Computational methods for optimal mass transport

Eldad Haber
with: Tauseef Rehman and Allen Tannenbaum

Overview

- Overview and Motivation
- Two local methods
- Obtaining a global method
- Discretization
- Solution of linear systems
- Preliminary results
- Outlook

Overview

The Monge-Kantorovich mass transfer: how to transfer a "pile of dirt" from one place to another with minimum energy

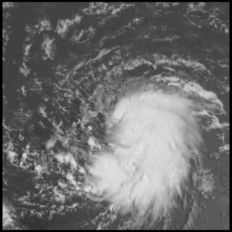
Overview

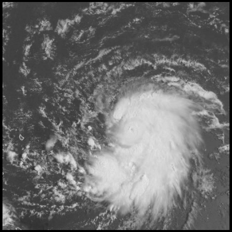
The Monge-Kantorovich mass transfer: how to transfer a "pile of dirt" from one place to another with minimum energy

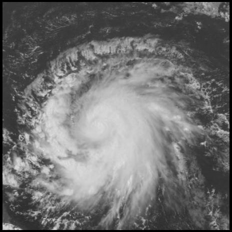
- Difficult problem with many local minima
- Applications in geoscience, image registration ...
- Analyzed by many see survey by Evans
- Very little numerical treatment Benamou and Brenier, Angenent et-al, Oberman, Oliker,

Optimal mass transport

Find a vector field u that moves a mass $\mu_0(x)$ into $\mu_1(x)$ in an optimal way



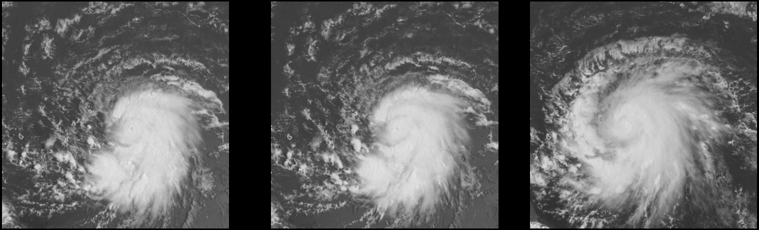




$$
\begin{aligned}
\min \; M(u) &= \int_\Omega |u - x|^2 \mu_0(x)\, dx = \|u - x\|^2_{\mu_0} \\
\text{s.t } c(u) &= \det(\nabla u)\mu_1(u) - \mu_0(x) = 0 \\
& \quad \mu_0, \mu_1 > 0
\end{aligned}
$$

Optimal mass transport

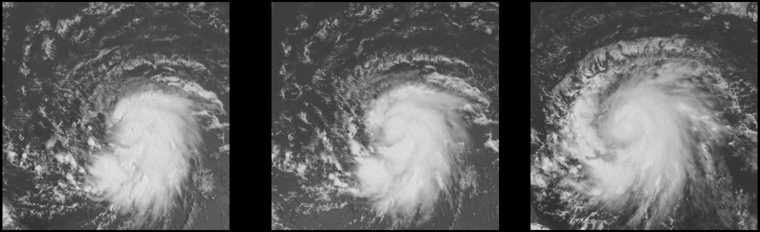
Find a vector field u that moves a mass $\mu_0(x)$ into $\mu_1(x)$ in an optimal way



$$
\begin{aligned}
\min \;\; M(u) &= \int_\Omega |u - x|^2 \mu_0(x)\, dx = \|u - x\|^2_{\mu_0} \\
\text{s.t } c(u) &= \det(\nabla u)\mu_1(u) - \mu_0(x) = 0 \\
&\quad \mu_0, \mu_1 > 0
\end{aligned}
$$

