Measure-valued continuous curves and processes in total variation norm

Jinghai Shao

Beijing Normal University

July 17, 2012

Theorem (Kolmogorov continuity theorem)

Let $X : [0, \infty) \times \Omega \to \mathbb{R}^n$ *be a stochastic process, and suppose for all time* $T > 0$, *there exist positive constants* α, β, K *such that*

$$\mathbb{E}\Big[|X_t - X_s|^{\alpha}\Big] \leq K|t - s|^{1+\beta}$$

for all $0 \leq s, t \leq T$. *Then there exists a continuous version of* X.

- A purely atomic measure-valued stochastic process has a continuous version in the total variation norm.
- geometric properties of the space of purely atomic probability measures.
- S., (2012), *Measure-valued continuous curves and processes in total variation norm*, J. Math. Analysis and Applications

Notations

Let (S, d) be a Polish space. Borel σ-algebra $\mathscr{B}(S)$.

- $\mathscr{M}(S)$: the set of all finite measures on S.
- $\mathscr{M}_a(S) \subset \mathscr{M}(S)$ contains all purely atomic ones. i.e.

$$\mathscr{M}_a(S) = \Big\{\mu = \sum_{i=1}^{\infty} p_i \delta_{x_i};\ p_i \geq 0,\ \sum_{i=1}^{\infty} p_i < \infty,\ x_i \in S\Big\}.$$

- $\mathscr{M}_1(S)$: probability measures on S.
- $\mathscr{M}_{1,a}(S)$ probability measures in $\mathscr{M}_a(S)$.

Total variation distance (norm):

$$\text{For } \mu,\ \nu \in \mathscr{M}(S),\ \|\mu - \nu\|_{\mathrm{Var}} = \sup_{B \in \mathscr{B}(S)} \big|\mu(B) - \nu(B)\big|.$$

Notations

Let (S, d) be a Polish space. Borel σ-algebra $\mathscr{B}(S)$.

- $\mathscr{M}(S)$: the set of all finite measures on S.
- $\mathscr{M}_a(S) \subset \mathscr{M}(S)$ contains all purely atomic ones. i.e.

$$\mathscr{M}_a(S) = \Big\{\mu = \sum_{i=1}^{\infty} p_i \delta_{x_i};\ p_i \geq 0,\ \sum_{i=1}^{\infty} p_i < \infty,\ x_i \in S\Big\}.$$

- $\mathscr{M}_1(S)$: probability measures on S.
- $\mathscr{M}_{1,a}(S)$ probability measures in $\mathscr{M}_a(S)$.

Total variation distance (norm):

$$\text{For } \mu,\ \nu \in \mathscr{M}(S),\ \|\mu - \nu\|_{\mathrm{Var}} = 2 \sup_{B \in \mathscr{B}(S)} \big|\mu(B) - \nu(B)\big|.$$

Proposition

(i) $(\mathscr{M}_{1,a}(S), \|\cdot\|_{\mathrm{Var}})$ *is separable iff* S *is countable.*

(ii) $(\mathscr{M}_{1,a}(S), \|\cdot\|_{\mathrm{Var}})$ *is compact iff* S *contains only finite number of points.*

Proposition

(i) $(\mathscr{M}_{1,a}(S), \|\cdot\|_{\mathrm{Var}})$ *is separable iff S is countable.*

(ii) $(\mathscr{M}_{1,a}(S), \|\cdot\|_{\mathrm{Var}})$ *is compact iff S contains only finite number of points.*

PROOF: Suppose S is uncountable, then for every countable subset of $\mathbb{D}$ of $\mathscr{M}_{1,a}(S)$, (since the totality of support points of all prob. in $\mathbb{D}$ is at most countable,) we can find a point $x \in S$, which doesn't belong to the support of any $\nu \in \mathbb{D}$. Put $\mu = \delta_x$, then for any $\nu \in \mathbb{D}$,

$$\|\mu - \nu\|_{\mathrm{Var}} = \sup_{B \in \mathscr{B}(S)} \big|\mu(B) - \nu(B)\big| \geq 1 - \nu(\{x\}) = 1.$$

Therefore, when S is uncountable, $(\mathscr{M}_{1,a}(S), \|\cdot\|_{\mathrm{Var}})$ is not separable.

