Estimates on transport maps onto log-concave measures.

Max Fathi
Joint work with Maria Colombo

LJLL & LPSM, Université de Paris

1er juin 2021

The optimal transport problem

Mathematical formulation : minimize $\int c(x,y)d\pi$ over all couplings π with given marginals.

Brenier-McCann theorem

Quadratic cost $c(x, y) = \frac{1}{2}|x - y|^2$ over $\mathbb{R}^d \times \mathbb{R}^d$.
If $\mu << \text{Leb}$, then the optimal coupling is deterministic : all mass at some point x is sent to $T(x)$ for some function T.
Moreover, T is the gradient of a convex function.

Monge-Ampère equation

From a change of variable, we can see that if $\mu = e^{-W}dx$ and $\nu = e^{-V}dx$, then $T = \nabla\varphi$ solves (in a weak sense) the Monge-Ampère PDE

$$e^{-W(x)} = e^{-V(\nabla\varphi(x))} \det \nabla^2\varphi(x).$$

Hence we can view the quadratic optimal transport problem as a way of describing (convex) solutions of such PDE.

The Caffarelli contraction theorem

In 2001, Caffarelli proved that if μ is the Gaussian measure $(2\pi)^{-d/2}\exp(-|x|^2/2)$ and ν is of the form $e^{-V}dx$ with convex support and $\nabla^2 V \geq \text{Id}$, then $T = \nabla\varphi$ is 1-lipschitz.
The original proof relies on the Monge-Ampère formulation and the maximum principle.
Other proof based on OT with Gozlan and Prod'homme.

Application 1 : Poincaré inequalities

The Gaussian measure satsifies the Poincaré inequality

$$\int f^2 d\gamma - \left(\int f d\gamma\right)^2 \leq \int |\nabla f|^2 d\gamma.$$

Many applications : concentration of measure, rates of convergence for diffusion processes, spectral theory of Markov processes,...
If ν is the image of the Gaussian by a 1-lipschitz map, then for a centered function f

$$\int f^2 d\nu = \int f^2 \circ T d\gamma \leq \int |\nabla T|_{op}^2 |\nabla f|^2 \circ T d\gamma \leq \int |\nabla f|^2 d\nu$$

so ν also satisfies a Poincaré inequality.
E. Milman : integrated estimates on ∇T when target is log-concave are enough to recover bounds on Poincaré constant.

Application 2 : Isoperimetry

The Gaussian measure satisfies a dimension-free isoperimetric principle (Tsirelson-Sudakov, Borell) : among all sets A with $\gamma(A)$ fixed, sets with smallest possible boundary measure

$$\gamma^+(\partial A) = \liminf \epsilon^{-1}(\gamma(A^\epsilon) - \gamma(A))$$

are half-spaces.
More explicitly, we have the lower bound

$$\gamma^+(\partial A) \geq \varphi(\phi^{-1}(\gamma(A)))$$

with

$$\varphi(t) = \frac{e^{-t^2/2}}{\sqrt{2\pi}}; \quad \phi(t) = \int_{-\infty}^{t} \varphi(s) ds.$$

Other proofs by Bobkov, Bakry-Ledoux, Barthe-Maurey.

This last inequality is stable by transfer by 1-lipschitz maps. Hence, we can use the Caffarelli contraction theorem to prove the Bakry-Ledoux isoperimetric inequality : all measures $\nu = \exp(-V)$ with $\nabla^2 V \geq \mathrm{Id}$ satisfy

$$\nu^+(\partial A) \geq \varphi(\phi^{-1}(\nu(A)))$$

for all sets A.
Particular instance of a more general theory on metric-measure spaces, for which this argument is not known to apply, but the original semigroup proof of Bakry & Ledoux does.

