Distances between transition probabilities of diffusions and applications.

Stanislav Shaposhnikov

Department of Mechanics and Mathematics
Moscow State University

Assume that there are two diffusion processes ξ_t^1 and ξ_t^2 such that

$$d\xi_t^1 = \sigma_1(\xi_t^1)dw_t + b_1(\xi_t^1)\,dt, \quad d\xi_t^2 = \sigma_2(\xi_t^2)dw_t + b_2(\xi_t^2)\,dt,$$

and $\xi_0^1 = \xi_0^2 = x_0$.

PROBLEM: to obtain an estimate of the difference between the corresponding distributions μ_t^1 and μ_t^2:

$$\mu_t^1(B) = P(\xi_t^1 \in B), \quad \mu_t^2(B) = P(\xi_t^2 \in B).$$

Applications:

- Nonlinear equations
- Optimal control problems
- Mean Field Games
- Computer simulations
- Ergodicity problems

Fokker–Planck–Kolmogorov equations

We recall that μ_t^1 and μ_t^2 satisfy the following equations

$$\partial_t \mu_t^{1,2} = \partial_{x_i}\partial_{x_j}\Big(a^{ij}_{1,2}\mu_t^{1,2}\Big) - \partial_{x_i}\Big(b^i_{1,2}\mu_t^{1,2}\Big),$$

where

$$A_{1,2} = \sigma_{1,2}\sigma^*_{1,2}/2$$

and we assume the summation over the repeated indexes.

Thus we would like to estimate the distance between μ_t^1 and μ_t^2 over the difference of $b_1 - b_2$ and $A_1 - A_2$.

We consider the Cauchy problem

$$\partial_t \mu = \partial_{x_i}\partial_{x_j}(a^{ij}\mu) - \partial_{x_i}(b^i\mu), \quad \mu|_{t=0} = \nu. \tag{1}$$

- $A(x,t) = (a^{ij}(x,t))_{1\le i,j\le d}$ is a positive symmetric matrix (called the diffusion matrix) with Borel measurable entries,
- $b(x,t) = (b^i(x,t))_{1\le i\le d}$ is a Borel measurable mapping (called the drift coefficient),
- ν is a probability measure on $\mathbb{R}^d$.

Set $Lu = a^{ij}\partial_{x_i}\partial_{x_j}u + b^i\partial_{x_i}u, \quad L^*u = \partial_{x_i}\partial_{x_j}(a^{ij}u) - \partial_{x_i}(b^iu)$.
Then the equation from (1) can be written shortly:

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

A solution $\mu = \mu_t(dx)\,dt$ is given by a family $(\mu_t)_{t\in(0,T)}$ of probability measures μ_t on $\mathbb{R}^d$ such that $t \mapsto \mu_t(B)$ is measurable for every Borel set $B \subset \mathbb{R}^d$ and for every $\varphi \in C_0^\infty(\mathbb{R}^d)$ the equality

$$\int_{\mathbb{R}^d} \varphi\, d\mu_t = \int_{\mathbb{R}^d} \varphi\, d\nu + \int_0^t \int_{\mathbb{R}^d} L\varphi\, d\mu_s\, ds$$

holds for almost all $t \in (0, T)$.

Suppose that

- b is locally bounded, i.e., for every ball $U \subset \mathbb{R}^d$, there is a number $B = B(U) \geq 0$ such that

$$|b(x,t)| \leq B(U) \quad \forall\, x \in U,\ t \in [0,T],$$

- A is locally Lipschitzian in x and locally strictly positive, i.e., for every ball $U \subset \mathbb{R}^d$, there exist numbers $\lambda = \lambda(U) \geq 0$, $\alpha = \alpha(U) > 0$ and $m = m(U) > 0$ such that

$$|a^{ij}(x,t) - a^{ij}(y,t)| \leq \lambda |x-y|, \quad \alpha \cdot \mathrm{I} \leq A(x,t) \leq m \cdot \mathrm{I}$$

for every $x, y \in U,\ t \in [0,T]$.

