Making transport more robust and interpretable by moving data through a small number of anchor points

Chi-Heng Lin¹, Mehdi Azabou², Eva L. Dyer²
¹School of Electrical and Computer Engineering, Georgia Tech, ²Machine Learning Program, Georgia Tech, ²Coulter Department of Biomedical Engineering, Georgia Tech, ²Coulter Department of Biomedical Engineering, Emory University

Abstract
Optimal transport (OT) is a widely used technique for distribution alignment, with applications throughout the machine learning, graphics, and vision communities. Without any additional structural assumptions on transport, however, OT can be fragile to outliers or noise, especially in high dimensions. Here, we introduce Latent Optimal Transport (LOT), a new approach for OT that simultaneously learns low-dimensional structure in data while leveraging this structure to solve the alignment task. The idea behind our approach is to learn two sets of "anchors" that constrain the flow of transport between a source and target distribution. In both theoretical and empirical studies, we show that LOT regularizes the rank of transport and makes it more robust to outliers and the sampling density. We show that by allowing the source and target to have different anchors, and using LOT to align the latent spaces between anchors, the resulting transport plan has better structural interpretability and highlights connections between both the individual data points and the local geometry of the datasets.

Approach: Latent Optimal Transport (LOT)

Latent Optimal Transport:
\[
\min_{P_{x}, P_{y}} \sum_{i,j} P_{x,i}C_{x}(x_{i}, y_{j}) + \sum_{i,j} P_{y,j}C_{y}(y_{j}, x_{i}) + \sum_{i,m} P_{y,m}C_{y}(y_{m}, y_{i})
\]

- Cluster the source and target points with anchors
- Transport data with anchors
- Optimize the anchors

Domain Adaptation Experiments:
A. Adapting a classifier trained on MNIST digits to classify USPS digits and corrupted MNIST digits respectively, by linearly updating the last layer.
B. LOT outperforms other methods at rectifying domain shift.
C. LOT provides an interpretable transport plan.

Conclusion & Future works/ References & Acknowledgement
LOT is a low-rank transport leveraging data structures of the source and target. LOT is theoretical grounded in the objective and has fast sampling rate.
\*We demonstrate that LOT can be effectively used in domain transfer application.
\*Future works include extension of LOT to graphical data and incorporation of metric learning.

Distribution Alignment with Optimal Transport

Optimal Transport:
\[
\min_{P} \sum_{i,j} P_{ij}C_{ij}(x_{i}, y_{j})
\]

- P: Transport plan (align data points)
- C: Transport cost (measure data dissimilarity)

Low-Rank Transport:

- Geometric Properties: LOT induces the latent Wasserstein distance \(W_{\text{LOT}}(\mu, \nu)\)
- Symmetry
- Quasi-triangle inequality

Connections to OT: LOT as a relaxation of OT if \(K_{s}K_{t}K_{s} \leq K\)

- Sampling Complexity:
- Time Complexity:

Visualized alignments with OT, LOT and factored couplings [1] with different number of anchors.