Two approaches for population Wasserstein barycenter problem: Stochastic Averaging versus Sample Average Approximation

Darina Dvinskikh, Alexander Gasnikov
Weierstrass Institute for Applied Analysis and Stochastics, Berlin,
MIPT, IITP, HSE

Abstract

In Machine Learning and Optimization community there are two main approaches for convex risk minimization problem: Stochastic Averaging (SA) and Sample Average Approximation (SAA). At the moment, it is known that both approaches are on average equivalent (up to a logarithmic factor) in terms of oracle complexity (required number of stochastic gradient evaluations). What is the situation with total complexity? The answer depends on specific problem. However, generally SA is assumed to be better than SAA. In this paper we show that SAA may outperform SA in the problem of calculating an estimation for population (entropy regularized) Wasserstein barycenter.

Literature

A. Kroshnin, D. Dvinskikh, P. Dvurechensky, A. Gasnikov, N. Tupitsa and C. A. Uribe. On the Complexity of Approximating Wasserstein Barycenter. Proceedings of the 36th International Conference on Machine Learning, PMLR 97:3530-3540, 2019.

P. Dvurechenskii, D. Dvinskikh, A. Gasnikov, C. Uribe and A. Nedich. Decentralize and randomize: Faster algorithm for Wasserstein barycenters. InAdvances in Neural Information Processing Systems. P. 10760–10770, 2018.

P. Dvurechensky, A. Gasnikov and A. Kroshnin. Computational Optimal Transport: Complexity by Accelerated Gradient Descent Is Better Than by Sinkhorn's Algorithm. Proceedings of the 35th International Conference on Machine Learning, PMLR 80:1367–1376, 2018.

C.A. Uribe, D. Dvinskikh, P. Dvurechensky, A. Gasnikov, A. Nedić. Distributed computation of Wasserstein barycenters over networks. In 2018 IEEE Conference on Decision and Control (CDC) (pp. 6544-6549). IEEE, December, 2018.