Under this assumptions there hold

- **Existence:** for every probability measure ν there exists a subprobability solution $\mu = \mu_t(dx)\,dt$ ($\mu_t \geq 0$ and $\mu_t(\mathbb{R}^d) \leq 1$) of the Cauchy problem (1). Moreover if at least one of the following two conditions is fulfilled:
  (i) $(1+|x|)^{-2}|a^{ij}|,\ (1+|x|)^{-1}|b|, \in L^1(\mathbb{R}^d \times [0,T], \mu)$,
  (ii) there exist a nonnegative function $V \in C^2(\mathbb{R}^d)$ (Lyapunov function) and a number $M \geq 0$ such that $\lim_{|x|\to\infty} V(x) = +\infty$ and $LV \leq MV$,
  then μ_t are probability measures ($\mu_t \geq 0$ and $\mu_t(\mathbb{R}^d) = 1$).
- **Uniqueness:** if at least one of the conditions (i) or (ii) is fulfilled, then such solution is unique.

Example

Let $A = I$. Suppose that for some numbers $\gamma_1 > 0$ and $\gamma_2 > 0$ we have

$$\langle b(x,t), x\rangle \le \gamma_1 + \gamma_2 |x|^2.$$

Then there exists a unique probability solution.

Example

There exists a smooth function B on $\mathbb{R}$ such that the probability solution ν of the equation $\nu'' - (B\nu)' = 0$ is not invariant measure for the corresponding semigroup T_t with the generator L but only subinvariant: $T_t^*\nu < \nu$ if $t > 0$.
Let now $C(y) = (C^1(y), C^2(y))$ be a smooth vector field on $\mathbb{R}^2$ for which there are two different probability solutions σ^1 and σ^2 of the equation $\Delta\sigma - \mathrm{div}(C\sigma) = 0$. Set

$$\mu_t^1 = \nu \otimes \sigma_1, \quad \mu_t^2 = (\nu - T_t^*\nu) \otimes (\sigma_2 - \sigma_1) + \nu \otimes \sigma_1.$$

We construct two different probability solutions of the Cauchy problem $\partial_t \mu_t = \Delta\mu - \mathrm{div}(b\mu)$, $\mu_0 = \nu \otimes \sigma_1$, where $b = (B(x), C^1(y), C^2(y))$.

Let us formulate the main result.

Let $\mu = \varrho_\mu(x,t)\,dxdt$ and $\sigma = \varrho_\sigma(x,t)\,dxdt$. Set

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

Let us introduce vector mappings

$$h_\mu = (h_\mu^i)_{i=1}^d, \quad h_\sigma = (h_\sigma^i)_{i=1}^d, \quad h_\mu^i = b_\mu^i - \sum_{j=1}^d \partial_{x_j} a_\mu^{ij},$$

$$h_\sigma^i = b_\sigma^i - \sum_{j=1}^d \partial_{x_j} a_\sigma^{ij}.$$

Set

$$\Phi = \frac{(A_\mu - A_\sigma)\nabla\varrho_\sigma}{\varrho_\sigma} - (h_\mu - h_\sigma).$$

The latter mapping is crucial: the distances between μ_t and σ_t will be estimated through the $L^2(\sigma)$-norm of $A_\mu^{-1/2}\Phi$. Observe that in case of equal diffusion matrices we obtain just the difference of the drifts:

$$\Phi = b_\sigma - b_\mu.$$

In case of equal drifts and constant diffusion matrices, only the first term of this mapping appears.

Theorem

Let $|A_\mu^{-1/2}\Phi| \in L^2(\mathbb{R}^d \times [0,T], \sigma)$*. Suppose also that at least one of the following two conditions is fulfilled*:

(a)

$(1+|x|)^{-2}|a_\mu^{ij}|,\ (1+|x|)^{-1}|b_\mu|,\ (1+|x|)^{-1}|\Phi| \in L^1(\mathbb{R}^d \times [0,T], \sigma)$.

(b) *there exist a nonnegative function* $V \in C^2(\mathbb{R}^d)$ *and a number* $M \geq 0$ *such that*

$$\lim_{|x|\to\infty} V(x) = +\infty, \quad L_{A_\mu, b_\mu} V \leq MV, \quad \frac{\langle \Phi, \nabla V\rangle}{1+V} \in L^1(\mathbb{R}^d \times [0,T], \sigma).$$

Then

$$H(\sigma_t|\mu_t) = \int_{\mathbb{R}^d} v \log v \, d\mu_t \leq \frac{1}{2}\int_0^t \int_{\mathbb{R}^d} \left|A_\mu^{-1/2}\Phi\right|^2 d\sigma_s \, ds. \qquad (2)$$

