Error estimates for invariant measures of diffusion processes, and stability for a Monge-Ampère type PDE.

Max Fathi
joint work with Dan Mikulincer (Weizmann Institute)

9 octobre 2020

Broad goal : given two ergodic diffusion processes of the form $(1 = 1, 2)$

$$dX_t^i = b_i(X_t^i)dt + \sqrt{2\sigma_i(X_t^i)}dB_t$$

with $b_i : \mathbb{R}^d \longrightarrow \mathbb{R}^d$ and $\sigma^i : \mathbb{R}^d \longrightarrow \mathcal{S}_d^{++}$, if the coefficients are close, are the invariant measures close ?

Many ways of tackling this (coupling, entropy estimates, stochastic calculus...)

From the Ito formula, evolution of the distribution is governed by

$$\frac{d}{dt}\mathbb{E}[f(X_t)] = \mathbb{E}[b(X_t) \cdot \nabla f(X_t) + \mathrm{Tr}(\sigma(X_t)\nabla^2 f(X_t))].$$

Can also be formulated in terms of a Fokker-Planck PDE.

Invariant measures are characterized as solutions of second-order linear PDE

$$L^*\mu = 0; Lf = b \cdot \nabla f + \mathrm{Tr}(\sigma \nabla^2 f).$$

Actual motivation

Given a probability density ρ with barycenter at the origin, we can look at (convex) solutions to the Monge-Ampère PDE

$$e^{-\varphi} = \rho(\nabla\varphi)\det(\nabla^2\varphi).$$

A solution to this PDE is a transport map sending the measure with density $e^{-\varphi}$ onto ρ.

Question : if two solutions are close, are the original data ρ_1, ρ_2 close ? And in what sense ?

Why are the two questions related ?

Up to issues about regularity, for any smooth f, for $\mu = \rho dx$,

$$\begin{aligned}\int x \cdot \nabla f(x) d\mu &= \int \nabla \varphi \cdot \nabla f(\nabla \varphi) e^{-\varphi} dx \\ &= \int \mathrm{Tr}(\nabla^2 \varphi \nabla^2 f(\nabla \varphi)) e^{-\varphi} dx \\ &= \int \mathrm{Tr}((\nabla^2 \varphi^*)^{-1} \nabla^2 f) d\mu\end{aligned}$$

where φ^* is the Legendre transform f φ, so that $\nabla \varphi^*$ is the inverse transport map.

In probabilistic terms, this identity says that μ is the invariant measure of the diffusion process

$$dX_t = -X_t dt + \sqrt{2(\nabla^2\varphi^*)^{-1}}(X_t)dB_t.$$

This SDE can be viewed as a deformation of the classical Ornstein-Uhlenbeck process

$$dX_t = -X_t dt + \sqrt{2}dB_t$$

whose invariant measure is the standard Gaussian $\gamma(dx) = (2\pi)^{-d/2}\exp(-|x|^2/2)dx$. Indeed, the Gaussian is the unique fixed point of the PDE.

What can we say about the PDE?

Theorem (Cordero-Erausquin & Klartag 2015)

If the barycenter of the density ρ is at the origin, then a weak solution exists. It is moreover unique under a regularity condition at the boundary.

Builds up on earlier works of Wang & Zhu (2004), Donaldson (2008) and Berman & Berndtson (2013).

Can be interpreted as a canonical bijection between convex functions and measures, in the spirit of the Minkowski problem relating convex sets and measures on the sphere.