Optimal mass transport

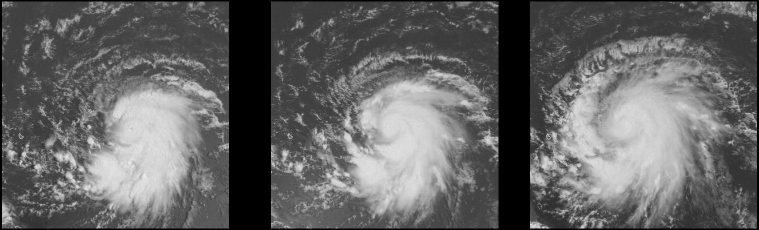
Find a vector field u that moves a mass $\mu_0(x)$ into $\mu_1(x)$ in an optimal way



$$
\begin{aligned}
\min \; M(u) &= \int_\Omega |u-x|^2 \mu_0(x)\, dx = \|u-x\|^2_{\mu_0} \\
\text{s.t } c(u) &= \det(\nabla u)\mu_1(u) - \mu_0(x) = 0 \\
& \mu_0, \mu_1 > 0
\end{aligned}
$$

Optimal mass transport

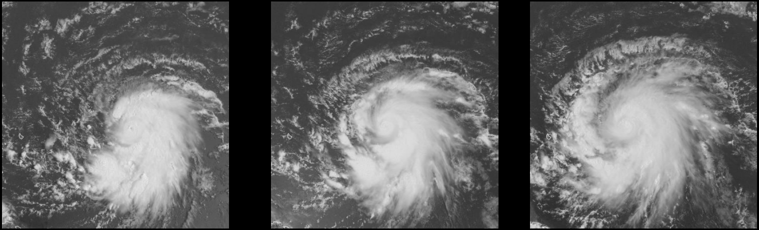
Find a vector field u that moves a mass $\mu_0(x)$ into $\mu_1(x)$ in an optimal way



$$
\begin{aligned}
\min \; M(u) &= \int_\Omega |u-x|^2 \mu_0(x)\, dx = \|u-x\|^2_{\mu_0} \\
\text{s.t } c(u) &= \det(\nabla u)\mu_1(u) - \mu_0(x) = 0 \\
& \quad \mu_0, \mu_1 > 0
\end{aligned}
$$

Optimal mass transport

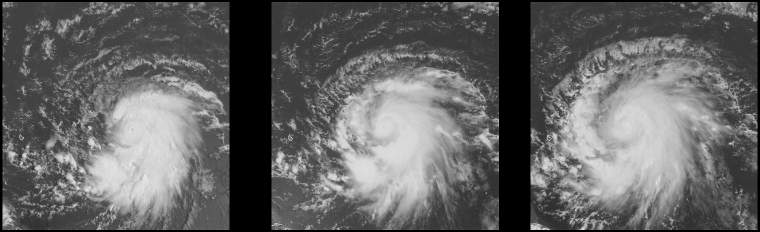
Find a vector field u that moves a mass $\mu_0(x)$ into $\mu_1(x)$ in an optimal way



$$\begin{aligned} \min \ M(u) &= \int_\Omega |u-x|^2 \mu_0(x)\,dx = \|u-x\|^2_{\mu_0} \\ \text{s.t } c(u) &= \det(\nabla u)\mu_1(u) - \mu_0(x) = 0 \\ & \mu_0, \mu_1 > 0 \end{aligned}$$

Optimal mass transport

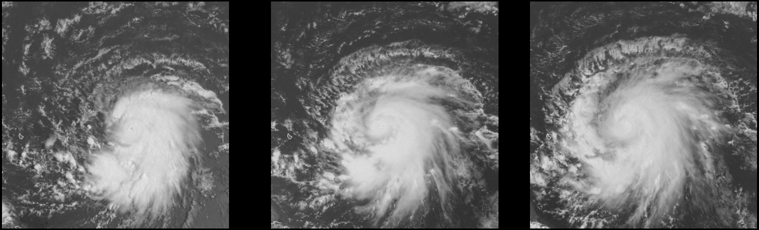
Find a vector field u that moves a mass $\mu_0(x)$ into $\mu_1(x)$ in an optimal way



$$
\begin{aligned}
\min\ M(u) &= \int_\Omega |u-x|^2 \mu_0(x)\, dx = \|u-x\|^2_{\mu_0} \\
\text{s.t } c(u) &= \det(\nabla u)\mu_1(u) - \mu_0(x) = 0 \\
&\quad \mu_0, \mu_1 > 0
\end{aligned}
$$

Optimal mass transport

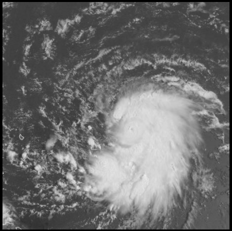
Find a vector field u that moves a mass $\mu_0(x)$ into $\mu_1(x)$ in an optimal way

