Variational and Information Flows in Machine Learning and Optimal Transport

19.-25.11.2023

December 15, 2022

Program We intend to address parts of the following topics which fit together in the context of learning generative models.

1. Optimal transport and Wasserstein spaces

   We will introduce optimal transport and its regularized, unbalanced, and multimarginal variants. This includes the study of the geometry of Wasserstein spaces, geodesics and their Riemannian-like structure as exhibited by the Benamou–Brenier functional. This will lead us to the notion of Wasserstein gradient flows which recently attracted interest in the theory of machine learning. We will also consider related novel developments such as the Stein variational gradient descent and neural optimal transport. Last but not least, computational aspects will be addressed.

2. Generalized normalizing flows

   A unified framework to (diffusion) normalizing flows and variational autoencoders can be given via Markov chains. More precisely, stochastic normalizing flows are a pair of Markov chains fulfilling some properties and many state-of-the-art models for data generation fit into this framework and enables the coupling of both deterministic layers as invertible neural networks and stochastic layers as Metropolis-Hasting layers, Langevin layers, variational autoencoders and diffusion normalizing flows in a mathematically sound way. Applications in inverse problems in imaging will be addressed.

3. Mean field games

   Mean-field games/control study the behavior of a large number of rational agents moving in the Euclidean spaces and recently also on on Riemannian manifolds. The formulation of the mean-field game Nash equilibrium, the equivalence between the PDE system and the optimality conditions of the associated variational form will be discussed as well as the design of proximal gradient method for variational mean-field games.

References:

[1] L. Ambrosio, N. Gigli, and G. Savare. *Gradient Flows.* Lectures in Mathematics ETH Zürich. Birkhäuser, Basel, 2005.

[2] R. Barboni and G. Peyré and F-X. Vialard, Global convergence of ResNets: From finite to infinite width using linear parameterization, NeurIPS 2022.

[3] Y.F. Chen and W.C. Li. Optimal transport natural gradient in statistical manifolds with continuous sample space. Information Geometry, 2020.

[4] M. Bruveris and F-X. Vialard, On completeness of groups of diffeomorphisms, Journal of the European Mathematical Society 5(19), 1507–1544, 2017.

[5] G. Carlier, V. Duval, G. Peyré and B Schmitzer, Convergence of entropic schemes for optimal transport and gradient flows, SIAM Journal on Mathematical Analysis 49 (2), 1385-1418

[6] P. Hagemann, J. Hertrich, and G. Steidl. Generalized normalizing flows via Markov chains. in: Elements in Non-local Data Interactions: Foundations and Applications, Cambridge University Press, 2021.

[7] R. Jordan, D. Kinderlehrer, and F. Otto. The variational formulation of the Fokker–Planck equation. *SIAM Journal on Mathematical Analysis*, 29(1):1–17, 1998.

[8] G. Peyré and M. Cuturi, Computational Optimal Transport, Foundations and Trends in Machine Learning 11(5-6), 355-607, 2019

[9] L. Ruthotto, S. Osher, W.C. Li, L. Nurbekyan, S. Fung. A Machine Learning Framework for Solving High-Dimensional Mean Field Game and Mean Field Control Problems. Proceedings of the National Academy of Sciences (PNAS), 2020.

[10] M. Bruveris and F-X. Vialard, On completeness of groups of diffeomorphisms, Journal of the European Mathematical Society 5(19), 1507–1544, 2017.