Distances between stationary distributions of diffusions

Stanislav Shaposhnikov
Department of Mechanics and Mathematics
Lomonosov Moscow State University

Let ξ_t be a diffusion process with generator

$$Lu = \mathrm{tr}(AD^2 u) + \langle b, \nabla u \rangle,$$

for example, a solution of the stochastic differential equation

$$d\xi_t = b(\xi_t)\, dt + \sqrt{2A(\xi_t)}\, dw_t.$$

Let $T_t f(x) = \mathbb{E}_x f(\xi_t)$ be the corresponding transition semigroup. A probability measure μ is a stationary distribution of ξ_t if

$$\int T_t f\, d\mu = \int f\, d\mu \quad \forall f \in C_b(\mathbb{R}^d).$$

The stationary distribution μ is a solution to the stationary Fokker–Planck–Kolmogorov equation

$$\partial_{x_i}\partial_{x_j}\big(a^{ij}\mu\big) - \partial_{x_i}\big(b^i\mu\big) = 0$$

or in the short form

$$L^*\mu = 0.$$

The equation is understood in the sense of distributions:

$$\int_{\mathbb{R}^d} Lu\,d\mu = 0 \quad \forall u \in C_0^\infty(\mathbb{R}^d).$$

Assume that μ and σ are stationary measures of diffusion processes ξ_t^1 and ξ_t^2 with generators

$$L_\mu = \mathrm{tr}(A_\mu D^2) + \langle b_\mu, \nabla\rangle, \quad L_\sigma = \mathrm{tr}(A_\sigma D^2) + \langle b_\sigma, \nabla\rangle.$$

PROBLEM:

to obtain an estimate of the difference between μ and σ.

MOTIVATION:

Investigation of dependence of solutions on the coefficients of the Fokker–Planck–Kolmogorov equation is important for the whole number of nonlinear problems:

- existence and uniqueness of solutions to stationary McKean–Vlasov equations,
- continuity and differentiability of distributions of diffusion processes with respect to a parameter,
- optimal control problems,
- regularity of invariant measures.

Assume, for example, in the case $A_\mu = A_\sigma = I$ we have obtained estimates

$$\|\mu - \sigma\|_{TV} \leq C\|b_\mu - b_\sigma\|_{L^1(\sigma)}$$

or

$$W_1(\mu, \sigma) \leq C\|b_\mu - b_\sigma\|_{L^1(\sigma)},$$

where $\|\cdot\|_{TV}$ is the total variation norm and W_1 is the Kantorovich distance.

Nonlinear Fokker-Planck-Kolmogorov equations

We can derive an existence and uniqueness theorem for the nonlinear equation

$$\Delta\mu - \operatorname{div}(b(x,\mu)\mu) = 0,$$

in which the drift b_μ now depends on the solution μ, for instance

$$b(x,\mu) = \int K(x,y)\,\mu(dy).$$

(H.P. McKean, T.Funaki, A.Yu. Veretennikov, S. Herrmann, J. Tugaut, O.A. Butkovsky, L.G. Tonoyan,...)

Let us consider a mapping F on the space of probability measures defined as follows: $F(\sigma) = \mu$ if $L^*_\sigma \mu = 0$, where L_σ is the operator with the drift b_σ. If

$$|b(x,\mu) - b(x,\sigma)| \leq qC^{-1}\|\mu - \sigma\|_{TV},$$

where $0 < q < 1$, then the mapping F satisfies the estimate

$$\|F(\sigma_1) - F(\sigma_2)\|_{TV} \leq q\|\sigma_1 - \sigma_2\|_{TV},$$

i.e., F is a contracting mapping

Differentiability with respect to a parameter
Let $\mu_\alpha = \varrho_\alpha(x)\,dx$ be a solution of the equation

$$\Delta\varrho_\alpha - \mathrm{div}(b_\alpha\varrho_\alpha) = 0,$$

where b_α depends on a parameter $\alpha \in \mathbb{R}$.
When the mapping $\alpha \to \varrho_\alpha$ is continuous or differentiable?
(E. Pardoux, A.Yu. Veretennikov,...)