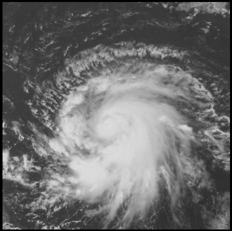


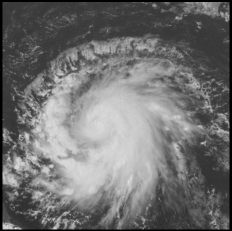
$$
\begin{aligned}
\min \; M(u) &= \int_\Omega |u - x|^2 \mu_0(x)\, dx = \|u - x\|^2_{\mu_0} \\
\text{s.t } c(u) &= \det(\nabla u)\mu_1(u) - \mu_0(x) = 0 \\
&\quad \mu_0, \mu_1 > 0
\end{aligned}
$$

Optimal mass transport

Find a vector field u that moves a mass $\mu_0(x)$ into $\mu_1(x)$ in an optimal way







$$\begin{aligned} \min \; M(u) &= \int_\Omega |u-x|^2 \mu_0(x)\, dx = \|u-x\|^2_{\mu_0} \\ \text{s.t } c(u) &= \det(\nabla u)\mu_1(u) - \mu_0(x) = 0 \\ & \mu_0, \mu_1 > 0 \end{aligned}$$

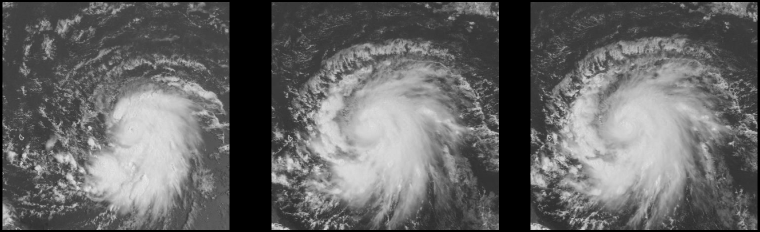
Optimal mass transport

Find a vector field u that moves a mass $\mu_0(x)$ into $\mu_1(x)$ in an optimal way

$$\begin{aligned} \min\ M(u) &= \int_\Omega |u-x|^2 \mu_0(x)\, dx = \|u-x\|^2_{\mu_0} \\ \text{s.t } c(u) &= \det(\nabla u)\mu_1(u) - \mu_0(x) = 0 \\ &\quad \mu_0, \mu_1 > 0 \end{aligned}$$

Optimal mass transport

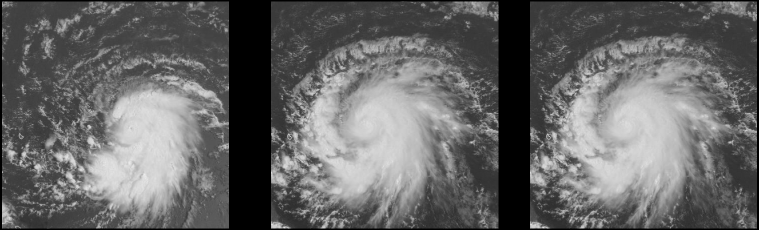
Find a vector field u that moves a mass $\mu_0(x)$ into $\mu_1(x)$ in an optimal way



$$\begin{aligned} \min\ M(u) &= \int_\Omega |u - x|^2 \mu_0(x)\, dx = \|u - x\|^2_{\mu_0} \\ \text{s.t } c(u) &= \det(\nabla u)\mu_1(u) - \mu_0(x) = 0 \\ & \mu_0, \mu_1 > 0 \end{aligned}$$

Optimal mass transport

Find a vector field u that moves a mass $\mu_0(x)$ into $\mu_1(x)$ in an optimal way



$$
\begin{aligned}
\min\ M(u) &= \int_\Omega |u-x|^2 \mu_0(x)\, dx = \|u-x\|^2_{\mu_0} \\
\text{s.t } c(u) &= \det(\nabla u)\mu_1(u) - \mu_0(x) = 0 \\
& \mu_0, \mu_1 > 0
\end{aligned}
$$

