Optimal transport via a Monge-Ampére optimization problem

Michael Lindsey

(joint work with Yanir Rubinstein)

May 29, 2020

Outline

Review of optimal transport and numerical methods

Recasting the Monge-Ampére equation for OT as a (convex) optimization problem

Discretization

Convergence result

Experiments

Modified discretization

Removing target condition

Section 1

Review of optimal transport and numerical methods

Monge problem

- Monge optimal transport problem:

$$
\begin{array}{ll}
\text{minimize} & \displaystyle\int_X c(x, T(x))\, d\mu(x) \\
\text{subject to} & T \text{ measurable, } T_{\#}\mu = \nu.
\end{array}
$$

- Kantorovich is a relaxation (always feasible)
- Exist unique coinciding solutions to each problem when $c(x, y) = |x - y|^2$, $\mu = f\,dx, \nu = g\,dy$ on $X = Y = \mathbb{R}^n$

Brenier's theorem

- $X = Y = \mathbb{R}^n$. Roughly speaking, under mild technical conditions on μ, ν, Monge-Kantorovich problem is uniquely solved by transport map $T = \nabla\varphi$, where φ is convex
- Conversely, gradients of convex functions are optimal maps from μ to $T_{\#}\mu$

Review of numerical methods

- LP, assignment problems
- semi-discrete (Lévy 2015)
- entropic regularization (Cuturi 2013)
- sparse LP (Oberman-Ruan 2015, Schmitzer 2016)
- PDE methods (Benamou-Brenier 2000, Benamou-Froese-Oberman 2014)

Section 2

Recasting the Monge-Ampére equation for OT as a (convex) optimization problem

The Monge-Ampère equation

- Ω, Λ open and bounded in $\mathbb{R}^n$, consider optimal transport problem from source measure μ on Ω with density f to target measure ν on Λ with density g
- Caffarelli's regularity theory for Λ convex
- Equivalently, find convex ψ solving second boundary value problem for the Monge–Ampère equation

$$\begin{aligned} \det\left(\nabla^2\psi(x)\right) &= \frac{f(x)}{g\left(\nabla\psi(x)\right)}, \quad x \in \Omega, \\ \nabla\psi(\Omega) &= \Lambda. \end{aligned} \tag{1}$$

A cute observation

- Suppose ψ convex is a 'subsolution' of the MAE, i.e.,

$$\begin{aligned} \det\left(\nabla^2\psi(x)\right) &\geq \frac{f(x)}{g\left(\nabla\psi(x)\right)}, \quad x \in \Omega, \\ \nabla\psi(\Omega) &\subset \Lambda. \end{aligned} \tag{2}$$

- Then ψ is a solution!

A cute transformation

- Equivalently, find convex ψ such that

$$-\det{}^{1/n}\left(\nabla^2\psi(x)\right) + f^{1/n}(x)g^{-1/n}\left(\nabla\psi(x)\right) \le 0, \quad x \in \Omega,$$
$$\nabla\psi(\Omega) \subset \Lambda. \tag{3}$$

- Notice that LHS of inequality is 'convex' in ψ if $g^{-1/n}$ is convex. This holds in particular if g is log-concave (which holds in particular if g is uniform on convex Λ)

A Monge-Ampére optimization problem

- So the convex potential is the solution of:

$$\begin{aligned} \underset{\psi}{\text{minimize}} \quad & \int_\Omega \max\left\{0, -\det{}^{1/n}\left(\nabla^2\psi(x)\right) + f^{1/n}(x) g^{-1/n}\left(\nabla\psi(x)\right)\right\} \\ \text{subject to} \quad & \psi \text{ convex} \\ & \nabla\psi(\Omega) \subset \Lambda \end{aligned}$$

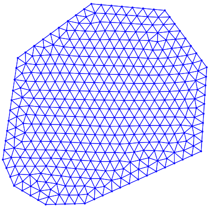
- If $g^{-1/n}$ is convex, then this can be thought of as a convex optimization problem

Section 3

Discretization

Discretization

- Many ways to discretize
- First we triangulate the domain into simplices $\{S_i\}_{i=1}^{M}$ with vertex set $\{x_j\}_{j=1}^{N}$, where $S_i = \mathrm{Conv}\{x_{i_0}, \ldots, x_{i_n}\}$
- Example:



A discrete Monge-Ampére optimization problem

The *discrete Monge–Ampère optimization problem (DMAOP)* associated to the data $(\Omega, \Lambda, f, g, \{x_j\}_{j=1}^N, \{S_i\}_{i=1}^M)$ is a natural discretization of the continuous MAOP. Optimization variables are:

- ψ_j variable for Brenier potential value at point x_j
- η_j variable for subdifferential of Brenier potential at x_j
- convexity enforced via $\psi_j \geq \psi_i + \langle \eta_i, x_j - x_i \rangle \,,\; i, j = 1, \ldots, N$
- convexity constraint inefficient, can be made more efficient in practice later

Section 4

Convergence result

What do we want to converge?

