Research Statement

Joseph D. Walsh III

1 Background

The theory of optimal transport dates back to Monge in 1781 [15]. In the 1940s, Kantorovich's papers [9, 10] relaxed Monge's formulation, creating the Monge-Kantorovich problem:

> Let $X, Y \subset \mathbb{R}^d$, and let μ and ν be probability densities defined on X and Y, respectively. Define Π, the set of *transport plans*: probability densities in the product space $X \times Y$ whose marginals equal μ and ν. Given $c(x, y) : X \times Y \to \mathbb{R}$, a measurable *ground cost* function, find the *optimal transport cost*: $$P^* := \inf_{\pi \in \Pi(\mu, \nu)} \int_{X \times Y} c(x, y) d\pi(x, y).$$

For a fixed ground cost function c, solutions to Monge-Kantorovich problems are related to the *Wasserstein metric*, $W_1 = W_1(\mu, \nu) = P^*$, a distance between probability distributions.

Monge and Kantorovich's equations are described as solving optimal transport problems because they developed them with an eye toward fundamental economic applications: optimizing the transport of materials from one location to another. Their broader applications were recognized in the 1990s, when Brenier discovered a link between optimal transport and fluid mechanics; see [5]. Further revelations followed in papers like [7], linking optimal transport to a wide range of dynamical systems and naturally-occurring partial differential equations.

As recently as 2013, an international group of leaders in computational optimal transport outlined what they considered the field's crucial research objectives; see [12]. Specific goals included: designing numerical methods capable of handling general ground costs (described as "completely open and very challenging"), and addressing the difficulty and expense of computing Wasserstein distances. The group's 2014 and 2015 reports, available as [13, 14], and the talks presented at their July 2016 research workshop[1], which I attended, suggest that these goals remain priorities.

Most computational optimal transport focuses on continuous Monge-Kantorovich problems with bounded support, which are first made semi-discrete by discretizing (Y, ν). My research has focused on this semi-discrete transport problem, with an eye toward the goals listed above.

2 Accomplishments

2.1 A new transport approach: the boundary method

Solving the semi-discrete Monge-Kantorovich problem described above is equivalent to partitioning X into regions based on transport destination. This can be expressed as an *optimal coupling*:

> Assume the support of X is contained in the bounded convex region A, and that ν has n non-zero values, located at $\{y_i\}_{i=1}^n \subset Y$. Define
> $$F(x) = \max_n \{a_i - c(x, y_i)\} \quad \text{and} \quad A_i = \{\, x \mid F(x) = a_i - c(x, y_i) \,\}.$$
> Find constants $\{a_i\}_{i=1}^n$ such that for all i, $\mu(A_i) = \nu(y_i)$.

If the Monge-Kantorovich equation corresponding to the optimal coupling problem is solved, the resulting optimal transport plan π^* "couples" points in the set A to one or more points y_i, partitioning A a.e. into n destination regions, each of which satisfies the definition of its set A_i. For many ground cost functions, including all norms, the partition is unique μ-a.e.

On the region boundaries — one place where uniqueness always fails — the definition of the regions gives the relation:

$$c(x, y_i) - c(x, y_j) = a_i - a_j \tag{2.1}$$

[1] *Computational Optimal Transportation* at le Centre de Recherches Mathématiques in Montréal, July 18–22, 2016.

for all $x \in A_i \cap A_j$, where $i, j \in \mathbb{N}_n$. There exist $n-1$ independent equations of the form given in Equation (2.1), so if $n-1$ appropriate points x can be identified, the set $\{a_i\}_{i=1}^n$ can be computed. (The set has exactly one degree of freedom, inherent in the definition of F, so a single a_i should be fixed before computation.) Given $\{a_i\}_{i=1}^n$, the partition sets A_i may be determined, and because $\varphi = -F$, if A_i and a_i are known for $1 \leq i \leq n$, the W_1 distance can be computed.

Using this information, I developed the boundary method. The optimal transport problem is used to distinguish interior from boundary regions based on the destinations of nearest neighbors. Interior points, irrelevant to Equation (2.1), may be discarded, and applying localized refinement keeps the problem size manageable. Transport needs only be approximate, as specific destinations are unimportant; boundary points will be reevaluated on the next refinement step. However, given the size and scaling of transport problems, it is vital that the transport solver be *fast*.

2.2 A new transport solver: the general auction

A cornucopia of continuous algorithms have been developed, based on fluid mechanics [5], Monge-Ampére [2], gradient descent [6, 8], and Bregman's iteration method. Unfortunately, to satisfy well-posedness requirements, the ground cost function must either be regularized, or satisfy the MTW conditions described in [11]. (Most norms, including ℓ_2, do not satisfy the MTW conditions.) More importantly, these methods can only compute over region interiors; the singularity of the boundaries prevents their accurate approximation by continuous methods.