Completeness

Proposition

$(\mathscr{M}_{1,a}(S), \|\cdot\|_{\mathrm{Var}})$ *is complete.*

In particular, a sequence of $(\mu_n)_n$ in $\mathscr{M}_{1,a}(S)$ converges in $\|\cdot\|_{\mathrm{Var}}$ to some $\mu \in \mathscr{M}_1(S)$, then μ is also in $\mathscr{M}_{1,a}(S)$.

$\mathscr{M}_{1,a}(S)$ is not complete under the weak topology.

Geodesic space

Proposition

$(\mathscr{M}_{1,a}(S), \|\cdot\|_{\mathrm{Var}})$ *is a geodesic space. Namely, for each pair of* $\mu, \nu \in \mathscr{M}_{1,a}(S)$, *there is a curve* $t \mapsto \mu_t$ *from [0,1] to* $\mathscr{M}_{1,a}(S)$ *s.t.* $\mu_0 = \mu$, $\mu_1 = \nu$ *and*

$$\|\mu_t - \mu_s\|_{\mathrm{Var}} = |t-s|\|\mu - \nu\|_{\mathrm{Var}}, \quad s, t \in [0,1].$$

For $\mu, \nu \in \mathscr{M}_{1,a}(S)$, it's said $\mu \ll \nu$ if $\nu(\{x\}) = 0$ yields $\mu(\{x\}) = 0$.
For $\mu \ll \nu$ in the form $\mu = \sum_i p_i \delta_{x_i}$, $\nu = \sum_i q_i \delta_{x_i}$ with $p_i \geq 0$ and $q_i > 0$, define

$$\mathrm{Ent}(\mu|\nu) = \sum_i p_i \log \frac{p_i}{q_i}, \quad 0 \log 0 := 0.$$

Proposition

Let $\mu, \nu \in \mathscr{M}_{1,a}(S)$, and $\mu \ll \nu$. Let $(\mu)_{t\in[0,1]}$ be a geodesic connecting μ to ν. Then

$$\mathrm{Ent}(\mu_t|\nu) \leq (1-t)\mathrm{Ent}(\mu|\nu) - 2t(1-t)\|\mu - \nu\|^2_{\mathrm{Var}}, \quad t \in [0,1]. \qquad (1)$$

- Lott-Villani-Sturm generalization of lower bound of Ricci curvature.

For μ, $\nu \in \mathscr{M}_{1,a}(S)$, it's said $\mu \ll \nu$ if $\nu(\{x\}) = 0$ yields $\mu(\{x\}) = 0$.
For $\mu \ll \nu$ in the form $\mu = \sum_i p_i \delta_{x_i}$, $\nu = \sum_i q_i \delta_{x_i}$ with $p_i \geq 0$ and $q_i > 0$, define

$$\mathrm{Ent}(\mu|\nu) = \sum_i p_i \log \frac{p_i}{q_i}, \quad 0 \log 0 := 0.$$

Proposition

Let μ, $\nu \in \mathscr{M}_{1,a}(S)$, and $\mu \ll \nu$. Let $(\mu)_{t\in[0,1]}$ be a geodesic connecting μ to ν. Then

$$\mathrm{Ent}(\mu_t|\nu) \leq (1-t)\mathrm{Ent}(\mu|\nu) - 2t(1-t)\|\mu - \nu\|_{\mathrm{Var}}^2, \quad t \in [0,1]. \tag{1}$$

- Lott-Villani-Sturm generalization of lower bound of Ricci curvature.

Pinsker's inequality

Corollary

Let $\mu, \nu \in \mathscr{M}_{1,a}(S)$ *and* $\mu \ll \nu$, *then*

$$2\|\mu - \nu\|_{\mathrm{Var}}^2 \leq \mathrm{Ent}(\mu|\nu). \tag{2}$$

Proof: If (2) does not hold, i.e. $\mathrm{Ent}(\mu|\nu) < 2\|\mu - \nu\|_{\mathrm{Var}}^2$, then $\exists\, \varepsilon > 0$ such that $\mathrm{Ent}(\mu|\nu) < 2\|\mu - \nu\|_{\mathrm{Var}}^2 - \varepsilon$. Substituting it into (1),

$$0 < \mathrm{Ent}(\mu_t|\nu) < (1-t)\big(2(1-t)\|\mu - \nu\|_{\mathrm{Var}}^2 - \varepsilon\big) < -(1-t)\varepsilon/2 < 0$$

as t close enough to 1 s.t. $2(1-t)\|\mu - \nu\|_{\mathrm{Var}}^2 < \varepsilon/2$.