Difficulties

Difficulties

- Highly nonlinear constraint
- May have local minima
- Consistent discretization is difficult
- Large scale applications

Difficulties

Difficulties

- Highly nonlinear constraint
- May have local minima
- Consistent discretization is difficult
- Large scale applications

A short and not inclusive list:

- Reformulation Kantorovich 48
- Solving using a Monge-Ampre formulation Oliker 94
- Control problem in space time Benamou and Brenier 00
- Polar factorization Angenent Haker & Tannenbaum 04

More about the problem

$$
\begin{aligned}
\min\ M(u) &= \|u - x\|_{\mu_0}^2 \\
\text{s.t } c(u) &= \det(\nabla u)\mu_1(u) - \mu_0(x) = 0
\end{aligned}
$$

Can be shown that

- Smooth
- Possible local solutions
- Global min is curl free $\nabla \times u = 0 \quad \Leftrightarrow \quad u = \nabla \phi$

The obvious option - Sequential Quadratic Programming

$$\begin{aligned}\min\ M(u) &= \|u-x\|^2_{\mu_0}\\ \text{s.t}\ c(u) &= \det(\nabla u)\mu_1(u)-\mu_0(x)=0\end{aligned}$$

- Can be solved using SQP
- May converge to a local minima (not curl free)

More about SQP

$$
\begin{aligned}
\min \; M(u) &= \|u - x\|^2_{\mu_0} \\
\text{s.t } c(u) &= \det(\nabla u)\mu_1(u) - \mu_0(x) = 0
\end{aligned}
$$

- Have been developed in the 80's
- Commonly used for many non-convex problems
- Successfully used for problems evolving from PDE's
- Nontrivial globalization mechanism

More about SQP

Closely related to Newton's method on the Lagrangian

$$\mathcal{L}\ M(u) = \|u - x\|^2_{\mu_0} + \int_\Omega p(\det(\nabla u)\mu_1(u) - \mu_0(x))\, dx$$

At each iteration approximately solve

$$\begin{aligned} \min \quad & (\delta u, A(u,p)\delta u) + (\delta u, g(u,p)) \\ \text{s.t} \quad & B\delta u + c(u) = 0 \end{aligned}$$

More about SQP

$$\begin{aligned} \min\ M(u) &= \|u - x\|^2_{\mu_0} \\ \text{s.t } c(u) &= \det(\nabla u)\mu_1(u) - \mu_0(x) = 0 \end{aligned}$$

- Smooth
- Mesh independence properties
- Need to deal with KKT systems
- Preconditioners necessary

Modified Objective function

Use the properties of the solution to get better global attraction

$$\begin{aligned} \min\ M(u) &= \|u-x\|^2_{\mu_0} + \alpha \int_\Omega |\nabla \times\ u|^2\,dx \\ \text{s.t } c(u) &= \det(\nabla u)\mu_1(u) - \mu_0(x) = 0 \end{aligned}$$

- Does not change the minima, but bias towards the global one
- May converge to a local minima (not curl free)

The method of Angenent Haker and Tannenbaum (AHT)

Use the properties of the solution to obtain a different problem

- Find an initial MP map u_0 such that

$$\det(\nabla u_0)\mu_1(u_0) = \mu_0(x)$$

- Set $u(s) = u_0$ and solve

$$\begin{aligned} \min_s \ M(s) &= \|u_0(s^{-1}) - x\|^2_{\mu_0} \\ \text{s.t } c(s) &= \det(\nabla s)\mu_0(s) = \mu_0(x) \end{aligned}$$

The method of AHT

$$\begin{aligned} \min_s \ M(s) &= \|u_0(s^{-1}) - x\|^2_{\mu_0} \\ \text{s.t } c(s) &= \det(\nabla s)\mu_0(s) = \mu_0(x) \end{aligned}$$

Assuming the constraint is feasible linearize to obtain

$$\nabla \cdot \ \mu_0 \delta s(s^{-1}) = 0$$

$$\delta u = -(\nabla u)\delta s(s^{-1})$$

The method of AHT

$$\begin{aligned} \min_s \; M(s) &= \|u_0(s^{-1}) - x\|^2_{\mu_0} \\ \text{s.t } c(s) &= \det(\nabla s)\mu_0(s) = \mu_0(x) \end{aligned}$$