If $\alpha \to b_\alpha(x)$ is $L^1(\mu_{\alpha_0})$–differentiable at α_0, then one can estimate L^1-norm of

$$\delta_k \varrho = h_k^{-1}\big(\varrho_{\alpha_0+h_k} - \varrho_{\alpha_0}\big), \quad h_k \to 0.$$

Elliptic regularity theory gives the existence of $h_{k_m} \to 0$ such that $\delta_{k_m}\varrho$ tends to some function w in L^1_{loc} and

$$\Delta w - \mathrm{div}(b_\alpha w) = \mathrm{div}(\varrho_{\alpha_0}\partial_\alpha b).$$

The uniqueness of a solution with zero integral gives that the limit of $\delta_{k_m}\varrho$ does not depend on $\{k_m\}$. Thus we proof that ϱ_α is differentiable in α at α_0. It can be applied to the following optimal control problem

$$\min_\alpha \Big(\int |\varrho_\alpha(x) - \sigma(x)|^2 \, dx + \frac{\alpha^2}{2} \Big).$$

(M. Annunziato, A. Borzi,...)

Regularity of invariant measures

The estimates

$$W_1(\mu,\sigma)\le C\|b_\mu-b_\sigma\|_{L^1(\sigma)},\quad \|\mu-\sigma\|_{TV}\le C\|b_\mu-b_\sigma\|_{L^1(\sigma)}$$

can be used to prove some analogs of the transport inequality and the Poincaré inequality for solutions to the stationary Fokker–Planck–Kolmogorov equation.
(D. Bakry, F.Barthe, P. Cattiaux, A. Guillin, S.G. Bobkov, F. Bolley, I. Gentil, F.-Y. Wang,...)

It suffices to apply the above estimates to the measures μ and $\sigma = f \cdot \mu$, where σ satisfies the equation with the drift

$$b_\sigma = \frac{\nabla f}{f} + \frac{\nabla \varrho_\mu}{\varrho_\mu}.$$

Then we obtain

$$W_1(\mu, \sigma) \leq C \int |\nabla f| \, d\mu + C \int \Big| \frac{\nabla \varrho_\mu}{\varrho_\mu} - b_\mu \Big| f \, d\mu$$

or

$$\int |1 - f| \, d\mu \leq C \int |\nabla f| \, d\mu + C \int \Big| \frac{\nabla \varrho_\mu}{\varrho_\mu} - b_\mu \Big| f \, d\mu.$$

Suppose now that we are given two Borel probability measures $\mu = \varrho_\mu \, dx$ and $\sigma = \varrho_\sigma \, dx$ satisfying the equations

$$\Delta\mu - \mathrm{div}(b_\mu \mu) = 0 \quad \text{and} \quad \Delta\sigma - \mathrm{div}(b_\sigma \sigma) = 0.$$

Set

$$L_\mu = \Delta + \langle b_\mu, \nabla \rangle, \quad L_\sigma = \Delta + \langle b_\sigma, \nabla \rangle.$$

Set

$$v(x) = \frac{\varrho_\sigma(x)}{\varrho_\mu(x)}, \quad \sigma = v \cdot \mu.$$

Theorem

Suppose that $|b_\mu - b_\sigma| \in L^2(\sigma)$ *and at least one of the following two conditions is fulfilled*

(i) $(1+|x|)^{-1}|b_\mu(x)| \in L^1(\mu)$,

(ii) *there exists a nonnegative function* $V \in C^2(\mathbb{R}^d)$ *such that* $L_\mu V(x) \leq MV(x)$ *for all* x *and some* $M > 0$ *and*

$$\lim_{|x|\to+\infty} V(x) = +\infty, \quad \langle b_\mu - b_\sigma, \nabla V\rangle(1+V)^{-1} \in L^1(\sigma).$$

Then the estimate

$$\int_{\mathbb{R}^d} \frac{|\nabla v|^2}{v}\, d\mu \leq \int_{\mathbb{R}^d} |b_\mu - b_\sigma|^2\, d\sigma$$

holds, which yields, in particular, the inclusion $\sqrt{v} \in W^{2,1}(\mu)$.