Linear programming methods look more promising, as any well-defined ground cost and feasible discretization can be applied. However, LP methods cannot *approximate* solutions; they are exact solvers. If the solution process is interrupted, their state can be arbitrarily far from optimal. Thus, these solvers operate at a fixed speed. This speed can be unacceptably slow.

As an alternative to the continuous and LP approaches, consider the distributed relaxation methods known as auction algorithms: members of one set, the "buyers," attempt to purchase members of the other set, the "lots," through a process resembling a real-world sales auction. Buyers bid on lots, occasionally attempting to outbid each other. Each new bid must exceed the existing price by a minimum bid *step size* ε. Successful bids link buyers to lots. With an appropriate pricing method, incorporating the ground cost function, the successful bids at the end of the auction establish a transport plan that approximates the transport solution.

Auction algorithms offer controlled approximation, because the error in the W_1 distance is bounded as a function of the step size ε. Computation time is manageable, as the speed of the method is inversely proportional to ε; see [3] for details. The auction method also accepts any well-defined ground cost and discretization. While this appears ideal, consider the limitations of the only published auction method for optimal transport, in [4]. Ground cost function values must be rationally-related, as must the combined set of masses generated by ν and the discretized μ. Other values are not supported. The Monge-Kantorovich problem must be transformed by expansion into an integer-valued assignment problem (transport with integer-valued costs and masses equal to 1). This multiplies problem size by the GCD of the costs and masses, which can be arbitrarily large.

Because of these issues, I developed a new auction method: the general auction. It is so named because it is based directly on the (more general) real-valued transport problem, rather than the integer-valued assignment problem. The general auction is guaranteed to terminate, even with real-valued masses and ground costs, and the error in the result is bounded as a function of the step size ε. In fact, the auction method for integer-valued assignment turns out to be a special case of the general auction method. Tests have shown that, for integer-valued transport problems, the general auction matches or outperforms other auction methods. For real-valued transport, the general auction is the only such method. The details of the general auction are given in [17]. I shared my source code for the general auction in an open source project, publicly available as [16].

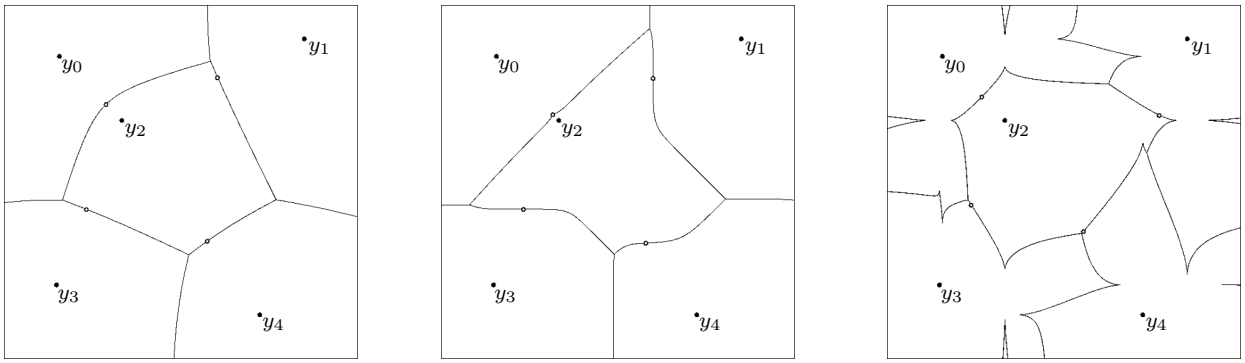
2.3 Performance for general ground costs: $c = \ell_p$ with $p = 2$, 10, ½

I also implemented the boundary method itself, in order to test real-world performance. Consider the following standard problem, taken from [1], to be solved using my software:

Example. *Let $X = Y = [0, 1]^2$, and $c = \ell_p$ for some $p \in \mathbb{R}^+ \cup \infty$. For all measurable sets $S \subseteq A$, $\mu(S) = |S|$, and ν has uniform discrete probability density, so $\nu(y_i) = 1/n$ for $1 \leq i \leq n$.*

Take $n = 5$, with the five points $\{y_i\}_{i=1}^n$ distributed as shown in Figure 1, and let the target width for X equal 2^{-13}. Fully discretizing X would generate 67 108 864 $\left(2^{26}\right)$ data points; by comparison, solving the boundary method uses fewer than 39 000 points, a ratio of 1700 to 1.

Partitions were generated for three different ℓ_p distances: ℓ_2, ℓ_{10}, and $\ell_{1/2}$. Note that $\ell_{1/2}$ is a general ground cost function; it violates the triangle inequality, so $\ell_{1/2}$ is not even a norm. Results are shown in Figure 1. It took less than one second to compute each solution on standard equipment (a six-year old laptop), including calculating W_1 distances with error bounds of magnitude 10^{-4}.



