UNIVERSITÉ PARIS SUD

BV estimates in optimal transportation with applications

Département de Mathématiques d'Orsay

Alpár Richárd Mészáros
Laboratoire de Mathématiques d'Orsay
Université Paris-Sud, Orsay, France
alpar.meszaros@math.u-psud.fr
(joint work with G. De Philippis, F. Santambrogio and B. Velichkov)

Introduction

Optimal transportation and Wasserstein spaces

The problem of Monge

- OT begins with the following problem of G. Monge[a] back in 1781: given two probability densities $f, g \geq 0$ on $\mathbb{R}^d$ ($\int f = \int g = 1$), find a map $T : \mathbb{R}^d \to \mathbb{R}^d$, which "transports" (pushes forward) f onto g and minimizes the transportation cost

$$M(T) := \int_{\mathbb{R}^d} \frac{1}{2}|x - T(x)|^2 f(x)dx.$$

Kantorovich's approach by relaxation

- Kantorovich solved the existence question[b] in the full generality in 1942 by a relaxation method.
- For μ, ν probability measures on X and Y (here X and Y are $\mathbb{R}^d$, compact subsets of $\mathbb{R}^d$ or more general Polish spaces) we consider

$$(PK) \quad \inf\left\{\int_{X\times Y} \frac{1}{2}|x-y|^2 d\gamma(x,y) : \gamma \in \Pi(\mu,\nu)\right\},$$

the set $\Pi(\mu,\nu) := \{\gamma \in \mathcal{P}(X \times Y) : (\pi^x)_\# \gamma = \mu, (\pi^y)_\# \gamma = \nu\}$ is called transport plans, where π^x and π^y are the two projections of $X \times Y$ onto X and Y respectively.
- If the optimal transport plan is of the form $\gamma = (id \times T)_\# \mu$ for a measurable map $T : X \to Y$, then T is an optimal transport map.
- The dual problem is:

$$(PD) \ \sup\left\{\int_X \phi d\mu + \int_Y \psi d\nu : \ \phi(x) + \psi(y) \leq \frac{1}{2}|x-y|^2\right\}$$

- ϕ and ψ are called Kantorovich potentials.

Brenier's theorem

- If $X = Y = \mathbb{R}^d$ or $X = Y = \Omega \subset \mathbb{R}^d$ is a compact set and $\mu \ll \mathcal{L}^d$, then there exists a unique optimal transport map, which is the gradient of a convex function. It is linked to the Kantorovich potential via

$$T(x) = x - \nabla\phi(x) = \nabla\left(\frac{1}{2}|x|^2 - \phi(x)\right).$$

Wasserstein spaces and gradient flows

- For an $\Omega \subset \mathbb{R}^d$ compact set we can equip the space of probability measures $\mathcal{P}(\Omega)$ with a metric, called Wasserstein metric:

$$W_2(\mu,\nu) := \inf\left\{\int_{\Omega\times\Omega} \frac{1}{2}|x-y|^2 d\gamma : \ \gamma \in \Pi(\mu,\nu)\right\}^{\frac{1}{2}}.$$

- Gradient flows: the ODE

$$\begin{cases} x'(t) = -\nabla F(x(t)), \\ x(0) = x_0. \end{cases}$$

- Generalization to metric spaces via a discrete implicit Euler scheme: for a time step $\tau > 0$ consider x_0 and

$$x_{k+1} := argmin_x F(x) + \frac{1}{2\tau} d(x, x_k)^2.$$

[a] G. Monge, *Mémoire sur la théorie des déblais et des remblais, Hist. de l'Acad. Roy. des Sci. de Paris*, 666-704, (1781)
[b] L. Kantorovich, On the transfer of masses, *Dokl. Acad. Nauk. USSR*, (37), 7-8, 1942.

Main results

Assumptions

- Let $\Omega \subset \mathbb{R}^d$ be a convex, compact subset.
- Let $H : \Omega \to \mathbb{R}$ be a smooth convex function, such that $H(z) = H(|z|)$ (hence its gradient preserves the direction of the vectors).
- The statement of our main result is the following:

Main theorem

Theorem 1 *Let $\rho, g \in \mathcal{P}(\Omega)$ sufficiently smooth probability densities, which are away from 0 and infinity. Then we have the following inequality*

$$-\int_\Omega \nabla\rho \cdot \nabla H(\nabla\varphi) \leq \int_\Omega \nabla g \cdot \nabla H(\nabla\psi), \tag{1}$$

where φ and ψ are the Kantorovich potentials (for the quadratic $\frac{1}{2}|x-y|^2$ cost) from ρ to g and from g to ρ respectively.

