Adaptive Approximation of the Monge-Kantorovich Problem

Sören Bartels

Department of Applied Mathematics
University of Freiburg, Germany

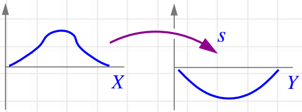
Joint work with **Patrick Schön** (University of Freiburg)

Mafelap 2016, Brunel, June 14-17

Optimal mass transfer

Goal: Move pile of soil into hole with minimal work/cost

- μ^+, μ^- measures on X, Y
- $c : X \times Y \to \mathbb{R}$ cost function



Monge ~1780: Find bijection $s : X \to Y$ rearranging μ^+ into μ^-,

$$s_\# \mu^+ = \mu^- \quad \iff \quad \int_X h \circ s \, \mathrm{d}\mu^+ = \int_Y h \, \mathrm{d}\mu^- \quad \forall h \in C(Y),$$

such that total cost I is minimal

$$I[s] = \int_X c(x, s(x)) \, \mathrm{d}\mu^+(x).$$

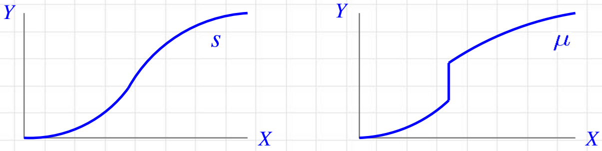
If $\mu^\pm = f^\pm \, \mathrm{d}x$ then s solves Monge–Ampère equation

$$f^+ = (f^- \circ s) \det(Ds).$$

Relaxation

Problem: Nonlinear constraint but no compactness

Kantorovich '42: Consider measure μ on $X \times Y$ instead of s



transport map s — transport plan μ

Relaxation: Find optimal transport plan $\mu \in \mathcal{M}(X \times Y)$ for

$$J[\mu] = \int_{X \times Y} c(x, y)\,\mathrm{d}\mu(x, y)$$

subject to $\mathrm{proj}_X \mu = \mu^+$ and $\mathrm{proj}_Y \mu = \mu^-$.

- Existence via direct method; consistent relaxation
- Discretization high-dimensional, need reduction

Duality I

Discretization leads to large linear program:

$$\text{Min } J_h[\mu_{ij}] = \sum_{i,j} \gamma_{ij} c_{ij} \mu_{ij} \quad \text{s.t. } \sum_j \mu_{ij} = \mu_i^+, \ \sum_i \mu_{ij} = \mu_j^-, \\ \mu_{ij} \geq 0, \ c_{ij} = c(x_i, y_j)$$

Imposing constraints via multipliers u_i and v_j leads to dual:

$$\text{Max } K_h[u_i, v_j] = \sum_i u_i \mu_i^+ + \sum_j v_j \mu_j^- \quad \text{s.t. } u_i + v_j \leq c_{ij}$$

Passing to limit $h \to 0$ gives rise to continuous dual:

$$\text{Max } K[u, v] = \int_X u \, \mathrm{d}\mu^+ + \int_Y v \, \mathrm{d}\mu^- \quad \text{s.t. } u(x) + v(y) \leq c(x, y).$$

Caffarelli: u, v price labels for shipping costs

Linear cost

For $X = Y = \Omega$, cost $c(x, y) = |x - y|$, densities $\mu^\pm = f^\pm \,\mathrm{d}x$:

$$\text{Max } K[u, v] = \int_\Omega u\, f^+ \,\mathrm{d}x + \int_\Omega v\, f^- \,\mathrm{d}y$$
$$\text{s.t. } u(x) + v(y) \le |x - y|.$$

Observation: May increase u, v so that $u(x) + v(y) = |x - y|$

- for $x = y$ follows that $v = -u$ and $u \ge 0$
- 1-Lipschitz continuity $u(x) - u(y) \le |x - y|$

With $f = f^+ - f^-$ problem reduced to ∞-Laplace problem:

$$\boxed{\text{Max } K[u] = \int_\Omega f\, u \,\mathrm{d}x \quad \text{s.t. } |\nabla u(x)| \le 1}$$

- EL equations $-\operatorname{div}(a \nabla u) = f$ with nonlinear multiplier $a(u)$

Duality II

Impose $|\nabla u| \leq 1$ via multiplier p, i.e.,

$$\text{Max}_u \, \text{Min}_p \; L[u,p] = \int_\Omega f\, u \, \mathrm{d}x + \int_\Omega |p| - p \cdot \nabla u \, \mathrm{d}x$$