(a) ℓ_2 ground cost (b) ℓ_{10} ground cost (c) $\ell_{1/2}$ ground cost

Figure 1: Optimal coupling into regions of equal area

The locations of the boundary points used to compute $\{a_i\}_{i=1}^n$ are indicated by open circles. All points and boundaries are emphasized to aid visibility; at this scale, the boundaries are actually 0.0125 millimeters thick. As Figure 1(c) shows, even general ground cost problems can be solved. For ℓ_p costs with $p < 1$, regions may become discontinuous, as exemplified by the "spikes" on the exterior walls. (The lower spike on the right is coupled with y_3; the others are coupled with y_2.)

3 Research Plans

My future research will develop applications in the following areas:

Cluster analysis. One of the commonly used tools in machine learning is cluster analysis: in essence, grouping similar concepts based on distance with respect to some metric. Determining clusters using numerical optimal coupling offers multiple advantages over traditional clustering methods. In particular, it takes into account both the density and centroid of the set(s), without requiring *a priori* assumptions about distribution or connectivity. I will develop the boundary method into an optimal coupling-based clustering technique. Any application that relies on computerized decision making or data analysis could potentially benefit from my new cluster-based identification algorithm: examples range from entity resolution to mathematical chemistry.

Voronoi diagrams. The images in Figure 1 are generalized Voronoi diagrams. Unlike standard Voronoi diagrams, they control for area *and* minimize distance. Each of the images I shared divides areas equally, but in fact my implementation can assign any desired ratio of areas (and cost functions, as the subfigures illustrate). Already, the boundary methods's scaling behavior is comparable to the best of the standard Voronoi algorithms. I will pursue applications of generalized Voronoi construction methods to computational chemistry, computational fluid dynamics, networking, and machine learning. Because the boundary method constructs generalized Voronoi diagrams equally well in 3-dimensions, I will also explore applications in ultra-light materials design.

References

[1] John W. Barrett and Leonid Prigozhin. A mixed formulation of the Monge-Kantorovich equations. *ESAIM: M2AN*, 41:1041–1060, 2007.

[2] Jean-David Benamou, Brittany D. Froese, and Adam M. Oberman. Numerical solution of the optimal transportation problem using the Monge-Ampère equation. *Journal of Computational Physics*, 260:107–126, 2014.

[3] Dimitri P. Bertsekas. *Network Optimization: Continuous and Discrete Models*. Athena Scientific, Belmont, Massachusetts, 1998. `http://web.mit.edu/dimitrib/www/books.htm`.

[4] Dimitri P. Bertsekas and David A. Castañón. The auction algorithm for the transportation problem. *Annals of Operations Research*, 20(1):67–96, 1989.

[5] Yann Brenier. The dual least action problem for an ideal, incompressible fluid. *Archive for Rational Mechanics and Analysis*, 122(4):323–351, 1993.

[6] Rick Chartrand, Brendt Wohlberg, Kevin R. Vixie, and Erik M. Bollt. A gradient descent solution to the Monge-Kantorovich problem. *Applied Mathematical Sciences (Ruse)*, 3:1071–1080, 2009.

[7] Michael J. P. Cullen and R. James Purser. An extended Lagrangian theory of semi-geostrophic frontogenesis. *Journal of the Atmospheric Sciences*, 41:1477–1497, 1984.

[8] Richard Jordan, David Kinderlehrer, and Felix Otto. The variational formulation of the Fokker-Planck equation. *SIAM Journal on Mathematical Analysis*, 29(1):1–17, 1998.

[9] Leonid V. Kantorovich. On the translocation of masses. *C.R. (Doklady) Acad. Sci. URSS (N.S.)*, 37:199–201, 1942.

[10] Leonid V. Kantorovich. On a problem of Monge. *Uspekhi Mat. Nauk*, 3:225–226, 1948.

[11] Xi-Nan Ma, Neil S. Trudinger, and Xu-Jia Wang. Regularity of potential functions of the optimal transportation function. *Arch. Ration. Mech. Anal.*, 177(2):151–183, 2005.

[12] Mokaplan. Inria international program: Associate team proposal 2014–2016, 2013. `https://team.inria.fr/mokaplan/files/2014/09/MOKALIEN_Proposal_2013.pdf`.

[13] Mokaplan. Mokalien associate team report 2014 (1st year), 2014. `https://team.inria.fr/mokaplan/files/2014/09/MOKALIEN_Report_2014.pdf`.

[14] Mokaplan. Mokalien associate team report 2015 (2nd year), 2015. `https://team.inria.fr/mokaplan/files/2014/09/MOKALIEN_Report_2015.pdf`.

[15] Gaspard Monge. Mémoire sur la théorie des déblais et des remblais. pages 666–704, 1781.

[16] J.D. Walsh III. The AUCTION ALGORITHMS IN C++ project. Computer software, `http://gatech.jdwalsh03.com/coding.html`.

[17] J.D. Walsh III and Luca Dieci. General auction method for real-valued optimal transport. Preprint, `http://gatech.jdwalsh03.com/index.html`, 2016.