Recall the classical Pinsker–Csiszár–Kullback inequality

$$\|\mu-\sigma\|_{TV}^2 \le 2H(\sigma|\mu)$$

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

$$\|\varphi(\mu-\sigma)\|_{TV}^2 \le 2\left(1+\log\left(\int_{\mathbb{R}^d} e^{\varphi^2}\,d\mu\right)\right)\int_{\mathbb{R}^d} v\log v\,d\mu \qquad (3)$$

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

Corollary

Let $A_\mu = A_\sigma = \mathrm{I}$. *Under the assumptions of the theorem, for every nonnegative Borel measurable function* φ *on* $\mathbb{R}^d \times [0, T]$, *we have*

$$\|\varphi(\mu_t - \sigma_t)\|_{TV}^2 \leq \big(1 + \log \alpha(t)\big) \int_0^t \int_{\mathbb{R}^d} |b_\mu - b_\sigma|^2 \, d\sigma_s \, ds,$$

where

$$\alpha(t) := \int_{\mathbb{R}^d} e^{\varphi^2(x,t)} \, \mu_t(dx).$$

Finally, in case $\varphi = 1$ *these bounds hold with* 1 *in place of* $1 + \log \alpha(t)$.

Corollary

Let $A_\mu = A_\sigma = I$. Suppose that for some numbers $\gamma_1 > 0$ and $\gamma_2 > 0$ we have

$$\langle b_\mu(x,t), x\rangle \leq \gamma_1 + \gamma_2 |x|^2.$$

Then

$$\|\mu_t - \sigma_t\|_{TV}^2 \leq \int_0^t \int_{\mathbb{R}^d} |b_\mu - b_\sigma|^2 \, d\sigma_s \, ds.$$

Moreover, for any $p \geq 1$ and $K > 0$ the following estimate holds:

$$\|(1+|x|^p)(\mu_t - \sigma_t)\|_{TV}^2 \leq N(t) \int_0^t \int_{\mathbb{R}^d} |b_\mu - b_\sigma|^2 \, d\sigma_s \, ds,$$

where

$$N(t) = 2K^{-1}\left(1 + \log\Bigl(\int_{\mathbb{R}^d} e^{K(1+|x|^p)^2} \mu_t(dx)\Bigr)\right).$$

Note that if $A_\mu = I$ and for some $p \geq 1$, $K > 0$, $\gamma_1 > 0$ and $\gamma_2 > 2pK$ we have

$$\langle b_\mu(x,t), x\rangle \leq \gamma_1 - \gamma_2 |x|^{2p},$$

then for some $C > 0$ and all $t \in [0, T]$ one has by Gronwall's inequality

$$\int_{\mathbb{R}^d} e^{K|x|^{2p}}\, \mu_t(dx) \leq e^{Ct} + e^{Ct} \int_{\mathbb{R}^d} e^{K|x|^{2p}}\, \nu(dx).$$

Let α, m and Λ do not depend on U and $\nu = \varrho_0\, dx$. The following estimate of the $L^2(\sigma)$-norm of $\nabla \varrho_\sigma / \varrho_\sigma$ holds:

$$\int_0^\tau \int_{\mathbb{R}^d} \frac{|\nabla \varrho_\sigma|^2}{\varrho_\sigma}\, dx\, dt \leq$$
$$\leq C\Big(1 + \|b_\sigma\|^2_{L^2(\sigma)} + \int_{\mathbb{R}^d} \varrho_0 \ln \varrho_0\, dx +$$
$$\int_{\mathbb{R}^d} \ln(\max|x|, 1)\varrho_\sigma(x, \tau)\, dx\Big).$$

(Bogachev V.I., Röckner M., Shaposhnikov S.V 2005)

Corollary

Assume also that $|x|^{2m} \in L^1(\nu)$, $\nu = \varrho_0\, dx$, $\varrho_0 \ln \varrho_0 \in L^1(\mathbb{R}^d)$ *and*

$$\langle b_\mu(x,t), x\rangle \le \gamma_1 + \gamma_2 |x|^2, \quad |b_\sigma(x,t)| \le \gamma_3 + \gamma_4 |x|^m$$

for some numbers $m, \gamma_i \ge 0$. *Then*

$$\|\mu_t - \sigma_t\|_{TV}^2 \le \sup_{x,t} \|A_\mu - A_\sigma\|^2 C(T) + 2\alpha^{-1} \int_0^t \int_{\mathbb{R}^d} |h_\mu - h_\sigma|^2 \, d\sigma_s \, ds,$$

where $C(T)$ *depends on* $T, m, \alpha, \Lambda, \gamma_i, \int |x|^{2m} \, d\nu$, *and* $\|\varrho_0 \ln \varrho_0\|_{L^1(\mathbb{R}^d)}$.