Further versions

Remark 1 *This theorem remains true if we work in the whole $\mathbb{R}^d$ (instead of Ω), but with compactly supported densities g and ρ. In this case we also can drop the assumption on H, that $H(z) = H(|z|)$, hence $\nabla H(-z) \neq -\nabla H(z)$,*

$$-\int_{\mathbb{R}^d} \nabla\rho \cdot \nabla H(\nabla\varphi) \leq -\int_{\mathbb{R}^d} \nabla g \cdot \nabla H(-\nabla\psi).$$

Remark 2 *Another interesting example is the quadratic case for H, i.e. $H(z) = \frac{1}{2}|z|^2$. In this particular case our inequality has the form of*

$$-\int_{\mathbb{R}^d} \nabla\rho \cdot \nabla\varphi \leq \int_{\mathbb{R}^d} \nabla g \cdot \nabla\psi.$$

The proof of this inequality is immediate using the geodesic convexity of the entropy functional. Indeed it is well-known that the entropy functional $\mathcal{E}(\rho) : \mathcal{P}_2(\mathbb{R}^d) \ni \rho \mapsto \int_{\mathbb{R}^d} \rho \log \rho$ if $\rho \ll \mathcal{L}^d$ and $+\infty$ otherwise is convex along any (absolutely continuous) geodesics in the Wasserstein space. In particular if we consider the geodesic ρ_t connecting ρ and g and we calculate $\frac{d}{dt}\mathcal{E}(\rho_t)$ which is increasing, we obtain the desired result.

- One more interesting particular case is for $H(z) = |z|$ (obtaining it by an approximation $H_\varepsilon(z) := \sqrt{|z|^2 + \varepsilon}$)

Theorem 2 *Let $\rho, g \in \mathcal{P}_2(\mathbb{R}^d)$ sufficiently smooth, compactly supported probability densities, which are away from 0 and infinity. Then we have the following inequality*

$$-\int_{\mathbb{R}^d} \nabla\rho \cdot \frac{\nabla\varphi}{|\nabla\varphi|} \leq \int_{\mathbb{R}^d} \nabla g \cdot \frac{\nabla\psi}{|\nabla\psi|}, \tag{2}$$

where φ and ψ are the Kantorovich potentials (for the quadratic $\frac{1}{2}|x-y|^2$ cost) from ρ to g and from g to ρ respectively.

Remark 3 *By approximation arguments both Theorem 1 and Theorem 2 will remain true for $W^{1,1}$ densities.*

Applications

BV estimates for projection and obstacle-like problems

- Consider $K_1 := \{\rho \in \mathcal{P}_2(\mathbb{R}^d) : \rho \leq 1 \quad a.e\}$, or if we work on a compact set $\Omega \subset \mathbb{R}^d$, this set is the same as $\{\rho \in \mathcal{P}(\Omega) : \rho \leq 1 \ a.e\}$.
- The projection operator of a density $g \in \mathcal{P}(\Omega)$ is defined as

$$P_{K_1}[g] := argmin_{\rho \in K_1} \frac{1}{2} W_2^2(g, \rho).$$

- We have the following estimate:

Theorem 3 *If $g \in BV(\Omega)$ is a probability density, the total variation (in Ω) of its projection is less than its own total variation (in Ω), i.e.*

$$TV\left(P_{K_1}[g], \Omega\right) \leq TV(g, \Omega).$$

- We study a more general problem, the projection below a given positive BV function $f : \mathbb{R}^d \to \mathbb{R}$, which could represent an obstacle.
- For this we consider the set $K_f := \{\rho \in \mathcal{P}(\Omega) : \rho \leq f \quad a.e\}$ and the problem

$$\min\left\{\frac{1}{2}W_2^2(\rho, g) : \ \rho \in K_f\right\}. \tag{3}$$

Theorem 4 *If $f, g \in BV(\Omega)$ we have the following estimation for the solution $\rho \in \mathcal{P}(\Omega)$ of* (3)

