Computing Wasserstein Barycenter via operator splitting: the Method of Averaged Marginals

D. Mimouni[*1,3], P. Malisani[†1], J. Zhu[‡2], and W. de Oliveira[§3]

[1]IFP Energies nouvelles, Dpt. Applied Mathematics
[2]IFP Energies nouvelles, Dpt. Control and Signal Processing
[3]Mines Paris, Center for Applied Mathematics

Keywords: Wassertein Barycenters, Optimal Transport, Douglas-Rachford Method

The Wasserstein barycenter (WB) is an important tool for summarizing sets of probabilities. It finds applications in applied probability, clustering, image processing, etc. When the probability supports are finite and fixed, the problem of computing a WB is formulated as a linear optimization problem whose dimensions generally exceed standard solvers' capabilities. For this reason, the WB problem is often replaced with a simpler nonlinear optimization model constructed via an entropic regularization function so that specialized algorithms can be employed to compute an approximate WB efficiently. Contrary to such a widespread inexact scheme, we propose an exact approach based on the Douglas-Rachford splitting method applied directly to the WB linear optimization problem for applications requiring accurate WB. Our algorithm, which has the interesting interpretation of being built upon averaging marginals, operates series of simple (and exact) projections that can be parallelized and even randomized, making it suitable for large-scale datasets. As a result, our method achieves good performance in terms of speed while still attaining accuracy. Furthermore, the same algorithm can be applied to compute generalized barycenters of sets of measures with different total masses by allowing for mass creation and destruction upon setting an additional parameter. Our contribution to the field lies in the development of an exact and efficient algorithm for computing barycenters, enabling its wider use in practical applications. The approach's mathematical properties are examined, and the method is benchmarked against the state-of-the-art methods on several data sets from the literature.

[*]daniel.mimouni@ifpen.fr
[†]paul.malisani@ifpen.fr
[‡]jiamin.zhu@ifpen.fr
[§]welington.oliveira@minesparis.psl.eu