Recall that a probability measure μ satisfies the logarithmic Sobolev inequality with constant C if

$$\mathrm{Ent}_\mu(f^2) := \int_{\mathbb{R}^d} f^2 \ln f^2 \, d\mu - \int_{\mathbb{R}^d} f^2 \, d\mu \ln \int_{\mathbb{R}^d} f^2 \, d\mu \leq C \int_{\mathbb{R}^d} |\nabla f|^2 \, d\mu.$$

for every $f \in W^{2,1}(\mu)$.
Note that the condition

$$\langle b(x) - b(y), x - y \rangle \leq -\kappa |x - y|^2$$

ensures that a probability solution μ of the equation

$$\Delta\mu - \mathrm{div}(b\mu) = 0$$

satisfies the logarithmic Sobolev inequality with constant $2/\kappa$.

Corollary

Suppose that, in addition to the hypotheses of Theorem, it is assumed that the solution μ satisfies the logarithmic Sobolev inequality with constant C. Then

$$\mathrm{Ent}_\mu v \leq \frac{C}{4} \int_{\mathbb{R}^d} |b_\mu - b_\sigma|^2 \, d\sigma.$$

Example

If $\langle b_\mu(x) - b_\mu(y), x - y \rangle \leq -\kappa |x - y|^2$ and $|b_\mu - b_\sigma| \in L^2(\sigma)$, then

$$\mathrm{Ent}_\mu v \leq \frac{1}{2\kappa} \int_{\mathbb{R}^d} |b_\mu - b_\sigma|^2 \, d\sigma.$$

Assume that $|x| \in L^2(\mu + \sigma)$. If the measure μ satisfies the logarithmic Sobolev inequality with constant C then the probability measure $\sigma = v \cdot \mu$ satisfies the so-called transport inequality

$$W_2^2(\mu, \sigma) \leq C \mathrm{Ent}_\mu v,$$

where the Kantorovich distance $W_2(\mu, \sigma)$ is defined as the infimum of

$$\Big(\int \int |x - y|^2 \, d\pi \Big)^{1/2}$$

over all probability measures π on $\mathbb{R}^d \times \mathbb{R}^d$ with projections μ and σ on the factors.

Corollary

Suppose that, in addition to the hypotheses of Theorem, it is assumed that the solutions μ and σ have second moments and the measure μ satisfies the logarithmic Sobolev inequality with constant C. Then

$$W_2(\mu, \sigma) \leq \frac{C}{2} \|b_\mu - b_\sigma\|_{L^2(\sigma)}.$$

Example

If $\langle b_\mu(x) - b_\mu(y), x - y\rangle \leq -\kappa|x - y|^2$ and $|b_\mu - b_\sigma| \in L^2(\sigma)$, then

$$W_2(\mu, \sigma) \leq \frac{1}{\kappa} \|b_\mu - b_\sigma\|_{L^2(\sigma)}.$$

Recall the classical Pinsker–Csiszár–Kullback inequality

$$\|\mu - v \cdot \mu\|_{TV}^2 \leq 2\mathrm{Ent}_\mu v$$

or the estimate established by F. Bolley and C. Villani (2005):

$$\|\varphi(\mu - v \cdot \mu)\|_{TV}^2 \leq 2\Big(1 + \log\Big(\int_{\mathbb{R}^d} e^{\varphi^2}\, d\mu\Big)\Big)\mathrm{Ent}_\mu v. \qquad (1)$$

for two probability measures μ and $\sigma = v \cdot \mu$ on $\mathbb{R}^d$ and a Borel function $\varphi \geq 0$.

