Evolution of the Wasserstein distance between the marginals of two Markov processes

Jacopo Corbetta (École des Ponts - ParisTech)

Joint work with: Aurélien Alfonsi and Benjamin Jourdain

January 10, 2017

The Wasserstein distance

Definition

The $\varrho-$Wasserstein distance between two probability measures P and $\widetilde{P}$ on $\mathbb{R}^d$ is given by

$$W_\varrho(P, \widetilde{P}) = \left(\inf_{\pi \in \Pi(P, \widetilde{P})} \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^\varrho \, \pi(\mathrm{d}x, \mathrm{d}y) \right)^{\frac{1}{\varrho}}$$

The Wasserstein distance

Definition

The $\varrho-$Wasserstein distance between two probability measures P and $\widetilde{P}$ on $\mathbb{R}^d$ is given by

$$W_\varrho(P,\widetilde{P}) = \left(\inf_{\pi \in \Pi(P,\widetilde{P})} \int_{\mathbb{R}^d \times \mathbb{R}^d} |x-y|^\varrho \, \pi(\mathrm{d}x,\mathrm{d}y) \right)^{\frac{1}{\varrho}}$$

Dual Representation

$$W_\varrho^\varrho(P,\widetilde{P}) = \sup \left\{ -\int_{\mathbb{R}^d} \phi(x) P(\mathrm{d}x) - \int_{\mathbb{R}^d} \widetilde{\phi}(y) \widetilde{P}(\mathrm{d}y) \right\}$$

A couple $(\psi,\widetilde{\psi})$ obtaining the sup is called Kantorovich potentials.

A generic heuristic formula

Let $\{X_t\}_{t\geq 0}$ and $\{\widetilde{X}_t\}_{t\geq 0}$ be two $\mathbb{R}^d$-valued Markov processes. Then

$$W_\varrho^\varrho(P_t, \widetilde{P}_t) = -\int_{\mathbb{R}^d} \psi_t(x) P_t(\mathrm{d}x) - \int_{\mathbb{R}^d} \widetilde{\psi}_t(y) \widetilde{P}_t(\mathrm{d}y)$$

A generic heuristic formula

Let $\{X_t\}_{t\geq 0}$ and $\{\widetilde{X}_t\}_{t\geq 0}$ be two $\mathbb{R}^d$-valued Markov processes. Then

$$W_\varrho^\varrho(P_t,\widetilde{P}_t) = -\int_{\mathbb{R}^d}\psi_t(x)P_t(\mathrm{d}x) - \int_{\mathbb{R}^d}\widetilde{\psi}_t(y)\widetilde{P}_t(\mathrm{d}y)$$

For all $0\leq t$

$$\frac{\mathrm{d}}{\mathrm{d}t}W_\varrho^\varrho(P_t,\widetilde{P}_t) = -\int_{\mathbb{R}^d}L\psi_t(x)P_t(\mathrm{d}x) - \int_{\mathbb{R}^d}\widetilde{L}\,\widetilde{\psi}_t(x)\widetilde{P}_t(\mathrm{d}x)\,.$$

A generic heuristic formula

Let $\{X_t\}_{t\geq 0}$ and $\{\widetilde{X}_t\}_{t\geq 0}$ be two $\mathbb{R}^d$-valued Markov processes. Then

$$W_\varrho^\varrho(P_t,\widetilde{P}_t) = -\int_{\mathbb{R}^d}\psi_t(x)P_t(\mathrm{d}x) - \int_{\mathbb{R}^d}\widetilde{\psi}_t(y)\widetilde{P}_t(\mathrm{d}y)$$

For all $0 \leq t$

$$\frac{\mathrm{d}}{\mathrm{d}t}W_\varrho^\varrho(P_t,\widetilde{P}_t) = -\int_{\mathbb{R}^d}L\psi_t(x)P_t(\mathrm{d}x) - \int_{\mathbb{R}^d}\widetilde{L}\,\widetilde{\psi}_t(x)\widetilde{P}_t(\mathrm{d}x)\,.$$

Integral formulation: for all $0 \leq s \leq t$

$$W_\varrho^\varrho(P_t,\widetilde{P}_t) - W_\varrho^\varrho(P_s,\widetilde{P}_s) = \\ -\int_s^t\left[\int_{\mathbb{R}^d}L\psi_r(x)P_r(\mathrm{d}x) + \int_{\mathbb{R}^d}\widetilde{L}\,\widetilde{\psi}_r(x)\widetilde{P}_r(\mathrm{d}x)\right]\mathrm{d}r\,.$$

Formal proof

For every $s, t \geq 0$

$$W_{\varrho}^{\varrho}(P_s, \widetilde{P}_s) \geq -\int_{\mathbb{R}^d} \psi_t(x)\, P_s(\mathrm{d}x) - \int_{\mathbb{R}^d} \widetilde{\psi}_t(x)\, \widetilde{P}_s(\mathrm{d}x)\,.$$

Formal proof

For every $s, t \geq 0$

$$W_\varrho^\varrho(P_s, \widetilde{P}_s) \geq -\int_{\mathbb{R}^d} \psi_t(x)\, P_s(\mathrm{d}x) - \int_{\mathbb{R}^d} \widetilde{\psi}_t(x)\, \widetilde{P}_s(\mathrm{d}x)\,.$$

In particular

$$\begin{aligned}\int_{\mathbb{R}^d} \psi_s(x)(P_s(\mathrm{d}x) - P_t(\mathrm{d}x)) + \int_{\mathbb{R}^d} \widetilde{\psi}_s(x)(\widetilde{P}_s(\mathrm{d}x) - \widetilde{P}_t(\mathrm{d}x)) \\ \leq W_\varrho^\varrho(P_t, \widetilde{P}_t) - W_\varrho^\varrho(P_s, \widetilde{P}_s) \\ \leq \int_{\mathbb{R}^d} \psi_t(x)(P_s(\mathrm{d}x) - P_t(\mathrm{d}x)) + \int_{\mathbb{R}^d} \widetilde{\psi}_t(x)(\widetilde{P}_s(\mathrm{d}x) - \widetilde{P}_t(\mathrm{d}x))\,.\end{aligned}$$