Proof.
Renormalized solutions
(I) (R.J. DiPerna, P.L. Lions 1989)
Let us consider the Cauchy problem for the continuity equation

$$\partial_t u + \mathrm{div}(bu) = 0, \quad u|_{t=0} = u_0.$$

We say that u is a renormalized solution if

$$\partial_t f(u) + \mathrm{div}(bf(u)) \leq (f(u) - uf'(u))\mathrm{div} b$$

for every convex function f. For example, if the above inequality holds, then for $f(u) = |u|$ we obtain (unformally)

$$\frac{d}{dt}\int_{\mathbb{R}} |u|\, dx \leq 0,$$

that implies the uniqueness.
(C. Le Bris, P.L. Lions 2008)

(II) (P. Bauman 1984)
Let us consider the elliptic equation of the double devergence form:

$$\partial_{x_i}\partial_{x_j}\left(a^{ij}u\right) = 0.$$

Fix a positive solution u and for another solution w we introduce the function $v = w/u$. The function v satisfies to the new equation with the matrix $A \cdot u$. It turns out that the renormalized solution v possesses many nice properties: the maximum principle, Harnack's inequality, Hölder's continuity with constants which are independent of the smoothness A.
(L. Escauriaza 2000)

The proof of the main theorem is based on the combination of this two methods.
Set $v = \varrho_\sigma/\varrho_\mu$. Then for every $f \in C^2((0,+\infty))$

$$\partial_t\big(\varrho_\mu f(v)\big) = L_\mu^*\big(\varrho_\mu f(v)\big) - \varrho_\mu f''(v)|\sqrt{A_\mu}\nabla v|^2 - f'(v)\mathrm{div}\big(\Phi\varrho_\sigma\big).$$

Multiplying this equation by the function $\psi \in C_0^\infty(\mathbb{R}^d)$ and integrating, we arrive at the equality

$$
\begin{aligned}
&\int_{\mathbb{R}^d} f(v(x,t))\psi(x)\varrho_\mu(x,t)\,dx+\int_0^t\int_{\mathbb{R}^d}\psi f''(v)|\sqrt{A_\mu}\nabla v|^2\varrho_\mu\,dx\,ds=\\
&\quad= f(1)\int_{\mathbb{R}^d}\psi\,d\nu+\int_0^t\int_{\mathbb{R}^d}\big[f(v)L_\mu\psi\big]\varrho_\mu\,dx\,ds+\\
&\quad+\int_0^t\int_{\mathbb{R}^d}\big[\langle\Phi,\nabla v\rangle f''(v)\psi\varrho_\sigma+f'(v)\langle\Phi,\nabla\psi\rangle\varrho_\sigma\big]\,dx\,ds.
\end{aligned}
$$

Assume that $\psi \geq 0$ and $f'' \geq 0$. Applying the Cauchy inequality we obtain

$$\int_{\mathbb{R}^d} f(v(x,t))\psi(x)\varrho_\mu(x,t)\,dx \leq f(1)\int_{\mathbb{R}^d} \psi\,d\nu +$$
$$+\frac{1}{2}\int_0^t\int_{\mathbb{R}^d} |A_\mu^{-1/2}\Phi|^2 f''(v)v\varrho_\sigma\,dx\,ds +$$
$$+\int_0^t\int_{\mathbb{R}^d}\big[f(v)L_\mu\psi\big]\varrho_\mu\,dx\,ds +$$
$$+\int_0^t\int_{\mathbb{R}^d}\big[f'(v)\langle\Phi,\nabla\psi\rangle\varrho_\sigma\big]\,dx\,ds.$$

Let ψ_N be such that

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

Replace in the above inequality ψ by ψ_N and tend $N \to \infty$ we obtain

$$\int_{\mathbb{R}^d} f(v(x,t))\varrho_\mu(x,t)\,dx \le f(1) + \frac{1}{2}\int_0^t \int_{\mathbb{R}^d} |A_\mu^{-1/2}\Phi|^2 f''(v) v \varrho_\sigma \,dx\,ds.$$

Setting $f(v) = v \ln v$ we derive the assertion of the theorem.