Variational viewpoint

Theorem (Santambrogio 2015)

Solutions arise as the unique critical point of the functional

$$\nu \longrightarrow -\frac{1}{2}W_2^2(\mu,\nu) + \mathrm{Ent}_{dx}(\nu) + \frac{1}{2}\int |x|^2 d\nu$$

where W_2 is the L^2 Wasserstein distance

$$\inf_{X\equiv\mu, Y\equiv\nu} \mathbb{E}[|X-Y|^2]$$

and Ent *the (negative) Boltzmann entropy functional.*

Regularity

When the density ρ is of the form e^{-V} with V convex, we can explicitly get regularity estimates for φ :

Theorem (Klartag & Kolesnikov 2015, Klartag 2014)

If V is convex, then

$$\forall p \geq 2, \int |\partial^2_{ee}\varphi|^p e^{-\varphi} dx \leq 8^p p^{2p} \left(\int (x \cdot e)^2 d\mu \right)^p$$

$$\int \langle (\nabla^2 \varphi)^{-1} \nabla \partial^2_{ee}\varphi, \nabla \partial_{ee}\varphi \rangle e^{-\varphi} dx \leq 32 \int (x \cdot e)^4 d\mu.$$

If moreover $c^{-1} \leq \text{Hess } V \leq c$, then $c^{-1/2} \leq \nabla^2 \varphi \leq c^{1/2}$.

Idea of the proof (Kolesnikov 2013) : the pushforward of the diffusion process by $\nabla\varphi^*$ can be interpreted as the Brownian motion on the weighted Riemannian space $(\mathbb{R}^d, (\nabla^2\varphi)^{-1}, e^{-\varphi})$.

Moreover, $\partial_e\varphi$ are eigenfunctions of the Laplacian, with eigenvalue 1.

It turns out that when V is convex, the Ricci curvature of this manifold is bounded from below by $1/2$.

The previous bounds can be interpreted as regularity estimates for eigenfunctions of the Laplacian on a positively curved manifold.

Other consequence of the curvature bound : exponential convergence to equilibrium for the SDE.

A first result on stability

Theorem (Ledoux, Nourdin & Peccati 2015)

If φ is the solution to the Monge-Ampère PDE with data μ, then

$$W_2(\mu, \gamma)^2 \leq \int |(\nabla^2 \varphi^*)^{-1} - \mathrm{Id}\,|^2_{HS} d\mu$$

This result proves stability of the fixed point of the PDE.
Important property for applications in probability : the estimate does *not* depend on the dimension of the ambient space.
The proof relies on a variant of Stein's method, that compares measures via integration by parts formula. Crucial property : we have an explicit basis of polynomial eigenfunctions for the Ornstein-Uhlenbeck generator. Very few examples satisfy this.

Proving a similar inequality for two non-Gaussian measures is in principle a (very) particular case of proving error estimates on invariant measures using L^2 distance between the coefficients.

There are many ways of doing this. Main difficulty here : all we have is $W^{1,2}$ regularity for the coefficients.

Works such as Le Bris & Lions (2008), Figalli (2008), Trevisan (2016) and Champagnat and Jabin (2018) deal with well-posedness for such irregular coefficients.

Consider two SDE

$$dX_t = -X_t dt + \sqrt{2\tau(X_t)} dB_t; dY_t = -Y_t dt + \sqrt{2\sigma(Y_t)} dB_t$$

which we coupled by taking the same Brownian motion. What follows would also work if we had different drifts.
If $\sqrt{\tau}$ was globally L-lipschitz, we could directly estimate the Wasserstein distance :

$$\begin{aligned}\frac{d}{dt}\mathbb{E}[|X_t - Y_t|^2] &= \mathbb{E}\left[-2|X_t - Y_t|^2 + ||\sqrt{\tau}(X_t) - \sqrt{\sigma}(Y_t)||_{HS}^2\right] \\ &\leq \mathbb{E}\left[-2|X_t - Y_t|^2 + 2||\sqrt{\tau}(X_t) - \sqrt{\tau}(Y_t)||_{HS}^2 + 2||\sqrt{\tau}(Y_t) - \sqrt{\sigma}(Y_t)||_{HS}^2\right] \\ &\leq (2L-2)\mathbb{E}[|X_t - Y_t|^2] + 2\mathbb{E}\left[||\sqrt{\tau}(Y_t) - \sqrt{\sigma}(Y_t)||_{HS}^2\right].\end{aligned}$$

We can then choose the same initial data, and so that (Y_t) is stationary. If the processes converge to equilibrium in W_2 distance, we can get an estimate on the invariant measures out of this.