Formal proof (2)

$$\int_{\mathbb{R}^d} \psi_t(x)(P_s(\mathrm{d}x) - P_t(\mathrm{d}x)) = -\int_s^t \int_{\mathbb{R}^d} L\psi_t(x) P_r(\mathrm{d}x)\mathrm{d}r$$

Formal proof (2)

$$\int_{\mathbb{R}^d} \psi_t(x)(P_s(\mathrm{d}x) - P_t(\mathrm{d}x)) = -\int_s^t \int_{\mathbb{R}^d} L\psi_t(x) P_r(\mathrm{d}x)\mathrm{d}r$$

$$\frac{1}{h}\Big(W_\varrho^\varrho(P_{t+h}, \widetilde{P}_{t+h}) - W_\varrho^\varrho(P_t, \widetilde{P}_t)\Big) \geq$$

$$\geq \frac{1}{h}\left(-\int_t^{t+h}\int_{\mathbb{R}^d} L\psi_t(x)P_r(\mathrm{d}x)\mathrm{d}r - \int_t^{t+h}\int_{\mathbb{R}^d} \widetilde{L}\widetilde{\psi}_t(x)P_r(\mathrm{d}x)\mathrm{d}r\right)$$

Formal proof (2)

$$\int_{\mathbb{R}^d} \psi_t(x)(P_s(\mathrm{d}x) - P_t(\mathrm{d}x)) = -\int_s^t \int_{\mathbb{R}^d} L\psi_t(x) P_r(\mathrm{d}x)\mathrm{d}r$$

$$\frac{1}{h}\Big(W_\varrho^\varrho(P_{t+h}, \widetilde{P}_{t+h}) - W_\varrho^\varrho(P_t, \widetilde{P}_t)\Big) \geq$$

$$\geq \frac{1}{h}\left(-\int_t^{t+h}\int_{\mathbb{R}^d} L\psi_t(x)P_r(\mathrm{d}x)\mathrm{d}r - \int_t^{t+h}\int_{\mathbb{R}^d} \widetilde{L}\widetilde{\psi}_t(x)P_r(\mathrm{d}x)\mathrm{d}r\right)$$

Taking the limit for $h \to 0^+$

$$\frac{\mathrm{d}}{\mathrm{d}t^+} W_\varrho^\varrho(P_t, \widetilde{P}_t) \geq -\int_{\mathbb{R}^d} L\psi_t(x)P_t(\mathrm{d}x) - \int_{\mathbb{R}^d} \widetilde{L}\,\widetilde{\psi}_t(x)\widetilde{P}_t(\mathrm{d}x)$$

In the same way:

$$\frac{\mathrm{d}}{\mathrm{d}t^-} W_\varrho^\varrho(P_t, \widetilde{P}_t) \leq -\int_{\mathbb{R}^d} L\psi_t(x)P_t(\mathrm{d}x) - \int_{\mathbb{R}^d} \widetilde{L}\,\widetilde{\psi}_t(x)\widetilde{P}_t(\mathrm{d}x)\,.$$

Main Issues

Technical problems:

- ψ_t, $L\psi_t$ integrability with respect to P_s.
- $r \mapsto W_\varrho^\varrho(P_r, \widetilde{P}_r)$ differentiability.

Pure jump: $Lf(x) = \lambda(x)\left(\int_{\mathbb{R}^d} k(x, \mathrm{d}y)\,(f(y) - f(x))\right)$

Theorem

Assume that

- $\sup_{x\in\mathbb{R}^d} \max(\lambda(x), \widetilde{\lambda}(x)) < \infty$
- $t \mapsto \mathrm{E}[|X_t|^{\varrho(1+\varepsilon)} + |\widetilde{X}_t|^{\varrho(1+\varepsilon)}]$ is locally bounded.

Then

- $t \mapsto \int_{\mathbb{R}^d} |L\psi_t(x)| P_t(\mathrm{d}x) + \int_{\mathbb{R}^d} |\widetilde{L}\,\widetilde{\psi}_t(x)| \widetilde{P}_t(\mathrm{d}x)$ is locally bounded
- $t \mapsto W_\varrho^\varrho(P_t, \widetilde{P}_t)$ is locally Lipschitz on $(0, +\infty)$ and for almost every $t \in (0, \infty)$

$$\frac{\mathrm{d}}{\mathrm{d}t} W_\varrho^\varrho(P_t, \widetilde{P}_t) = -\int_{\mathbb{R}^d} L\psi_t(x) P_t(\mathrm{d}x) - \int_{\mathbb{R}^d} \widetilde{L}\,\widetilde{\psi}_t(x) \widetilde{P}_t(\mathrm{d}x)\,.$$

- for every $t \geq 0$ the integral formula holds true.

Piecewise Deterministic Markov Processes

$$Lf(x) = V(x)\nabla f(x) + \lambda(x)\left(\int_{\mathbb{R}^d} k(x, \mathrm{d}y)\,(f(y) - f(x))\right).$$

The result still holds true

Piecewise Deterministic Markov Processes

$$Lf(x) = V(x)\nabla f(x) + \lambda(x)\left(\int_{\mathbb{R}^{\cancel{d}}} k(x, \mathrm{d}y)\,(f(y) - f(x))\right).$$

The result still holds true but:

- we have to restrict on the real line;
- more regularity on the marginals is required.

Thank you for your attention