Optimality:

$$0 \in \partial_u L(u,p) = f + \operatorname{div} p$$

$$0 \in \partial_p L(u,p) = \frac{p}{|p|} - \nabla u$$

Exchange min/max, eliminate u to obtain $1'$-Laplace problem:

$$\boxed{\text{Min} \;\; D[p] = \int_\Omega |p| \, \mathrm{d}x \quad \text{s.t.} \; -\operatorname{div} p = f, \; p \cdot n = 0}$$

Note strong duality relation:

$$\text{Max } K[u] = \text{Min } D[p]$$

Some references

Analysis:

- Caffarelli ’90, ’91, ’96
- Brenier ’87, ’91, ’93
- McCann ’95
- Jordan, Kinderlehrer & Otto ’98
- Evans & Gangbo ’99
- de Pascale & Pratelli ’04
- Villani ’09

Numerics:

- Barrett & Prigozhin ’09
- Oberman ’11
- Benamou & Carlier ’14
- Papadakis, Peyre, Oudet ’14

Applications: Image processing, data analysis, PDE

Discretization

$P1$ discretization of primal problem:

$$\text{Max } K[u_h] = \int_\Omega f\, u_h \,\mathrm{d}x \quad \text{s.t. } |\nabla u_h| \le 1$$

- existence straightforward
- impose vanishing mean on u_h
- no uniform convexity, error analysis?

$H(\mathrm{div};\Omega)$ conforming discretization of dual requires $f_h \approx f$

$$\text{Min } D[p_h] = \int_\Omega |p_h| \,\mathrm{d}x \quad \text{s.t. } -\operatorname{div} p_h = f_h,\ p_h \cdot n = 0$$

- discrete existence due to surjectivity of div
- nondifferentiability of objective

Primal-dual gap

For optimal u and admissible p_h and u_h have

$$
\begin{aligned}
K[u] - K[u_h] \leq D[p_h] - K[u_h] &= \int_\Omega |p_h| \,\mathrm{d}x - \int_\Omega f\, u_h \,\mathrm{d}x \\
&= \int_\Omega |p_h| \,\mathrm{d}x + \int_\Omega \operatorname{div} p_h\, u_h \,\mathrm{d}x \\
&= \int_\Omega |p_h| - p_h \cdot \nabla u_h \,\mathrm{d}x \\
&= \sum_{T \in \mathcal{T}_h} \int_T |p_h| - p_h \cdot \nabla u_h \,\mathrm{d}x = \sum_{T \in \mathcal{T}_h} \eta_T(u_h, p_h)
\end{aligned}
$$

- a posteriori error estimate with indicators $\eta_T \geq 0$
- But: no benefit from adapitvity with $\mathcal{R}T_0$ and $P1$!

A priori estimate

Noting $K[u_h] \geq K[v_h]$ for all v_h with $|\nabla v_h| \leq 1$ we have

$$K[u] - K[u_h] \leq K[u] - K[v_h] = \int_\Omega f\,(u - v_h)\,\mathrm{d}x.$$

Choose $v_h = \mathcal{I}_h u$ and note $\|\nabla \mathcal{I}_h u\|_{L^\infty} \leq \|\nabla u\|_{L^\infty}$ for $\ell = 0, 1$,

$$K[u] - K[u_h] \leq \|f\|_{L^1} \|u - u_h\|_{L^\infty} \leq c_{\mathcal{I}_h} \|f\|_{L^1} h^{1+\ell} \|D^{1+\ell} u\|_{L^\infty}.$$

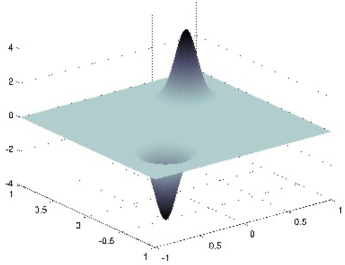
For dual problem only linear with $\mathcal{R}T_0$; more generally for $\mathcal{R}T_k$:

$$D[p_h] - D[p] = \int_\Omega |p_h| - |p|\,\mathrm{d}x \leq \|p - p_h\|_{L^1} = O(h^{k+1})$$

