University of Heidelberg
Department of Mathematics and Computer Science
Image & Pattern Analysis Group

Master-Thesis

A Stochastic Approach for Two-layer Semi-discrete Optimal Transport

Name:	Daniel Gonzalez
Enrolment number:	3112012
Supervisor:	Prof. Christoph Schnörr
Date:	20. Juli 2020

Abstract

The semi-discrete optimal transport problem arises in numerous applications as a natural way to calculate differences between two given measures. However, the computational burden to handle this problem tends to hinder practicable realizations, especially when working in high-dimensional spaces or with target measures that have large supports. In this work, we want to cope with this problem studying a stochastic optimization approach that uses the benefits of a hierarchical multi-scale strategy in order to approximate optimal transport maps. We investigate the two-layer approach as an unsupervised learning problem that seeks to find the best approximation by maximizing the two-layer expected reward. Our main goal will consist in understanding whether the numerical advantages that we might gain using this strategy compensate the possible deviation of these approximations from the optimal solution. For this, we study the underlying geometry of the two-layer power maps and use them as the class of functions in which the best possible approximation is to be found. Understanding them will give us a better intuition and insights on the optimization problem that arises when using this approach. We conclude this work by analyzing and motivating a version of the Average Stochastic Gradient Ascent algorithm which will turn out as a very efficient strategy to solve the two-layer stochastic optimization problem.

Table of Contents

Bibliography

[Aurenhammer, 1987] Aurenhammer, F. (1987). Power diagrams: properties, algorithms and applications. SIAM Journal on Computing, 16(1):78–96.

[Aurenhammer et al., 1998] Aurenhammer, F., Hoffmann, F., and Aronov, B. (1998). Minkowski-type theorems and least-squares clustering. Algorithmica, 20(1):61–76.

[Bercu and Bigot, 2020] Bercu, B. and Bigot, J. (2020). Asymptotic distribution and convergence rates of stochastic algorithms for entropic optimal transportation between probability measures. Annals of Statistics.

[Bigoni et al., 2019] Bigoni, D., Zahm, O., Spantini, A., and Marzouk, Y. (2019). Greedy inference with layers of lazy maps. arXiv preprint arXiv:1906.00031.

[Border, 2016] Border, K. (2016). Differentiating an integral: Leibniz'rule. Caltech Division of the Humanities and Social Sciences.

[Bottou et al., 2018] Bottou, L., Curtis, F. E., and Nocedal, J. (2018). Optimization methods for large-scale machine learning. Siam Review, 60(2):223–311.

[Brenier, 1991] Brenier, Y. (1991). Polar factorization and monotone rearrangement of vector-valued functions. Communications on pure and applied mathematics, 44(4):375–417.

[Chakraborty, 2015] Chakraborty, S. (2015). Generating discrete analogues of continuous probability distributions-a survey of methods and constructions. Journal of Statistical Distributions and Applications, 2(1):1–30.

[Chazal et al., 2011] Chazal, F., Cohen-Steiner, D., and Mérigot, Q. (2011). Geometric inference for probability measures. Foundations of Computational Mathematics, 11(6):733–751.

[Cuturi, 2013] Cuturi, M. (2013). Sinkhorn distances: Lightspeed computation of optimal transport. In Advances in neural information processing systems, pages 2292–2300.

[Elstrodt, 1996] Elstrodt, J. (1996). Maß-und Integrationstheorie, volume 7. Springer.

[Evans and Gariepy, 2015] Evans, L. C. and Gariepy, R. F. (2015). Measure theory and fine properties of functions. CRC press.

[Frogner et al., 2015] Frogner, C., Zhang, C., Mobahi, H., Araya, M., and Poggio, T. A. (2015). Learning with a wasserstein loss. In Advances in neural information processing systems, pages 2053–2061.

[Galerne et al., 2018] Galerne, B., Leclaire, A., and Rabin, J. (2018). A texture synthesis model based on semi-discrete optimal transport in patch space. SIAM Journal on Imaging Sciences, 11(4):2456–2493.

[Gangbo and McCann, 1996] Gangbo, W. and McCann, R. J. (1996). The geometry of optimal transportation. Technical report, SCAN-9604031.

[Garling, 2018] Garling, D. J. (2018). Analysis on Polish spaces and an introduction to optimal transportation, volume 89. Cambridge University Press.