Nonlinear equations

Suppose now that for every measure μ on $\mathbb{R}^d \times (0, T)$ given by a family $(\mu_t)_{t\in(0,T)}$ of probability measures on $\mathbb{R}^d$ we are given a locally bounded Borel measurable mapping

$$b(\mu, \cdot, \cdot)\colon\ \mathbb{R}^d \times [0, T] \to \mathbb{R}^d.$$

Then we can consider the Cauchy problem for the nonlinear Fokker–Planck–Kolmogorov equation

$$\partial_t\mu = \Delta\mu - \mathrm{div}(b(\mu, x, t)\mu), \quad \mu|_{t=0} = \nu. \tag{4}$$

(McKean, H.P. 1966, Funaki, T. 1984, Veretennikov A. Yu. 2006, 2016, Manita O.A., Romanov M.S., Shaposhnikov S.V. 2015)

Let $C^+[0,T]$ denote the set of nonnegative continuous functions on $[0,T]$. Suppose that $V \in C^2(\mathbb{R}^d)$ and $V \geq 1$. For $\alpha \in C^+[0,T]$ and $\tau \in (0,T]$ we set

$$\mathcal{M}_{\tau,\alpha}(V) = \Big\{ \mu(dxdt) = \varrho(x,t)\,dx\,dt\colon\ \varrho \geq 0, \int \varrho(x,t)\,dx = 1, \int_{\mathbb{R}^d} V(x)\varrho(x,t)\,dx \leq \alpha(t), t \in [0,\tau] \Big\}.$$

If $V(x) = e^{K|x|^{2p}}$, then the corresponding set $\mathcal{M}_{\tau,\alpha}(V)$ will be denoted by $\mathcal{M}^{K,p}_{\tau,\alpha}$.
Let $\|\varrho\|_{p,\tau}$ be the norm defined by

$$\|\varrho\|^2_{p,\tau} := \int_0^\tau \Big(\int (1+|x|^p)|\varrho(x,t)|\,dx \Big)^2\,dt.$$

Corollary

Let $p \geq 1$, $K > 0$ and suppose that for every function $\alpha \in C^+[0,T]$ there exist numbers $\gamma_1(\alpha) > 0$ and $\gamma_2(\alpha) > 2pK$ such that for every $\tau \in (0,T]$ and $\mu \in \mathcal{M}^{K,p}_{\tau,\alpha}$ one has

$$\langle b(\mu,x,t),x\rangle \leq \gamma_1(\alpha) - \gamma_2(\alpha)|x|^{2p} \quad \forall\,(x,t) \in \mathbb{R}^d \times [0,\tau].$$

Suppose also that

$$|b(\mu,y,t) - b(\sigma,y,t)| \leq Ce^{K|y|^{2p}/2}\|(1+|x|^p)(\mu_t - \sigma_t)\|_{TV}.$$

Then, for every probability measure ν on $\mathbb{R}^d$ such that $e^{K|x|^{2p}} \in L^1(\nu)$, there exist $\tau \in (0,T]$ and $\alpha \in C^+[0,T]$ such that a solution to the Cauchy problem (4) in the class of measures $\mathcal{M}^{K,p}_{\tau,\alpha}$ exists and is unique.

Example

Let

$$b(\mu, x, t) = \beta(x, t) + \int_{\mathbb{R}^d} K(x, y)\, \mu_t(dy),$$

where $\beta\colon \mathbb{R}^d \times [0, T] \to \mathbb{R}^d$ and $K\colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ are Borel measurable locally bounded mappings such that there exist numbers $C > 0, 2p > q > 0, \gamma_1 > 0, \gamma_2 > 2pK$ for which

$$|K(x, y)| \leq C(1 + |x|^q)(1 + |y|^p), \quad \langle \beta(x, t), x\rangle \leq \gamma_1 - \gamma_2 |x|^{2p}.$$

Then all conditions of the above corollary are fulfilled.