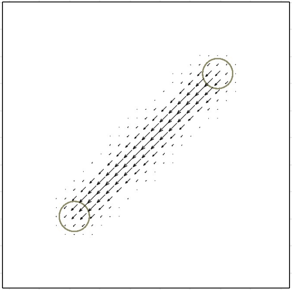
– provided sufficiently regular solution p exists

Experiment I

- Approximate f by continuous $P1$
- Discretize primal with continuous $P1$ functions
- Discretize dual with $\mathcal{R}T_1$ vector fields



$f = f^+ - f^-$

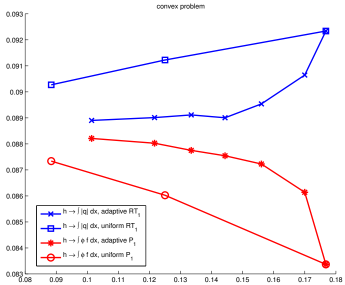


p_h

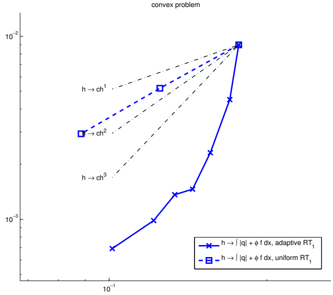
Experiment I (cont'd)

- Experimental convergence behaviour: convex case

Primal-dual gap



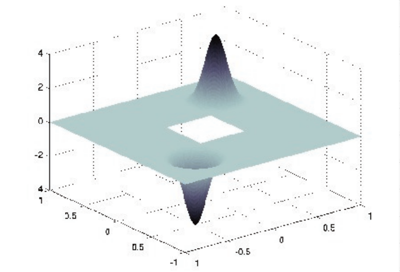
Gap estimator



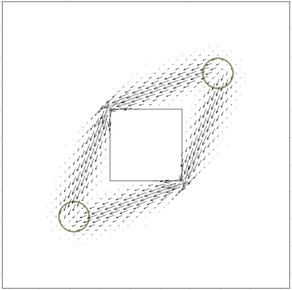
- Regularity $u \in W^{2,\infty}(\Omega)$ for ∞-Laplace on convex Ω ?

Experiment II

- Approximate f by continuous $P1$
- Discretize primal with continuous $P1$ functions
- Discretize dual with $\mathcal{R}T_1$ vector fields



$f = f^+ - f^-$

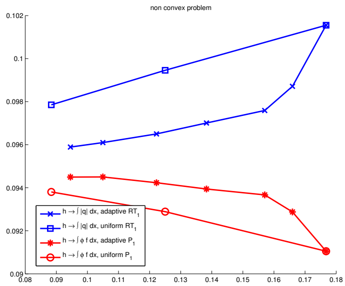


p_h

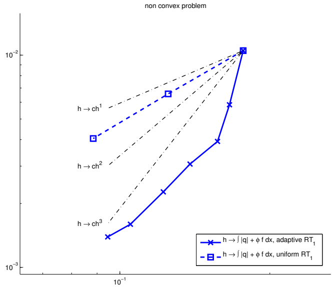
Experiment II (cont'd)

- Experimental convergence behaviour: nonconvex case

Primal-dual gap



Gap estimator



- Optimal choice of parameters?

Iterative solution: Proximity operators

Strongly convex minimization problem:

$$x \mapsto \tau|x| + \frac{|x-a|^2}{2}$$

Unique minimizer $x \in \mathbb{R}^n$ satisfies:

$$-\frac{1}{\tau}(x-a) \in \partial|x| = \begin{cases} x/|x| & |x| \neq 0, \\ K_1(0) & |x| = 0. \end{cases}$$

Implies that x is parallel to a and

$$x = (1 + \tau\partial|\cdot|)^{-1}(a) = \big(|a| - \tau\big)_+ \frac{a}{|a|}.$$

Similarly:

$$x = \operatorname{argmin} I_{K_1(0)}(x) + \frac{|x-a|^2}{2} \quad \Longleftrightarrow \quad x = \frac{a}{\max\{1, |a|\}}.$$

Iterative solution: Primal

Primal problem has standard structure

$$\text{Max } F(\nabla u) + G(u), \quad F(r) = -I_{K_1(0)}(r), \; G(u) = -\int_\Omega f\, u \,\mathrm{d}x$$

Decouple nonlinearity from gradient via splitting:

$$\text{Max}_{u,r}\, \text{Min}_\sigma \;\; L(u,r;\sigma) = F(r) + G(u) + (\sigma, \nabla u - r) - \frac{\tau}{2}\|\nabla u - r\|^2$$

Alg-2 (Fortin & Glowinski '83): Choose (u^0, r^0, σ^0) and $\tau > 0$.