Continuous curves in total variation norm

CONCLUSION: Given a curve $\mu : [0,\infty) \to \mathcal{M}_a(S)$, the following assertions are equivalent:

(i) $(\mu_t)_{t\geq 0}$ is continuous in total variation norm.

(ii) For every $B \in \mathscr{B}(S)$, $t \mapsto \mu_t(B)$ is continuous.

(iii) For every $x \in S$, $t \mapsto \mu_t(\{x\})$ is cont., and $t \mapsto \mu_t(S)$ is cont..

(iv) For every $B \in \mathscr{B}(S)$, the curves $t \mapsto \sup_{x\in B} \mu_t(\{x\})$ and $t \mapsto \mu_t(S)$ are cont..

Continuous curves in total variation norm

CONCLUSION: Given a curve $\mu : [0,\infty) \to \mathcal{M}_a(S)$, the following assertions are equivalent:

(i) $(\mu_t)_{t\geq 0}$ is continuous in total variation norm.

(ii) For every $B \in \mathscr{B}(S)$, $t \mapsto \mu_t(B)$ is continuous.

(iii) For every $x \in S$, $t \mapsto \mu_t(\{x\})$ is cont., and $t \mapsto \mu_t(S)$ is cont..

(iv) For every $B \in \mathscr{B}(S)$, the curves $t \mapsto \sup_{x\in B} \mu_t(\{x\})$ and $t \mapsto \mu_t(S)$ are cont..

Continuous curves in total variation norm

CONCLUSION: Given a curve $\mu : [0, \infty) \to \mathcal{M}_a(S)$, the following assertions are equivalent:

(i) $(\mu_t)_{t\geq 0}$ is continuous in total variation norm.

(ii) For every $B \in \mathscr{B}(S)$, $t \mapsto \mu_t(B)$ is continuous.

(iii) For every $x \in S$, $t \mapsto \mu_t(\{x\})$ is cont., and $t \mapsto \mu_t(S)$ is cont..

(iv) For every $B \in \mathscr{B}(S)$, the curves $t \mapsto \sup_{x\in B} \mu_t(\{x\})$ and $t \mapsto \mu_t(S)$ are cont..

CONCLUSION: Given a curve $\mu : [0,\infty) \to \mathcal{M}_a(S)$, the following assertions are equivalent:

(i) $(\mu_t)_{t\geq 0}$ is continuous in total variation norm.

(ii) For every $B \in \mathscr{B}(S)$, $t \mapsto \mu_t(B)$ is continuous.

(iii) For every $x \in S$, $t \mapsto \mu_t(\{x\})$ is cont., and $t \mapsto \mu_t(S)$ is cont..

(iv) For every $B \in \mathscr{B}(S)$, the curves $t \mapsto \sup_{x\in B} \mu_t(\{x\})$ and $t \mapsto \mu_t(S)$ are cont..

Continuous curves in total variation norm

CONCLUSION: Given a curve $\mu : [0, \infty) \to \mathcal{M}_a(S)$, the following assertions are equivalent:

(i) $(\mu_t)_{t\geq 0}$ is continuous in total variation norm.

(ii) For every $B \in \mathscr{B}(S)$, $t \mapsto \mu_t(B)$ is continuous.

(iii) For every $x \in S$, $t \mapsto \mu_t(\{x\})$ is cont., and $t \mapsto \mu_t(S)$ is cont..

(iv) For every $B \in \mathscr{B}(S)$, the curves $t \mapsto \sup_{x\in B} \mu_t(\{x\})$ and $t \mapsto \mu_t(S)$ are cont..

(i) $\Longrightarrow$ (ii) $\Longrightarrow$ (iii), and (iv) $\Longrightarrow$ (iii), are clear. To show (iii) $\Longrightarrow$ (i) and (i) $\Longrightarrow$ (iv) needs some trick.

Continuous curves in total variation norm

If assume further $\exists$ a positive measure m on S s.t. $m\big(B(x,r)\big) > 0$, and $m(S(x,r)) = 0$, $\forall\, x \in S,\ r > 0$. Then above assertions are also equivalent to

(v) $\exists$ dense subset $(x_i)_{i\in\mathbb{N}} \subset S$ and dense subset $(r_j)_{j\in\mathbb{N}} \subset [0,\infty)$ s.t. for every $t \geq 0$

1. $\mu_t\big(\bigcup_{ij} S_{ij}\big) = 0,\ i,\, j \in \mathbb{N}$, where $S_{ij} = \big\{y \in S;\ d(x_i, y) = r_j\big\}$;
2. $t \mapsto \mu_t$ is weakly continuous;
3. $t \mapsto \sup_{x\in B_{ij}} \mu_t(\{x\})$ is cont. where $B_{ij} = \big\{y \in S;\ d(x_i, y) < r_j\big\}$.