Assuming the constraint is feasible linearize to obtain

$$\nabla \cdot \; \mu_0 \delta s(s^{-1}) = 0$$

$$\delta u = -(\nabla u)\delta s(s^{-1})$$

Therefore

$$\delta u = -\mu_0^{-1}(\nabla u)\delta\zeta$$

$$\nabla \cdot \; \delta\zeta = 0$$

The method of AHT

$$\delta u = -\mu_0^{-1}(\nabla u)\delta\zeta$$
$$\nabla \cdot \delta\zeta = 0$$

First variation in the functional yields

$$\delta M = (\delta\zeta, u) + \mathrm{h.o.t}$$

AHT: Choose $\delta\zeta$ such that it is div-free and minimize $(\delta\zeta, u)$

Can be done by the Helmhotz decomposition of u

The method of AHT

Helmhotz decomposition of u

$$\begin{aligned} u &= \delta\zeta + \nabla\xi \\ 0 &= \nabla \cdot \delta\zeta \end{aligned}$$

$$\delta\zeta = (I - \nabla\Delta^{-1}\nabla\cdot\,)u$$

therefore

$$u_t = -\mu_0^{-1}(\nabla u)(I - \nabla\Delta^{-1}\nabla\cdot\,)u$$

The method of AHT

Helmhotz decomposition of u

$$\begin{aligned} u &= \delta\zeta + \nabla\xi \\ 0 &= \nabla \cdot \delta\zeta \end{aligned}$$

$$\delta\zeta = (I - \nabla\Delta^{-1}\nabla\cdot\,)u$$

therefore

$$u_t = -\mu_0^{-1}(\nabla u)(I - \nabla\Delta^{-1}\nabla\cdot\,)u$$

Generalization of AHT

- We have $\nabla \cdot \delta\zeta = 0 \quad \Leftrightarrow \quad \delta\zeta = \nabla \times \delta\eta$
- and therefore $\delta M = (u, \nabla \times \delta\eta) = (\nabla \times u, \delta\eta)$
- Steepest descent direction $\delta\eta = -\nabla \times u$
- But any direction $\delta\eta = -\mathcal{A}\nabla \times u$ with $\mathcal{A}$ SPD works

Lead to

$$\mu_0(\nabla u)^{-1}u_t = -\nabla \times \mathcal{A}\nabla \times u$$
$$u(0) = u_0$$

Generalization of AHT

Use artificial time to obtain the IVP

$$\mu_0(\nabla u)^{-1}u_t = -\nabla \times \Delta^{-1}\nabla \times u$$
$$u(0) = u_0$$

Theorem (AHT): The flow converges to the global minimizer of the problem

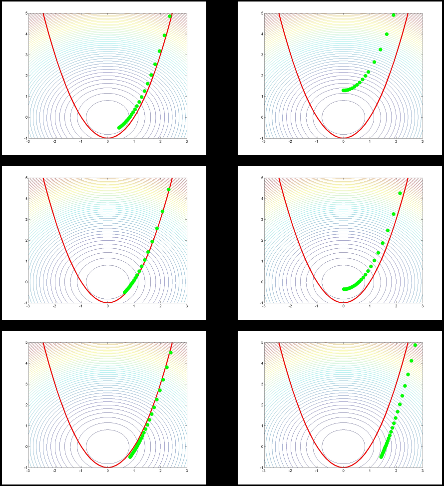
Is the AHT method global?

method $h = 1/100$ $h = 1/10$

RK4

BDF3

trap



For infinitesimal small steps yes - For any finite step no!