(1) Max. in u, i.e., $u^k = \mathrm{argmax}\, L(u, r^{k-1}; \sigma^{k-1})$
(2) Max. in r, i.e., $r^k = \mathrm{argmax}\, L(u^k, r; \sigma^{k-1})$
(3) Update σ via descent, i.e., $\sigma^k = \sigma^{k-1} - \tau\delta_\sigma L(u^k, r^k; \sigma^{k-1})$

- Maximization in r solved explicitly pointwise
- Global convergence for every $\tau > 0$ but good choice?

Iterative solution: Dual

Dual problem nonsmooth and given by

$$\text{Min } -F^*(p) - G^*(\operatorname{div} p), \quad F^*(p) = -\int_\Omega |p| \, \mathrm{d}x, \; G^*(u) = -I_{-f}(u)$$

Decouple nondifferentiability from divergence via splitting:

$$\text{Min}_{p,s} \, \text{Max}_\lambda \; M(p, s; \lambda) = \int_\Omega |s| + I_{-f}(\operatorname{div} p) \\ + (\lambda, p - s)_h + \frac{\tau}{2} \|p - s\|_h^2$$

Apply splitting:

- Minimization in s solved pointwise
- Minimization in p is weighted dual mixed Poisson
- Update of λ via ascent step
- Need discrete norm $\| \cdot \|_h$ since $p_h \in L^1$ only
- No direct connection to primal

Dual: $\mathcal{R}T$ projection

Problem: Pointwise character in s not inherited to $\mathcal{R}T$ spaces!

$$\mathrm{Min}_{p_h,s_h}\mathrm{Max}_{\lambda_h} \int_\Omega |s_h| \,\mathrm{d}x + I_{-f_h}(\operatorname{div} p_h) + (\lambda_h, p_h - s_h)_h + \frac{\tau}{2}\|p_h - s_h\|_h^2$$

- Use Raviart–Thomas $\mathcal{R}T_k$ for p
- Use discontinuous nodal FE space $\widehat{\mathcal{S}}^k$ for s and λ

Adjustment 1: With L^2 projection $\widehat{\Pi}_h^k : L^2(\Omega)^d \to [\widehat{\mathcal{S}}^k]^d$ consider

$$\int_\Omega \widehat{\mathcal{I}}_h |s_h| \,\mathrm{d}x + I_{-f_h}(\operatorname{div} p_h) + (\lambda_h, \widehat{\Pi}_h^k p_h - s_h)_h + \frac{\tau}{2}\|\widehat{\Pi}_h^k p_h - s_h\|_h^2$$

Minimization in p_h is stable variant of dual mixed Poisson:

$$\frac{1}{2}\int_\Omega |\widehat{\Pi}_h p_h|^2 \,\mathrm{d}x + L_h(p_h), \qquad -\operatorname{div} p_h = f_h,\ p_h \cdot n = 0$$

Dual: $\mathcal{RT}$ lumping

Adjustment 2: Nodewise minimization in s_h after quadrature

$$s_h \mapsto \int_\Omega \widehat{\mathcal{I}}_h |s_h| \,\mathrm{d}x + M_h(s_h) = \sum_{T\in\mathcal{T}_h} \sum_{z\in\mathcal{N}_h\cap T} \beta_z^T |s_h(z)| + M(s_h)$$

Quadrature reduces convergence for dual Poisson ($k \geq 1$)

$$\frac{1}{2}\int_\Omega \widehat{\mathcal{I}}_h |\widehat{\Pi}_h p_h|^2 \,\mathrm{d}x + L_h(p_h), \quad -\operatorname{div} p_h = f_h,\ p_h \cdot n = 0$$

but no reduction for linear growth OT problem

- Implementation: $\mathcal{BDFM}_1$

Summary

- Optimal transport leads to ∞-Laplace problem
- A posteriori error control for duality gap
- Nearly optimal convergence rates via adaptivity
- Iterative solution of nonlinear problems via splitting

- Future aspects
  - Optimal choice of step sizes
  - Generalizations: obstacles, other costs

- More information

  `http://aam.uni-freiburg.de/bartels`