- We need to extract something to converge
- Define

$$a_j(x) := \psi_j + \langle \eta_j, x - x_j \rangle, \qquad j = 1, \ldots, N, \tag{4}$$

  so a_j is the (unique) affine function with $a_j(x_j) = \psi_j$ and $\nabla a_j(x_j) = \eta_j$.
- Define the *optimization potential* by

$$\phi(x) := b + \max_{j=1,\ldots N} a_j(x), \tag{5}$$

  where $b \in \mathbb{R}$ is chosen such that $\phi(0) = 0$.

Convergence result

Theorem (L.-Rubinstein, arXiv:1603.07435)

Let $f \in C^{0,\alpha}(\overline{\Omega})$, $g \in C^{0,\alpha}(\overline{\Lambda})$ *and suppose* Λ *is convex. Let* $\{\{S_i^{(k)}\}_{i=1}^{M(k)}\}_{k\in\mathbb{N}}$ *be a 'nice' sequence of triangulations. Let* φ *be the the unique Brenier solution of our Monge–Ampère equation with* $\varphi(0) = 0$*, and let* $\phi^{(k)}$ *be the optimization potentials obtained from the DMAOP. Then*

$$\phi^{(k)} \to \varphi \text{ uniformly on } \overline{\Omega},$$

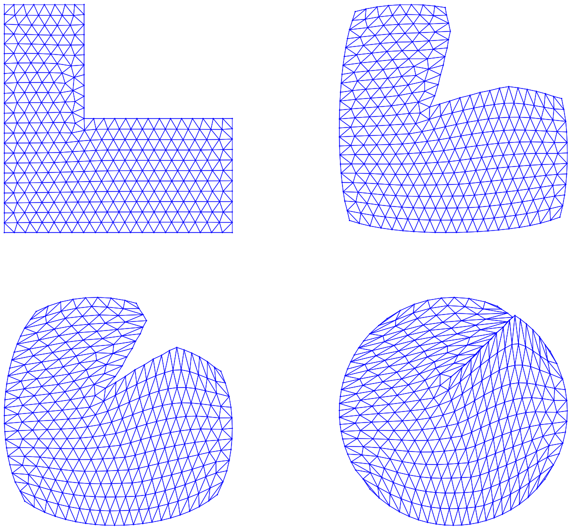
and $\partial\phi^{(k)} \to \nabla\varphi$ *pointwise on* $\overline{\Omega}$*. In particular,* $\nabla\phi^{(k)}$ *converges pointwise almost everywhere to the optimal transport map pushing forward* $\mu = f\,dx$ *to* $\nu = g\,dx$*.*