Proof.
Let us define a mapping $F\colon \mathcal{M}^{K,p}_{\tau,\alpha} \to \mathcal{M}^{K,p}_{\tau,\alpha}$ by

$$\mu = F(\sigma) \Longleftrightarrow \partial_t\mu = \Delta\mu - \mathrm{div}(b(\sigma)\mu),\ \mu|_{t=0} = \nu.$$

It turns out that there exist $\tau > 0$ and α such that the mapping F is contracting. Indeed, we have

$$\|(1+|x|^p)(\mu^1_t - \mu^2_t)\|^2_{TV} \le \widetilde{C}\int_0^t\int_{\mathbb{R}^d} |b(\sigma^1) - b(\sigma^2)|^2\, d\sigma \le \widehat{C}\|\sigma^1 - \sigma^2\|^2_{p,\tau},$$

where $\widehat{C}$ does not depend on τ, but only on T. Integrating in t over $[0,\tau]$, we find that

$$\|F(\sigma^1) - F(\sigma^2)\|^2_{p,\tau} \le \tau\widehat{C}\|\sigma^1 - \sigma^2\|^2_{p,\tau}.$$

The next example demonstrates that uniqueness depends on the given metric on the space of measures and also depends on the regularity of the initial condition. Moreover the term $|\nabla \varrho_\sigma|^2/\varrho_\sigma$ in the right side of our estimate is essential.

Example

Let $d = 1$, $A = a(t, \mu)$, $b = 0$ and $\nu = \delta_0$. Set $\mu^1 = \mu_t^1\, dt$ and $\mu^2 = \mu_t^2\, dt$, where

$$\mu_t^1 = (2\pi t)^{-1/2} e^{-x^2/2t}\, dx, \quad \mu_t^2 = (8\pi t)^{-1/2} e^{-x^2/8t}\, dx.$$

Note that $\|\mu_t^1 - \mu_t^2\|_{TV} = c_0 > 0$ and c_0 does not depend on t. Let

$$a(t, \mu) = 1 + \frac{3}{c_0} \|\mu_t - \mu_t^1\|_{TV}.$$

We have $a(t, \mu^1) = 1$, $a(t, \mu^2) = 4$ and $|a(t, \mu) - a(t, \sigma)| \leq \frac{3}{c_0} \|\mu_t - \sigma_t\|_{TV}$. The measures μ^1 and μ^2 are two different solutions to the Cauchy problem with this coefficient a and $\nu = \delta_0$.

Optimal control

Our next application is concerned with optimal control. For a given bounded probability density σ on $\mathbb{R}^d$ and $\tau \in (0,1)$, we consider the problem of minimization of the function

$$J(u) = \frac{1}{2}\int_{\mathbb{R}^d} |\varrho(x,\tau,u) - \sigma(x)|^2\,dx + \frac{u^2}{2}, \quad u \in \mathbb{R}, \qquad (5)$$

in the class of probability densities $x \mapsto \varrho(x,t,u)$ on $\mathbb{R}^d$ such that $\varrho(x,t,u)$ solves the Cauchy problem on $[0,\tau]$ for the Fokker–Planck–Kolmogorov equation

$$\partial_t\varrho = \Delta\varrho - \mathrm{div}(b(\cdot,\cdot,u)\varrho), \quad \varrho|_{t=0} = \varrho_0.$$

(Annunziato M., Borzi A. 2010, 2013, Pardoux E., Veretennikov A.Yu. 2003)

We assume that the initial condition ϱ_0 is a bounded probability density with $\varrho_0 \ln(4+|x|) \in L^1(\mathbb{R}^d)$ and the drift b depending on the parameter u satisfies the inequality

$$|b(x,t,u)| + |\partial_u b(x,t,u)| + |\partial_u^2 b(x,t,u)| \leq M$$

for every $(x,t) \in \mathbb{R}^d \times [0,1],\ u \in \mathbb{R}$.

Corollary

There is $\tau > 0$ *such that J from* (5) *has a unique point of minimum.*

Proof.

Indeed, the function J is continuous and tends to $+\infty$ as $|u| \to +\infty$, which implies the existence of a point of minimum. The function ϱ is differentiable in u. Hence the function J is differentiable, and at the point of minimum

$$J'(u) = \int_{\mathbb{R}^d} (\varrho(x,\tau,u) - \sigma(x))\partial_u \varrho(x,\tau,u)\, dx + u = 0.$$

Let us consider the mapping $G\colon \mathbb{R} \to \mathbb{R}$ given by

$$G(u) = -\int_{\mathbb{R}^d} (\varrho(x,\tau,u) - \sigma(x))\partial_u \varrho(x,\tau,u)\, dx.$$

The points of minimum of J are fixed points of G. It turns out that for sufficiently small $\tau > 0$ the mapping G is contracting, which yields the uniqueness of a point of minimum.