Corollary

Suppose that, in addition to the hypotheses of Theorem, it is assumed that the solution μ satisfies the logarithmic Sobolev inequality with constant C. Then

$$\|\mu - \sigma\|_{TV}^2 \leq \frac{C}{2} \int_{\mathbb{R}^d} |b_\mu - b_\sigma|^2 \, d\sigma.$$

Example

If $\langle b_\mu(x) - b_\mu(y), x - y \rangle \leq -\kappa |x - y|^2$ and $|b_\mu - b_\sigma| \in L^2(\sigma)$, then

$$\|\mu - \sigma\|_{TV}^2 \leq \frac{1}{\kappa} \int_{\mathbb{R}^d} |b_\mu - b_\sigma|^2 \, d\sigma.$$

We shall say that a probability measure μ satisfies the Poincaré inequality with the constant C_P if

$$\int_{\mathbb{R}^d} \left| f - \int_{\mathbb{R}^d} f \, d\mu \right|^2 d\mu \leq C_P \int_{\mathbb{R}^d} |\nabla f|^2 \, d\mu$$

for every function $f \in C_0^\infty(\mathbb{R}^d)$.
We recall that the Hellinger integral is the quantity

$$H(\mu, \sigma) = \int_{\mathbb{R}^d} \sqrt{\varrho_\mu \varrho_\sigma} \, dx.$$

Corollary

If in addition to the conditions of Theorem it is known that the measure μ satisfies the Poincaré inequality with the constant C_P, then the following estimates are valid:

$$1 - H(\mu,\sigma)^2 \leq \frac{C_P}{4} \int_{\mathbb{R}^d} |b_\mu - b_\sigma|^2 \, d\sigma,$$

$$\|\mu - \sigma\|_{TV}^2 \leq C_P \int_{\mathbb{R}^d} |b_\mu - b_\sigma|^2 \, d\sigma.$$

Proof.

By the Poincaré inequality

$$\int_{\mathbb{R}^d} \left| \sqrt{v} - \int_{\mathbb{R}^d} \sqrt{v}\, d\mu \right|^2 d\mu \leq \frac{C_P}{4} \int_{\mathbb{R}^d} \frac{|\nabla v|^2}{v}\, d\mu.$$

For the proof of the first inequality it suffices to observe that

$$\int_{\mathbb{R}^d} \left| \sqrt{v} - \int_{\mathbb{R}^d} \sqrt{v}\, d\mu \right|^2 d\mu = 1 - H(\mu, \sigma)^2.$$

The second inequality follows from the first one and the inequality $\|\mu - \sigma\|_{TV} \leq 2\sqrt{1 - H(\mu, \sigma)^2}$. □

Theorem proof

1. The function v satisfies the equation

$$\mathrm{div}(\varrho_\mu \nabla v - va - vh\varrho_\mu) = 0,$$

where

$$a = \varrho_\mu b_\mu - \nabla \varrho_\mu, \quad \mathrm{div} a = 0, \quad h = b_\sigma - b_\mu.$$

2. Let $f \in C^1((0,+\infty))$ and $f'' \geq 0$. The following estimate

$$\int_{\mathbb{R}^d} |\nabla v|^2 f''(v)\psi \, d\mu \leq$$
$$\leq 2\int_{\mathbb{R}^d} f(v) L_\mu \psi \, d\mu + 2\int_{\mathbb{R}^d} |h||\nabla \psi| f'(v) v \, d\mu$$
$$+ \int_{\mathbb{R}^d} |h|^2 f''(v) v \psi \, d\sigma$$

holds for every $\psi \in C_0^\infty(\mathbb{R}^d)$, $\psi \geq 0$.

3. Let

$$\psi_N \to 1, \quad |\nabla \psi_N| \to 0, \quad L_\mu \psi_N \to 0.$$

We obtain

$$\int_{\mathbb{R}^d} |\nabla v|^2 f''(v)\, d\mu \le \int_{\mathbb{R}^d} |h|^2 f''(v) v\, d\sigma.$$

Finally, we take $f(v) = v \ln v$.

