Résoudre efficacement le problème de transport optimal

Jean FEYDY, École Normale Supérieure

Mots-clés : GPU, transport optimal

Computing distances between weighted point clouds is a fundamental problem in applied mathematics, with applications ranging from Computer Vision (SLAM, reconstruction from LIDAR...) to Medical Imaging (image, mesh, fiber track registration) and Machine Learning (GANs, VAEs, geometric deep learning...). In practice, researchers strive to endow spaces of measures with distances that must:

- **lift a ground metric** defined on pairs of points as faithfully as possible;
- define **informative gradients** with respect to the points' weights and positions;
- be affordable, from a **computational** point of view.

I will first present the three families of distances that have shaped the literature since the 50's:

- **Soft-Hausdorff** (aka. Iterative Closest Points, loglikelihoods of mixture models) losses, which rely on (soft) nearest-neighbour projections.
- **Kernel** (aka. Sobolev, Electrostatic, Maximum Mean Discrepancies, blurred SSD) norms, which rely on (discrete) convolutions.
- **Optimal Transport** (OT, aka. Earth-Mover, Wasserstein, Linear Assignment) costs, which often rely on linear or convex optimizers.

Focusing on OT, which is most appealing from a geometric perspective, I will then explain how to compute Wasserstein distances *efficiently*. Leveraging recent works (2016+) on entropic regularization, we will see that **fast multiscale strategies** now allow us to compute smooth, approximate solutions of the OT problem with *strong theoretical guarantees*. As evidenced by a live demonstration of the `GeomLoss` python library:

`https://www.kernel-operations.io/geomloss/`

on modern gaming hardware, **Wasserstein distances (and gradients!) can now be computed between millions of samples in a matter of seconds**.
Available through efficient (log-linear), robust and well-packaged GPU routines, Wasserstein distances can now be used as a drop-in replacement for the cheaper Hausdorff and Kernel distances. **But how useful can this tool be in applied settings?** This talk will be followed by that of Pierre Roussillon, who will introduce applications to mesh registration and the study of white matter tractograms.

Références

[1] Kosowsky, Yuille, *The invisible hand algorithm: Solving the assignment problem with statistical physics*, Neural Networks, 1994.

[2] Peyré, Cuturi, *Computational Optimal Transport*, `optimaltransport.github.io/book`, 2018.

[3] Schmitzer, *Stabilized Sparse Scaling Algorithms for Entropy Regularized Transport Problems*, SIAM Journal on Scientific Computing, 2016.

[4] Feydy, Trouvé, *Global divergences between measures: from Hausdorff distance to Optimal Transport*, ShapeMI 2018.

[5] Feydy, Séjourné, Vialard, Amari, Trouvé, Peyré, *Interpolating between Optimal Transport and MMD using Sinkhorn Divergences*, AiStats2019.

Jean FEYDY, DMA, École Normale Supérieure, 45 rue d'Ulm, 75005, Paris
`jean.feydy@ens.fr`