From 2 local methods to a global method

- The AHT method is a PDE with an invariant
- Well known that without projection cannot converge
- Examples - harmonic oscillator does not preserve energy after discretization, div free for Navier Stokes
- Need to include the constraint directly in the computation

From 2 local methods to a global method

Combine with SQP - local convergence to the constraint

From 2 local methods to a global method

Combine with SQP - local convergence to the constraint

Algorithm

- Initialize

$$\min \ \|u - x\|_{\mu_0}^2 + \alpha\|\nabla \times u\|^2 \quad \text{s.t } c(u) = 0$$

- If needed
  - Update

$$\widehat{u}_{k+1} = u_k - \mu_0^{-1}(\nabla u)\nabla \times \Delta^{-1}\nabla \times u$$

  - Project using SQP

$$\begin{aligned} \min \quad & \|u_{k+1} - \widehat{u}_{k+1}\|_{\mu_0}^2 \\ \text{s.t} \quad & \det(\nabla u_{k+1})\mu_1(u_{k+1}) = \mu_0(x) \end{aligned}$$

Discretization

$$
\begin{aligned}
\min \ M(u) &= \|u-x\|^2_{\mu_0} \\
\text{s.t } c(u) &= \det(\nabla u)\mu_1(u) - \mu_0(x) = 0
\end{aligned}
$$

Two options

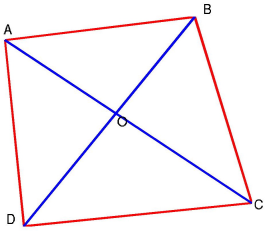
- Discretize and optimize
- Optimize and discretize

Do not commute! I prefer the first whenever possible

Consistent discretization

Discretization of the constraint, use finite-volume approach (H & Modersitzki 04)

$$\int_\Omega \det(\nabla u)\mu_1(u)\ dx = \int_{\Omega(u)} \mu_1(x)\, dx$$



Use staggered grid for stable discretization of differential operators

Solution of linear systems

At each iteration solve

- Vector Laplacian (use MG)
- A KKT system of the form

$$\begin{pmatrix} M(\mu) & C_u^\top \\ C_u & 0 \end{pmatrix}$$

- Solve the reduced system

$$A = C_u M(\mu)^{-1} C_u^\top$$

- Possible to show that A corresponds to a second order elliptic operator

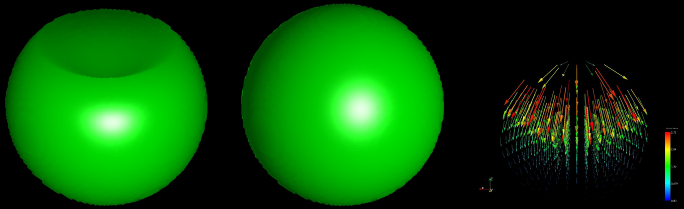
Further challenges

- High contrast $\det(\nabla u)\mu_1(u) = \mu_0(x)$

$$
\begin{aligned}
&\mu_0(\nabla u)^{-1}u_t = -\nabla \times \Delta^{-1}\nabla \times u\\
&u(0) = u_0
\end{aligned}
$$

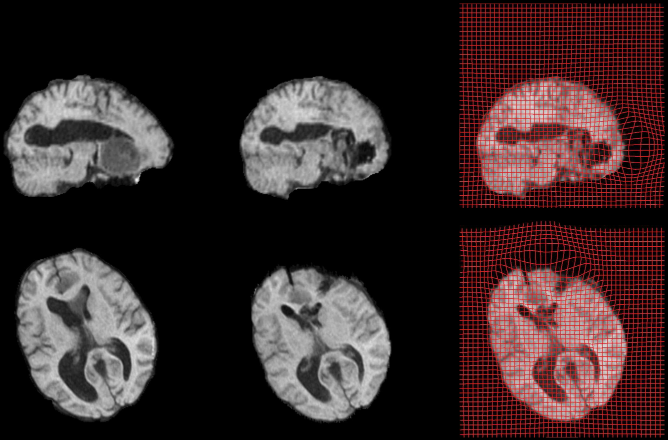
- Multigrid for the KKT system
- Adaptive mesh refinement

Results



- Grid size 128^3
- Problem solved in the first stage

Preliminary results



- Grid size 128^3
- Problem not solved in the first stage
- 2 iterations of AHT needed

Conclusion and Future work

- Developed a new method for the OMT problem
- Initial MP map by stabilized SQP
- Projection - globalize the problem
- Conservative discretization

Conclusion and Future work

- Developed a new method for the OMT problem
- Initial MP map by stabilized SQP
- Projection - globalize the problem
- Conservative discretization

- In most cases, initial map is the solution to the problem
- A few correction steps needed if not