Let's introduce a family of disjoint partition $\mathscr{A} = \big(A_j^n\big)_{j,n=1}^{\infty}$ of S.

1. for every fixed n, $(A_j^n)_{j=1\infty}$ is a disjoint partitions of S.
2. $\mathrm{diam}(A_j^n) \leq 1/n$, for $j \geq 1$.
3. $(A_j^{n+1})_{j=1}^{\infty}$ is a refinement of $(A_j^n)_{j=1}^{\infty}$.

With this $\mathscr{A}$,

$$\inf_n \sup_j \mu(A_j^n \bigcap B) = \sup_{x\in B} \mu\big(\{x\}\big), \quad B \in \mathscr{B}(S).$$

L. Overbeck, M. Röckner, B. Schmuland, *An analytic approach to Fleming-Viot processes with interactive selection*. Ann. Probab. 1995.

Kolmogorov's type continuity theorem

Theorem: Let $(X_t)_{t\geq 0}$ be a $\mathcal{M}_a(S)$-valued process, continuous in weak topology. Suppose $\exists$ dense subsets $\{x_i\}_{i\geq 1}\subset S$ and $\{r_j\}_{j\geq 1}\subset [0,\infty)$ so that

$$\forall\, t\geq 0,\quad X_t\big(\bigcup_{ij} S_{ij}\big)=0,\quad S_{ij}=\{y\in S;\, d(x_i,y)=r_j\}.$$

If $\exists$ positive const. α, β, $C_L, (L=1,2,\ldots)$ such that

$$\sum_{j=1}^{\infty}\mathbb{E}\Big|X_t\big(A_j^n\bigcap B_{k\ell}\big)-X_s\big(A_j^n\bigcap B_{k\ell}\big)\Big|^{\alpha}\leq C_L|t-s|^{1+\beta},$$

$\forall\, n,\, k,\ell\in\mathbb{N}$ and $t,\, s\in[0,L]$ $(L=1,2,\ldots)$. Then $(X_t)_{t\geq 0}$ has a version $(\hat{X}_t)_{t\geq 0}$ owning cont. pathes in $\|\cdot\|_{\mathrm{Var}}$.

An example of its application

T. Shiga (1990): A stochastic equation based on a Poisson system for a class of measure-valued diffusion processes. J. Math. Kyoto Univ. 30, 245-279. constructed measure-valued branching diffusions with immigrations generated by operator $\mathcal{L}$,

$$\mathcal{L}F(\mu) = \frac{1}{2}\int_S \mu(\mathrm{d}x)\beta(x)\frac{\delta^2 F(\mu)}{\delta\mu(x)^2} + \int_S \big(\mu(\mathrm{d}x)\gamma(x) + V(\mathrm{d}x)\big)\frac{\delta F(\mu)}{\delta\mu(x)},$$

Theorem: Under suitable condition on μ and V, the $\mathcal{M}_a(S)$-valued branching diffusion $(X_t)_{t>0}$ corresponding to operator $\mathcal{L}$ admits a continuous version w.r.t. $\|\cdot\|_{\mathrm{Var}}$.

An example of its application

T. Shiga (1990): constructed measure-valued branching diffusions with immigrations generated by operator $\mathcal{L}$,

$$\mathcal{L}F(\mu) = \frac{1}{2}\int_S \mu(\mathrm{d}x)\beta(x)\frac{\delta^2 F(\mu)}{\delta\mu(x)^2} + \int_S \big(\mu(\mathrm{d}x)\gamma(x) + V(\mathrm{d}x)\big)\frac{\delta F(\mu)}{\delta\mu(x)},$$

Theorem: Under suitable condition on μ and V, the $\mathcal{M}_a(S)$-valued branching diffusion $(X_t)_{t>0}$ corresponding to operator $\mathcal{L}$ admits a continuous version w.r.t. $\|\cdot\|_{\mathrm{Var}}$.

THE END

Thank you for your attention!