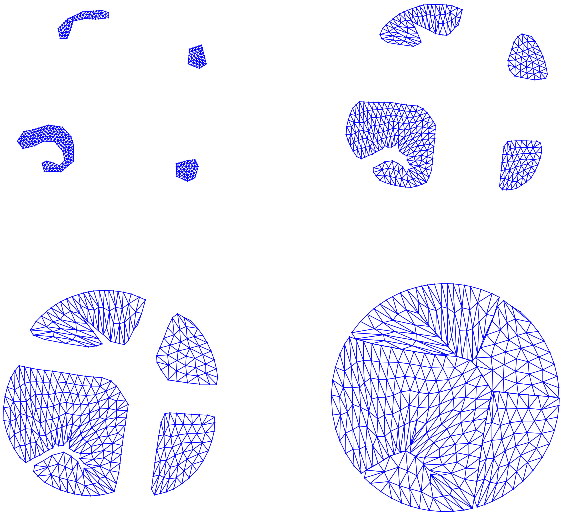
Main ideas of proof

- Show that optimal costs $c_k \to 0$.
- Assume $\phi^{(k)} \to \phi$ uniformly for some ϕ. WTS $\phi = \varphi$.
- Mollify the sequence to get $\phi_\varepsilon^{(k)} \to \phi_\varepsilon$ $(\star)$.
- Use $c_k \to 0$ together with the convergence $(\star)$ to establish that ϕ_ε does not 'excessively contract' mass, up to some error that is $o(1)$ in ε.
- Roughly speaking, define $\nu_\varepsilon := (\nabla\phi_\varepsilon)\#\mu$, use previous point to show that $\nu \geq \nu_\varepsilon + o(1)$ as $\varepsilon \to 0$.
- But $\nu(\Lambda) = \nu_\varepsilon(\Lambda)$, so $\nu_\varepsilon \to \nu$.
- Conclude by stability of optimal transport.

Section 5

Experiments



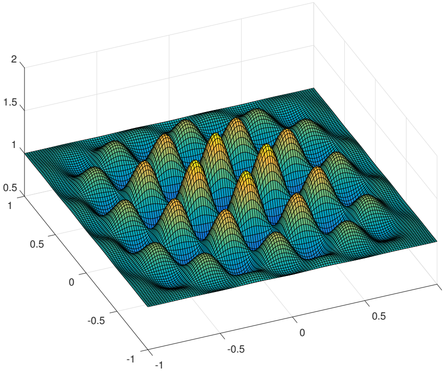


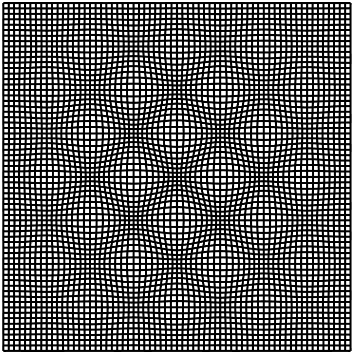
Section 6

Modified discretization

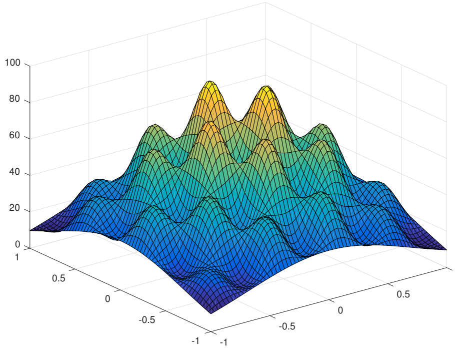
Recall we want to solve

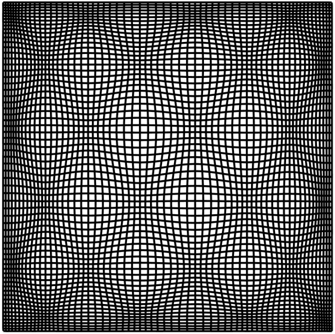
$$\begin{aligned} \underset{\psi}{\text{minimize}} \quad & \int_{\Omega} \max\left\{0, -\det{}^{1/n}\left(\nabla^2\psi(x)\right) + f^{1/n}(x) g^{-1/n}\left(\nabla\psi(x)\right)\right\} \\ \text{subject to} \quad & \psi \text{ convex} \\ & \nabla\psi(\Omega) \subset \Lambda \end{aligned}$$


2
1.5
1
0.5
1
0.5
0
-0.5
-1
-1
-0.5
0
0.5



h	L^∞ error	L^2 error	Time (s)
2^{-3}	0.0648	0.0249	1.53
2^{-4}	0.0045	0.0020	1.38
2^{-5}	0.0021	0.0010	2.15
2^{-6}	0.0007	0.0003	8.20
2^{-7}	0.0005	0.0002	38.58


100
80
60
40
20
0
1
0.5
0
-0.5
-1
-1
-0.5
0
0.5



h	L^∞ error	L^2 error	Time (s)
2^{-3}	0.0375	0.0179	1.27
2^{-4}	0.0119	0.0049	1.51
2^{-5}	0.0034	0.0012	1.90
2^{-6}	0.0013	0.0007	9.51
2^{-7}	–	–	42.31

Section 7

Removing target condition

- Recall that we want to solve

$$\det\left(\nabla^2\varphi(x)\right) = \frac{f(x)}{g\left(\nabla\varphi(x)\right)}, \qquad x \in \Omega, \qquad (6)$$
$$\nabla\varphi(\Omega) = \Lambda.$$