[Genevay et al., 2016] Genevay, A., Cuturi, M., Peyré, G., and Bach, F. (2016). Stochastic optimization for large-scale optimal transport. In Advances in neural information processing systems, pages 3440–3448.

[Kantorovich, 1942] Kantorovich, L. V. (1942). On the translocation of masses. In Dokl. Akad. Nauk. USSR (NS), volume 37, pages 199–201.

[Kantorovich, 1948] Kantorovich, L. V. (1948). On a problem of monge. In CR (Doklady) Acad. Sci. URSS (NS), volume 3, pages 225–226.

[Kitagawa, 2014] Kitagawa, J. (2014). An iterative scheme for solving the optimal transportation problem. Calculus of Variations and Partial Differential Equations, 51(1-2):243–263.

[Kitagawa et al., 2019] Kitagawa, J., Mérigot, Q., and Thibert, B. (2019). Convergence of a newton algorithm for semi-discrete optimal transport. Journal of the European Mathematical Society, 21(9):2603–2651.

[Leclaire and Rabin, 2019] Leclaire, A. and Rabin, J. (2019). A fast multi-layer approximation to semi-discrete optimal transport. In International Conference on Scale Space and Variational Methods in Computer Vision, pages 341–353. Springer.

[Leclaire and Rabin, 2020] Leclaire, A. and Rabin, J. (2020). A stochastic multi-layer algorithm for semi-discrete optimal transport with applications to texture synthesis and style transfer. Journal of Mathematical Imaging and Vision, pages 1–27.

[Lévy, 2015] Lévy, B. (2015). A numerical algorithm for l2 semi-discrete optimal transport in 3d. ESAIM: Mathematical Modelling and Numerical Analysis, 49(6):1693–1715.

[Lipman and Daubechies, 2011] Lipman, Y. and Daubechies, I. (2011). Conformal wasserstein distances: Comparing surfaces in polynomial time. Advances in Mathematics, 227(3):1047–1077.

[Mérigot, 2011] Mérigot, Q. (2011). A multiscale approach to optimal transport. In Computer Graphics Forum, volume 30, pages 1583–1592. Wiley Online Library.

[Monge, 1781] Monge, G. (1781). Mémoire sur la théorie des déblais et des remblais. Histoire de l'Académie Royale des Sciences de Paris.

[Moulines and Bach, 2011] Moulines, E. and Bach, F. R. (2011). Non-asymptotic analysis of stochastic approximation algorithms for machine learning. In Advances in Neural Information Processing Systems, pages 451–459.

[Oberman and Ruan, 2015] Oberman, A. M. and Ruan, Y. (2015). An efficient linear programming method for optimal transportation. arXiv preprint arXiv:1509.03668.

[Peyré et al., 2019] Peyré, G., Cuturi, M., et al. (2019). Computational optimal transport: With applications to data science. Foundations and Trends® in Machine Learning, 11(5-6):355–607.

[Pratelli, 2007] Pratelli, A. (2007). On the equality between monge's infimum and kantorovich's minimum in optimal mass transportation. In Annales de l'Institut Henri Poincare (B) Probability and Statistics, volume 43, pages 1–13. Elsevier.

[Rabin and Papadakis, 2015] Rabin, J. and Papadakis, N. (2015). Convex color image segmentation with optimal transport distances. In International Conference on Scale Space and Variational Methods in Computer Vision, pages 256–269. Springer.

[Robbins and Monro, 1951] Robbins, H. and Monro, S. (1951). A stochastic approximation method. The annals of mathematical statistics, pages 400–407.

[Rockafellar and Wets, 2009] Rockafellar, R. T. and Wets, R. J.-B. (2009). Variational analysis, volume 317. Springer Science & Business Media.

[Rubner et al., 2000] Rubner, Y., Tomasi, C., and Guibas, L. J. (2000). The earth mover's distance as a metric for image retrieval. International journal of computer vision, 40(2):99–121.

[Santambrogio, 2015] Santambrogio, F. (2015). Optimal transport for applied mathematicians. Birkäuser, NY, 55(58-63):94.

[Villani, 2003] Villani, C. (2003). Topics in optimal transportation. Number 58. American Mathematical Soc.

[Villani, 2008] Villani, C. (2008). Optimal transport: old and new, volume 338. Springer Science & Business Media.