To go beyond the globally Lipschitz case, one possible tool is the Luzin-Lipschitz property of Sobolev functions : for a.e. x, y

$$|f(x) - f(y)| \leq C(M|\nabla f|(x) + M|\nabla f|(y))|x - y|$$

with M the Hardy-Littlewood maximal operator
$Mg(x) = \sup_{r>0} |B(0,r)|^{-1} \int_{B(0,r)} g(x + y)dy$.
We have $||Mf||_2 \leq C||f||_2$.
Application to stability of transport equations going back to Crippa & De Lellis (2008). Applied to SDE by Li & Luo, and Champagnat & Jabin (2018).

Problem here : maximal estimates are w.r.t. the Lebesgue measure, while the natural reference measure here is the invariant measure of the SDE.

Theorem (Ambrosio, Brué & Trevisan 2018)

If the reference measure μ is log-concave, there exists a continuous operator M on L^2 such that the Luzin-Lipschitz property holds.

So when the invariant measure is log-concave, we can still get estimates that are stable.

Scheme of proof

The idea of Crippa and De Lellis is to control the evolution in time of $\log(1+|X_t - Y_t|^2/\delta)$, with a small parameter δ to be tuned. Letting $g = M|\sqrt{\tau}|$,

$$
\begin{aligned}
\frac{d}{dt}\mathbb{E}[\log(1+|X_t - Y_t|^2/\delta)] &= -2\mathbb{E}\left[\frac{|X_t - Y_t|^2}{\delta + |X_t - Y_t|^2}\right] \\
&+ \mathbb{E}\left[\frac{||\sqrt{\tau}(X_t) - \sqrt{\sigma}(Y_t)||_{HS}^2}{|X_t - Y_t|^2 + \delta}\right] \\
&\leq 2\mathbb{E}[g^2(X_t) + g^2(Y_t)] + \frac{4}{\delta}\mathbb{E}[||\sqrt{\tau}(Y_t) - \sqrt{\sigma}(Y_t)||_{HS}^2].
\end{aligned}
$$

$\mathbb{E}[g(Y_t)^2]$ is controlled by the Sobolev norm of $\sqrt{\tau}$, with respect to ν. Since we only contorle the average w.r.t. μ, we need to somehow transfer the estimate.

So we end up assuming that we have an L^∞ control on the density of μ w.r.t. ν. This is preserved by the flow, but is very strong restriction.

We can then control the Wasserstein distance by $\exp(\mathbb{E}[\log(1 + |X_t - Y|^2/\delta)])$ with a small δ, if we have an a priori control on the second moments. We can then approximate the invariant measure by the law at some large time using a rate of convergence to equilibrium.

Theorem

If $\rho_1 = e^{-V}$ *with* $c^{-1} \leq \text{Hess}\, V \leq c$ *then*

$$W_2(\rho_1, \rho_2)^2 \leq C||\rho_2/\rho_1||_\infty \left(\log(1 + 1/||(\nabla^2\varphi_1^*)^{-1} - (\nabla^2\varphi_2^*)^{-1}||_{L^2(\nu)})\right)^{-1}$$

with a constant C *that explicitly depends on the dimension,* c *and the second moments of* μ *and* ν*.*

The constant obtained effectively scales in $d^4 \log d$, compared to the dimension-free behavior in the Gaussian case.

- Better regularity estimates of φ would help weaken the assumptions, and improve the dependence on the dimension. The geometric interpretation does not seem to help prove $W^{3,p}$ estimates if $p > 2$.
- Does not use regularization from the noise, probably hinting that there is a better method.
- No idea how to do anything in the non-log-concave setting.

The method works for other diffusion. Ingredients : log-concavity of the invariant measure, Sobolev estimates, and a rate of convergence to equilibrium in W_2 distance.

Thanks for your attention !