Sketch of proof

From the Monge-Ampère PDE, we have

$$\frac{1}{2}|x|^2 + C = V(\nabla\varpi(x)) - \log\det\nabla^2\varphi(x).$$

We can differentiate twice to get equations relating third and fourth-order derivatives of φ.
If the function

$$(x, e) \in \mathbb{R}^d \times \mathbb{S}^{d-1} \longrightarrow \partial^2_{ee}\varphi(x)$$

reaches a maximum at some point (x_0, e_0), we know the gradient is zero and the Hessian is non-positive.
Combining the equations leads to a formal a priori estimate

A priori, φ is not smooth enough to make this argument work. So we can consider instead

$$(x, e) \longrightarrow t^{-2}(\varphi(x + te) + \varphi(x - te) - 2\varphi(x))$$

and seek an estimate that is uniform for small enough t.
To ensure the function reaches a maximum, start with compactly-supported target distributions, and approximate.

Question today : what happens if V is convex, but without uniform bounds ?
Cannot expect a globally lipschitz estimate, since that's not the case for the exponential measure. But is that the worst possible case ?

In dimension one, with Courtade and Pananjady, we showed that if V is convex and the target measure centered and isotorpic, then

$$\varphi'(x) \leq c\sqrt{1+|x|^2}$$

for some universal numerical constant c.
Proof : the transport map from a measure with density ρ onto the two-sided exponential measure satsifies

$$T'(x) = \frac{2\rho(x)}{1-\rho(]-\infty, x])}$$

and we can use the Cheeger isoperimetric inequality, that compares the density and the cumulative distribution function. We conclude using the explicit map transporting the Gaussian measure onto the exponential.
What happens in higher dimension ?

We can use the fact that the transport map is the gradient of a convex funciton to localise a region where the mass is sent, relative to a reference point.

Isotropic log-concave measures satisfy concentration inequalities. Here we used the inequality for 1-lipschitz functions

$$\mathbb{P}(f \geq \mathbb{E}[f] + cr) \leq \exp(-cr^2/(r + \sqrt{d}))$$

with a universal, dimension-free constant c, which is due to Lee and Vempala (2018).

From this argument we get the estimate

$$|\nabla\varphi(x)| \leq C(d + |x|^2)$$

for any transport map from the Gaussian onto an isotropic, log-concave measure.
The order of magnitude is sharp, since this is what happens when the target measure is uniform on an ℓ^1 ball. But the worst region is quite small.
If we assume the target measure satisfies a Gaussian concentration inequality, we get instead

$$|\nabla\varphi(x)| \leq C\sqrt{d + |x|^2}$$

with C only depending on the concentration constant.

Using this estimate, we can recycle Caffarelli's argument, allowing for non-constant bonds on $\nabla^2 V$. For example, we get

Theorem

If e^{-V} is a log-concave isotropic probaility measure and

$$c_1 \geq \nabla^2 V(x) \geq \frac{c_2}{d + |x|}$$

then

$$||\nabla^2 \varphi(x)||_{op} \leq C \max\left(1, \frac{c_1^2}{c_2^2}\right)(d + |x|^2)^2.$$

To get this estimate, we seek to bound the penalized second-order difference

$$\frac{\varphi(x+te)+\varphi(x-te)-2\varphi(x)}{(d+|x|^2)^2}$$

Of course, can work with another denominator, which would require other assumptions on V.
When differentiating twice the Monge-Ampère PDE, the denominator and $\nabla^2 V$ should compensate.

We can also get more satisfactory (but hopefully still not optimal) L^p estimates by adapting an argument of Kolesnikov :

Theorem

Under the same assumptions as above

$$\left\| \frac{\partial^2_{ee}\varphi}{\sqrt{d+|x|^2}} \right\|_{L^p(\gamma)} \leq \frac{C}{c_2}\left(1 + p\frac{\sqrt{c_1}}{4\sqrt{d}}\right).$$

If we consider a log-concave product measure, tensorizing the one-dimensional estimate gives

$$||\nabla^2 \varphi||_{op} \leq c(1 + ||x||_\infty)$$

which suggests we should not work with the ℓ^2 norm. However the ℓ^∞ is not rotationally invariant...
Maybe should start with unconditional measures ?

Thanks!