Trigila[2], Wenjun Zhao[1]

[1] Courant Institute of Mathematical Sciences, New York University; [2] Baruch College, The City University of New York

Overview

A data driven procedure is developed to compute the optimal map between two conditional probabilities $\rho(x|z_1, \dots, z_L)$ and $\mu(y|z_1, \dots, z_L)$ depending on a set of covariates z_i.

The procedure is tested on:

- ACIC Data Analysis Challenge 2017 dataset;
- non uniform lightness transfer between images;
- fresco restoration for Sistine Chapel.

(a): OT

(b): COT

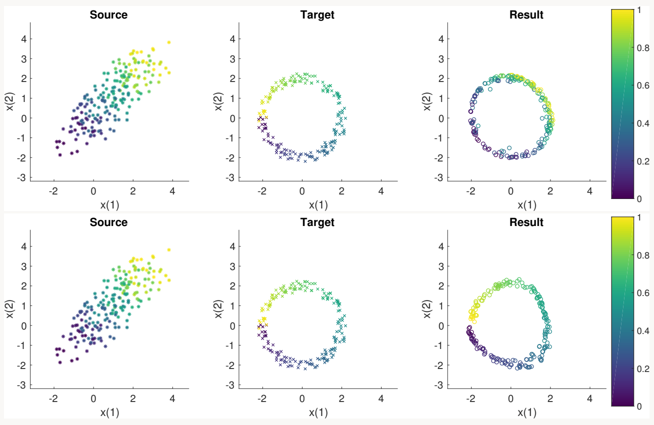


Figure: Illustration of OT/COT from a segment to a circle.

Formulations

Consider the conditional optimal transport between $\rho(:|z)$ and $\mu(:|z)$.

- **Equality constrained formulation**:

$$\forall z \begin{cases} \min_{T(:,z)} \int c(T(x,z),x)\rho(x|z)dx \\ T\#\rho(:|z) = \mu(:|z), \text{or } : D_{KL}(T\#\rho(:|z), \mu(:|z)) = 0 \end{cases}$$

Donsker-Varadhan variational formula:

$$D_{KL}(\rho, \mu) = \max_{g} \left[\int g(x)\rho(x)dx - \log\left(\int e^{g(x)}\mu(x)\, dx \right) \right],$$

Chain rule of KL divergence:

$$D_{KL}(\rho_1(x|z)||\rho_2(x|z)) = D_{KL}(\rho_1(x,z)||\rho_2(x,z)) - D_{KL}(\gamma_1(z)||\gamma_2(z)).$$

- **Unconstrained Minimax formulation**:

$$\min_{T} \max_{g,\lambda} \int c(T(x,z),x)d\rho(x,z) + \lambda \left[\int g(T(x,z),z)d\rho(x,z) - \log\left(\int e^{g(y,z)} d\mu(y,z) \right) \right]$$

- **Data-driven version**:

$$\min_{T} \max_{g,\lambda} \left[\frac{1}{n} \sum_i \Big(c(T(x_i,z_i),x_i) + \lambda g(T(x_i,z_i),z_i) \Big) - \lambda \log\left(\frac{1}{m} \sum_j e^{g(y_j,z_j)} \right) \right].$$

Parameterization of flows

$$T^n(x^i,z^i) = E^n\left(T^{n-1}(x^i,z^i), z^i \right).$$

- Evlovling Gaussian mixtures;
- Extended map compositions;
- Neural networks.

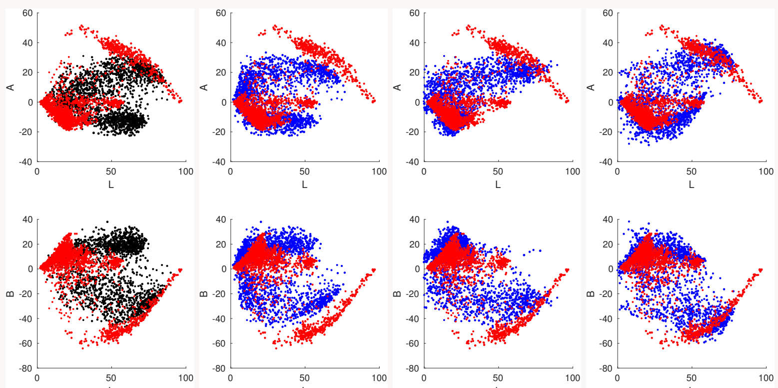
Motivation

- Optimal transport can be used to quantify the changes in probability distribution of quantities.
- The distributions depend on many covariates, hence one seeks the transport as a function of them.
- The need for conditional optimal transport is particularly apparent when the distributions for the covariates for the source and target distributions are unbalanced.
- Conditional transport provides a very flexible toolbox for data analysis, as the choice of which variables are conditioned to which others is left at the discretion of the analyst.

Application – lightness transfer

1D OT: lightness transferred;
COT: lightness transfer conditioned on color contrasts;
3D OT: lightness/color contrasts all transferred.

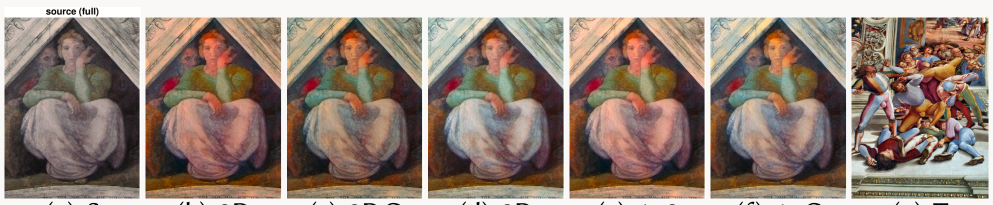
(a) Source (b) 1D OT (c) COT (d) 3D OT (e) Target



(f) S-LAB (g) 1DOT-LAB (h) COT-LAB (i) 3DOT-LAB

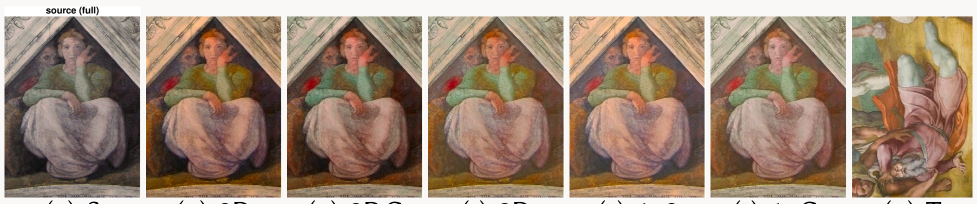
Application – fresco restoration

Source: Michelangelo's Jesse Spandrel from the Sistine Chapel;
Target: Cathedral of Orvieto/actual restoration/Pauline Chapel



(a) S (b) 2D (c) 2DC (d) 3D (e) 1+2 (f) 1+C (g) T

(h) S (i) 2D (j) 2DC (k) 3D (l) 1+2 (m) 1+C (n) T



(o) S (p) 2D (q) 2DC (r) 3D (s) 1+2 (t) 1+C (u) T