$$TV(\rho, \Omega) \leq TV(g, \Omega) + 2TV(f, \Omega). \tag{4}$$

Remark 4 *The constant 2 in the previous inequality is sharp.*

BV estimates for the porous medium equation

- Let us consider the problem

$$\begin{cases} \partial_t \rho_t = \Delta\left(\rho_t^m\right), \text{ in } (0,T] \times \mathbb{R}^d, \\ \rho(0,\cdot) = \rho_0, \quad \text{ in } \mathbb{R}^d, \end{cases} \tag{5}$$

where ρ_0 is a non-negative BV probability density and $m > 0$ is fixed.
- Since the seminal work of F. Otto [a] we know that the problem (5) can be seen as a gradient flow of the functional $\mathcal{F}(\rho) := \frac{1}{m-1}\int_{\mathbb{R}^d} \rho^m$ in the space $(\mathcal{P}(\mathbb{R}^d), W_2)$.
- As a corollary of Theorem 2 we obtain the estimate $\forall t, s \in [0,T], t \geq s$:

$$TV(\rho_t) \leq TV(\rho_s),$$

and in particular for any $t \in [0,T]$

$$TV(\rho_t) \leq TV(\rho_0).$$

Remark 5 *For $m = 1$, i.e. for the heat equation considering the gradient flow of the entropy functional, $\mathcal{F}(\rho) := \mathcal{E}(\rho)$ we get the same estimates.*

[a] F. Otto, The geometry of dissipative evolution equations: the porous medium equation, *Commun. in PDE*, 26 (2001), No. 1-2, 101-174.

Set evolution problems

Let us consider the following problem:

- For a set $A \subset \mathbb{R}^d$ of finite perimeter with $|A| = 1$, we define $\rho_0 = I_A$, the uniform density on the set A, which gives a probability measure.
- For a time interval $[0,T]$ and a time step $\tau > 0$ (and $N + 1 := \left[\frac{T}{\tau}\right]$) we consider the following scheme $\rho_0^\tau := \rho_0$ and

$$\rho_{k+1}^\tau := P_{K_1}\left[(1+\tau)\rho_k^\tau\right], \ k \in \{0, \ldots, N-1\}, \tag{6}$$

and want to study the convergence of this algorithm as $\tau \to 0$.
- Theorem 2 will ensure that after each step of the algorithm (6) we will get a BV density and moreover the total variation of the new density decreases.
- This allows us to pass to the limit as $\tau \to 0$ and we have a strong L^1 convergence, which tells us that in the limit we will have precisely indicator functions of sets of finite perimeter, hence during the evolution of a set A we always have sets.

Crowd movements with congestion

- Similar estimates are very important to get compactness in some second order macroscopic crowd motion models with density constraints (similar to the one studied in [a]).
- The analysed Fokker-Planck type equation for a given ρ_0 BV initial density and u_t smooth enough velocity field is the following,

$$\partial_t \rho_t - \Delta\rho_t + \nabla \cdot \left(P_{adm(\rho_t)}[u_t]\rho_t\right) = 0,$$

where $P_{adm(\rho)}$ is a projection operator onto the set of admissible vector field w.r.t. ρ, the ones with positive divergence on the saturated set $\{\rho = 1\}$.
- In particular a key ingredient is an estimate of the form $W_p(\mu_0, \mu_t) \leq Ct$ when μ_t is a solution of the Fokker-Planck equation without projection. This requires $\mu_0 \in BV$, and our BV estimate ensures that the projection does not worsen the BV behaviour.
- This is the subject of a future work (see [b])

[a] B. Maury, A. Roudneff-Chupin, F. Santambrogio, A macroscopic crowd motion model of gradient flow type, *Math. Models and Meth. in Appl. Sci.*, **20** (2010), No. 10, 1787-1821.
[b] A. R. Mészáros, F. Santambrogio, A second order model for macroscopic crowd movements with congestion, *in preparation*.

Spring School "Microscopic descriptions and mean-field equations in physics and social sciences"
University of Bath, United Kingdom, May 12-16 2014

This research was supported by the European Union and the State of Hungary, co-financed by the European Social Fund in the framework of TÁMOP-4.2.4.A/ 2-11/1-2012-0001 National Excellence Program.