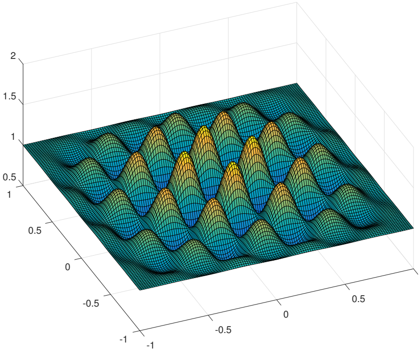
- Let $V = \mathrm{vol}(\Lambda)$, $f^{(0)} = f$, and let $\varphi^{(0)}$ be the solution of

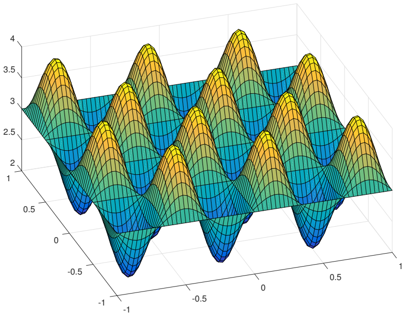
$$\det\left(\nabla^2\varphi(x)\right) = V f^{(0)}(x), \qquad x \in \Omega,$$
$$\nabla\varphi(\Omega) = \Lambda.$$

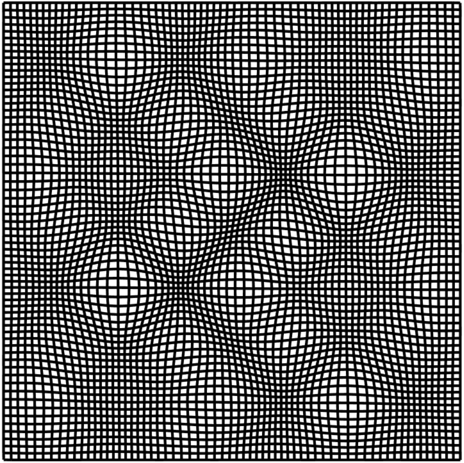
- Define $\varphi^{(i+1)}$ to be the solution of

$$\det\left(\nabla^2\varphi(x)\right) = V f^{(i+1)}(x) \qquad x \in \Omega,$$
$$\nabla\varphi(\Omega) = \Lambda,$$

where $f^{(i+1)} := \tilde{f}^{(i+1)} \Big/ \left(\int_\Omega \tilde{f}^{(i+1)}\, dx\right)$ for $\tilde{f}^{(i+1)} := f/(g \circ \nabla\varphi^{(i)})$.

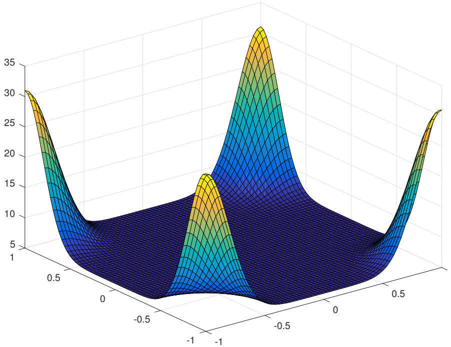

2
1.5
1
0.5
1
0.5
0
-0.5
-1
-1
-0.5
0
0.5

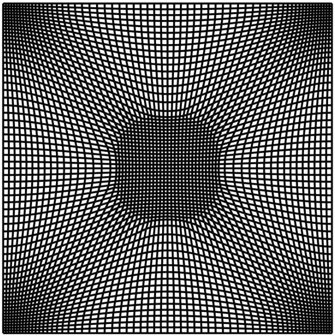

4
3.5
3
2.5
2
1
0.5
0
-0.5
-1
-1
-0.5
0
0.5
1



h	L^∞ error	L^2 error	Iterations	Time (s)
2^{-3}	0.0814	0.0317	3	3.36
2^{-4}	0.0175	0.0095	4	5.59
2^{-5}	0.0056	0.0031	4	8.97
2^{-6}	0.0017	0.0008	5	45.16
2^{-7}	–	–	5	207.02

Source density: $f = 1.5 - \mathbf{1}_{B(0,0.5)}$.





h	L^∞ error	L^2 error	Iterations	Time (s)
2^{-3}	0.0312	0.0192	6	7.42
2^{-4}	0.0156	0.0069	7	10.53
2^{-5}	0.0091	0.0041	7	14.11
2^{-6}	0.0045	0.0025	9	78.90
2^{-7}	–	–	10	381.73