Mean Field Games

We consider yet another possible application, which concerns the so-called mean field games. A typical model for mean field games is the system

$$\begin{cases} \partial_t u + \Delta u - H(x, \nabla u) = F(x, \mu_t), \\ \partial_t \mu_t - \Delta \mu_t + \operatorname{div}(b(x, \nabla u)\mu_t) = 0, \quad (x,t) \in \mathbb{R}^d \times (0,T), \end{cases} \tag{6}$$

with initial-terminal conditions $u(x, T) = G(x, \mu_T)$ and $\mu_t|_{t=0} = \nu$, where ν is a Borel probability measure, H, F, G are given functions and b is a vector field, usually $b(x, p) = \partial H(x, p)/\partial p$, but we do not assume this relation. (Gueant O., Lasry, J.M., Lions P.L. 2011, Gomes D.A., Saude J. 2014, Cardaliaguest P. 2013)

Let $\mathcal{P}(\mathbb{R}^d)$ be the space of all Borel probability measures on $\mathbb{R}^d$. Suppose that F and G are functions on $\mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d)$, H is a function on $\mathbb{R}^d \times \mathbb{R}^d$ and $b\colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ is a vector field such that

(C1) F and G are continuous, $G(x,\mu)$ is continuously differentiable in x and there exist numbers $L_0 > 0$, $L_1 > 0$ such that

$$|F(x,\mu)| + |G(x,\mu)| + |\nabla_x G(x,\mu)| \le L_0,$$

$$\begin{aligned}|F(x,\mu) - F(x,\sigma)| + |G(x,\mu) - G(x,\sigma)| \\ + |\nabla_x G(x,\mu) - \nabla_x G(x,\sigma)| \le L_1 \|\mu - \sigma\|_{TV}\end{aligned}$$

for all $x \in \mathbb{R}^d$, $\mu, \sigma \in \mathcal{P}(\mathbb{R}^d)$;

(C2) H and b are continuous and for every $R > 0$ there exist numbers $M_0(R) > 0$ and $M_1(R) > 0$ such that, whenever $x \in \mathbb{R}^d$ and $|p| \leq R$,

$$|H(x,p)| \leq M_0(R),$$

$$|H(x,p) - H(x,q)| + |b(x,p) - b(x,q)| \leq M_1(R)|p - q|.$$

A solution to (6) is a pair consisting of a mapping $u \in C([0,T], C_b^1(\mathbb{R}^d))$ and a flow of probability measures μ_t on $\mathbb{R}^d$ such that μ_t is a solution to the Cauchy problem

$$\partial_t \mu_t - \Delta \mu_t - \mathrm{div}(H_p(x, \nabla u)\mu_t) = 0, \quad \mu_t|_{t=0} = \nu,$$

and u satisfies the identity

$$u(x,t) = \int_{\mathbb{R}^d} Z(x-y, T-t) G(y, \mu_T)\, dy$$
$$+ \int_t^T \int_{\mathbb{R}^d} Z(x-y, \tau - t)\big(H(y, \nabla u(y,\tau)) + F(y, \mu_\tau)\big)\, dy\, d\tau,$$

where

$$Z(x,t) = (4\pi t)^{-d/2} \exp(-|x|^2/4t).$$

Corollary

There is $T > 0$ *such that* (6) *has a unique solution on* $[0, T]$.

Proof.

We apply the contracting mapping theorem to the mapping

$$\mathcal{F}\colon C([0,T], C_b^1(\mathbb{R}^d)) \to C([0,T], C_b^1(\mathbb{R}^d)),$$

where $C([0,T], C_b^1(\mathbb{R}^d))$ is equipped with its natural norm

$$\|v\| = \sup_{t\in[0,T]} \sup_{x} \big[|v(x,t)| + |\nabla_x v(x,t)| \big],$$

defined as follows:

for each $v \in C([0,T], C_b^1(\mathbb{R}^d))$ we find a solution μ_t (which is unique under our assumptions) to the Cauchy problem

$$\partial_t \mu_t - \Delta \mu_t - \mathrm{div}(b(x, \nabla v)\mu_t) = 0, \quad \mu_t|_{t=0} = \nu,$$

and set

$$\mathcal{F}(v) = \int_{\mathbb{R}^d} Z(x-y, T-t) G(y, \mu_T)\, dy$$
$$- \int_t^T \int_{\mathbb{R}^d} Z(x-y, \tau - t)\big(H(y, \nabla v(y,\tau)) + F(y, \mu_\tau)\big)\, dy\, d\tau.$$