The another approach is based on the Poisson equation

$$Lu = \psi.$$

The idea of this approach is very simple. Suppose that for every bounded smooth function ψ we are able to solve the Poisson equation

$$L_\mu u = \widetilde{\psi}$$

with the right-hand side

$$\widetilde{\psi} := \psi - \int_{\mathbb{R}^d} \psi \, d\mu.$$

Then for the solution u we have

$$\int_{\mathbb{R}^d} \psi \, d(\mu - \sigma) = \int_{\mathbb{R}^d} \langle \nabla u, b_\mu - b_\sigma \rangle \, d\sigma.$$

In the last equality we write ψ in place of $\widetilde{\psi}$, since the integral of any constant against the measure $\mu - \sigma$ is zero. If, for example, we know that the boundedness of ψ yields the boundedness of $|\nabla u|$, then we immediately obtain the estimate

$$\|\mu - \sigma\|_{TV} \leq C\|b_\mu - b_\sigma\|_{L^1(\sigma)}.$$

If for the boundedness of $|\nabla u|$ the boundedness of $|\nabla \psi|$ is needed (this is the case if we obtain this estimate by differentiating the equation $Lu = \psi$ and applying the maximum principle), then we obtain the estimate

$$W_1(\mu, \sigma) \leq C\|b_\mu - b_\sigma\|_{L^1(\sigma)}.$$

However, it is possible to use the equation $Lu = \psi$ only for estimating $|u|$, and to derive estimates on the gradient $|\nabla u|$ from the equality

$$\int_{\mathbb{R}^d} |\nabla u|^2 \, d\sigma = - \int_{\mathbb{R}^d} [u\psi + u\langle b_\mu - b_\sigma, \nabla u\rangle] \, d\sigma.$$

In this case we obtain an estimate on $\|\mu - \sigma\|_{TV}$ via $\|b_\mu - b_\sigma\|_{L^2(\sigma)}$.

Theorem

Assume that

$$|b_\mu(x)| + |b_\sigma(x)| \leq C(1+|x|)^k \quad \forall x \in \mathbb{R}^d$$

and some $C > 0$ and $k > 0$. Assume also that

$$\langle b_\mu(x), x\rangle \leq -\gamma|x|^2 \quad \text{for some} \quad \gamma > 0.$$

If the measure σ has a finite moment of order $2k+2$, then

$$\|(1+|x|^2)(\mu-\sigma)\|_{TV} \leq C\|b_\mu - b_\sigma\|_{L^2(\sigma)}\|(1+|x|^{k+1})\|_{L^2(\sigma)}.$$

Let us consider the third method of obtaining estimates for the distance between stationary distributions, which relates such estimates with the properties of the semigroup generated by the operator L.

Important advantages of such estimates are the absence of dependence on the dimension (our estimates depend on the dimension precisely as the employed estimates for the semigroup). Note that the semigroup generated by L is usually applied for deriving the logarithmic Sobolev inequality or the Poincare inequality for the measure and for a priori estimates for the solution to the Poisson equation.

Let us explain the new method of obtaining bounds by a relatively simple, but important example.

Assume that

$$\mu = \frac{1}{(2\pi)^{d/2}} e^{-|x|^2/2}\, dx, \quad L_\mu = \Delta - \langle x, \nabla \rangle, \quad L_\sigma = \Delta + \langle b_\sigma, \nabla \rangle,$$

$$b_\sigma(x) = -x + h(x).$$

Thus L_μ is the Ornstein–Uhlenbeck operator.

Let T_t be the Ornstein–Uhlenbeck semigroup defined by the formula

$$T_t\varphi(x) = \int_{\mathbb{R}^d} \varphi\Big(e^{-t}x + \sqrt{1 - e^{-2t}}\, y\Big)\, \mu(dy).$$

The measure μ is the invariant measure for T_t and

$$\nabla T_t\varphi = e^{-t} T_t \nabla\varphi, \tag{2}$$

$$|\nabla T_t\varphi(x)| \leq \frac{e^{-t}}{\sqrt{1 - e^{-2t}}} \|\varphi\|_\infty, \tag{3}$$

and

$$\lim_{t\to\infty} T_t\varphi(x) = \int \varphi\, d\mu.$$

Assume that $|h| \in L^1(\sigma)$. Then for $t > 0$ from the Fokker–Planck–Kolmogorov equation one can easily derive the equality

$$\int L_\sigma T_t\varphi \, d\sigma = 0.$$

Therefore,

$$\int L_\mu T_t\varphi \, d\sigma = -\int \langle h, \nabla T_t\varphi \rangle \, d\sigma.$$

Taking into account the invariance of the measure μ and the equality

$$L_\mu T_t\varphi = \frac{d}{dt} T_t\varphi,$$

we obtain

$$-\frac{d}{dt}\int T_t\varphi \, d(\mu - \sigma) = \int \langle h, \nabla T_t\varphi \rangle \, d\sigma. \tag{4}$$

Note that

$$\int_0^\infty \frac{e^{-t}}{\sqrt{1-e^{-2t}}}\,dt = \int_0^1 \frac{ds}{\sqrt{1-s^2}} = \frac{\pi}{2}$$

Let $|\varphi| \leq 1$. Then on account of inequality (3) we have

$$\int_0^{+\infty} \int \langle h, \nabla T_t\varphi\rangle \, d\sigma \, dt \leq \frac{\pi}{2}\|h\|_{L^1(\sigma)}.$$

Integrating (4) in t from 0 to $+\infty$ and taking into account that

$$\lim_{t\to+\infty} T_t\varphi(x) = \int \varphi \, d\mu,$$

we deduce the estimate

$$\int \varphi \, d(\mu - \sigma) \leq \frac{\pi}{2}\|h\|_{L^1(\sigma)},$$

from which the estimate of the variation norm of the measure $\mu - \sigma$ follows, since φ was arbitrary.
In the same way one can obtain the estimate on the Kantorovich metric W_1

Theorem

Let μ be the standard Gaussian measure and let σ be a probability measure that is a solution to the equation with the operator $L_\sigma^* \sigma = 0$ with the operator

$$L_\sigma u(x) = \Delta u(x) + \langle -x + h(x), \nabla u(x) \rangle,$$

such that $|h| \in L^1(\sigma)$. Then

$$\|\mu - \sigma\|_{TV} \leq \frac{\pi}{2} \|h\|_{L^1(\sigma)}.$$

In addition,

$$W_1(\mu, \sigma) \leq \|h\|_{L^1(\sigma)}.$$

The above theorem can be generalised in differen way. For instance if

$$\langle b_\mu(x) - b_\mu(y), x - y\rangle \leq -\kappa|x-y|^2, \quad \kappa > 0,$$

then

$$\|\mu - \sigma\|_{TV} \leq C(\kappa)\|b_\mu - b_\sigma\|_{L^1(\sigma)}.$$

If $b_\mu(x) = -x|x|^k + h(x)$, where

$$\langle h(x) - h(y), x - y\rangle \leq \kappa|x-y|^2, \quad \kappa > 0,$$

and $2 < r < (k-2)/2$, then

$$\|(1 + |x|^r)(\mu - \sigma)\|_{TV} \leq C\|b_\mu - b_\sigma\|_{L^1(\sigma)}.$$

The presented talk is based on the following papers:

1. Bogachev V.I., Kirillov A.I., Shaposhnikov S.V. The Kantorovich and variation distances between invariant measures of diffusions and nonlinear stationary Fokker-Planck-Kolmogorov equations. Mathematical Notes, 2014, 96(6), 855-863.
2. Bogachev V.I., Kirillov A.I., Shaposhnikov S.V. Distances between stationary distributions of diffusions and solvability of nonlinear Fokker-Planck-Kolmogorov equations. Teor. Veroyatnost. i Primenen., 2017, 62(1), 16-43.
3. Bogachev V.I., Röckner M., Shaposhnikov S.V. The Poisson equation and estimates for distances between stationary distributions of diffusions. J. Math. Sci. (New York) 232 (3), 254–282 (2018).

4. Bogachev V.I., Miftakhov A.F., Shaposhnikov S.V. Differential Properties of Semigroups and Estimates of Distances between Stationary Distributions of Diffusions. Doklady Mathematics, 2019, Vol. 99, N. 2, p. 175–180.
5. V.I.Bogachev, N.V.Krylov, M.Röckner, S.V.Shaposhnikov, Fokker–Planck–Kolmogorov Equations, Amer. Math. Soc., Rhode Island, Providence, 2015.

THANK YOU!