Convergence to the stationary measure

Let us consider the Cauchy problem

$$\partial_t \mu_t = \Delta \mu_t - \mathrm{div}(b(x,\mu_t)\mu_t), \quad \mu_0 = \nu.$$

Assume that

$$b(x,\mu) = b_0(x) + \varepsilon b_1(x,\mu)$$

and

$$\langle b_0(x), x\rangle \le -\gamma |x|^2, \quad |b_1| \le M_1,$$

$$|b_1(x,\mu) - b_1(x,\sigma)| \le M_2 \|(1+|x|)(\mu-\sigma)\|_{TV}.$$

Let μ be a stationary measure.
Problem: to prove that $\mu_t \to \mu$ and to obtain the following estimate $\|(1+|x|)(\mu_t - \mu)\| \le \alpha_1 e^{-\alpha_2 t}$.
Note that it is often simpler, and in the case of a degenerate diffusion matrix more natural, to consider convergence in the Kantorovich metric. Results of this sort for non-gradient drift coefficients were apparently first obtained by N.U. Ahmed and X. Ding, and have been recently generalized by A. Eberle, A. Guillin, R. Zimmer, A. Yu. Veretennikov, F.-Y. Wang.
The gradient case, where $b = \nabla V$, has been studied in many papers, starting from D.A. Dawson, J. Gärtner, Y. Tamura and further studied in many papers on the theory of gradient flows by L. Ambrosio, N. Gigli, G. Savaré, F. Bolley, I. Gentil, A. Guillin, J.A. Carrillo, R.J. McCann, C. Villani, ...
Here we discuss the convergence in variation.
(O.A. Butkovsky, A. Eberle, ...)

It is known that

$$\|(1+|x|)(\mu-\sigma_t)\|_{TV} \leq \lambda_1 e^{-\lambda_2 t}\|(1+|x|)(\mu-\nu)\|_{TV},$$

where

$$\partial_t \sigma_t = \Delta\sigma_t - \mathrm{div}(b(x,\mu)\sigma_t), \quad \sigma_0 = \nu.$$

Let $t \in [0, T]$. According to the above estimate we have

$$\|(1+|x|)(\mu_t-\sigma_t)\|_{TV}^2 \leq C(T)\varepsilon \int_0^t \|(1+|x|)(\mu-\mu_s)\|_{TV}^2\, ds,\ t \in [0, T].$$

Then

$$
\begin{aligned}
&\|(1+|x|)(\mu-\mu_t)\|_{TV}^2 \leq \\
&\qquad 2\|(1+|x|)(\mu-\sigma_t)\|_{TV}^2 + 2\|(1+|x|)(\sigma_t-\mu_t)\|_{TV}^2 \leq \\
&\qquad\qquad \leq 2\lambda_1^2 e^{-2\lambda_2 t}\|(1+|x|)(\mu-\nu)\|_{TV}^2 + \\
&\qquad\qquad\qquad \varepsilon C(T)\int_0^t \|(1+|x|)(\mu-\mu_s)\|_{TV}^2\, ds.
\end{aligned}
$$

Apply the Gronwall inequality we obtain

$$\|(1+|x|)(\mu-\mu_t)\|_{TV}^2 \leq \|(1+|x|)(\mu-\nu)\|_{TV}^2 (2\lambda_1^2 e^{-2\lambda_2 t} + \varepsilon C(T)).$$

For sufficiently small ε we have

$$\|(1+|x|)(\mu-\mu_T)\|_{TV} \leq q\|(1+|x|)(\mu-\nu)\|_{TV}, \quad 0 < q < 1.$$

Thus

$$\|(1+|x|)(\mu-\mu_t)\|_{TV} \leq \alpha_1 e^{-\alpha_2 t}.$$

1. V.I.Bogachev, N.V.Krylov, M.Röckner, S.V.Shaposhnikov, Fokker–Planck–Kolmogorov Equations, Amer. Math. Soc., Rhode Island, Providence, 2015.
2. V.I.Bogachev, M.Röckner, S.V.Shaposhnikov, Distances between transition probabilities of diffusions and applications to nonlinear Fokker–Planck–Kolmogorov equations. Journal of Functional Analysis, 2016, V. 271, p. 1262–1300.
3. V.I.Bogachev, M.Röckner, S.V.Shaposhnikov, Convergence in variation of solutions of nonlinear Fokker–Planck–Kolmogorov equations to stationary measures. 2017, Preprint 17027 (CRC 2183) Bielefeld University.

Thank you!