archives-ouvertes.fr

Transport equation with source and generalized Wasserstein distance

Benedetto Piccoli

To cite this version:

Benedetto Piccoli. Transport equation with source and generalized Wasserstein distance. NETCO 2014 - New Trends in Optimal Control, Jun 2014, Tours, France. <hal-01031203>

HAL Id: hal-01031203
https://hal.inria.fr/hal-01031203

Submitted on 22 Jul 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Transport equation with source and generalized Wasserstein distance

Benedetto Piccoli

Rutgers University - Camden

joint work with
Francesco Rossi

NetCo 2014
Tours, June 23rd 2014

Outline

Optimal transport

Given μ measure and measurable $\gamma : \mathbb{R}^n \to \mathbb{R}^n$:

$$(\gamma \# \mu)(E) = \mu(\gamma^{-1}(E)).$$

Given μ, ν probability measures, look for γ such that $\gamma \# \mu = \nu$ and minimizes the cost

$$\int c(x, \gamma(x))\, d\mu.$$

Optimal transport first proposed by Monge problem in 1781.
Particular case is given by the cost $c(x, y) = |x - y|^p$ with $p \geq 1$, defining the Wasserstein distance:

$$W_p(\mu, \nu) = \inf_{\gamma \# \mu = \nu} \left(\int_{\mathbb{R}^n} |\gamma(x) - x|^p \, d\mu \right)^{1/p}.$$

Optimal transport

Given μ measure and measurable $\gamma : \mathbb{R}^n \to \mathbb{R}^n$:

$$(\gamma\#\mu)(E) = \mu(\gamma^{-1}(E)).$$

Given μ, ν probability measures, look for γ such that $\gamma\#\mu = \nu$ and minimizes the cost

$$\int c(x, \gamma(x))\, d\mu.$$

Optimal transport first proposed by Monge problem in 1781. Particular case is given by the cost $c(x, y) = |x - y|^p$ with $p \geq 1$, defining the Wasserstein distance:

$$W_p(\mu, \nu) = \inf_{\gamma\#\mu=\nu} \left(\int_{\mathbb{R}^n} |\gamma(x) - x|^p \, d\mu \right)^{1/p}.$$

Optimal transport

Given μ measure and measurable $\gamma : \mathbb{R}^n \to \mathbb{R}^n$:

$$(\gamma\#\mu)(E) = \mu(\gamma^{-1}(E)).$$

Given μ, ν probability measures, look for γ such that $\gamma\#\mu = \nu$ and minimizes the cost

$$\int c(x, \gamma(x))\, d\mu.$$

Optimal transport first proposed by Monge problem in 1781.
Particular case is given by the cost $c(x, y) = |x - y|^p$ with $p \geq 1$, defining the Wasserstein distance:

$$W_p(\mu, \nu) = \inf_{\gamma\#\mu=\nu} \left(\int_{\mathbb{R}^n} |\gamma(x) - x|^p \, d\mu \right)^{1/p}.$$

Wasserstein distance

Take $\mu = \delta_1$, $\nu = \frac{1}{2}\delta_0 + \frac{1}{2}\delta_2$: but there exists no γ with $\nu = \gamma\#\mu$, since γ cannot separate masses.

A probability measure π on $\mathbb{R}^d \times \mathbb{R}^d$, one can interpret it as a method to transfer a measure. μ is sent to ν if:

$$|\mu| \int_{\mathbb{R}^d} d\pi(x,\cdot) = d\mu(x), \qquad |\nu| \int_{\mathbb{R}^d} d\pi(\cdot,y) = d\nu(y).$$

Monge-Kantorovich problem. Wasserstein distance:

$$W_p(\mu,\nu) = \left(|\mu| \min_{\pi\in\Pi(\mu,\nu)} \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^p \, d\pi(x,y) \right)^{1/p}.$$

where $\Pi(\mu,\nu)$ is the set of transference plans from μ to ν.

Wasserstein distance

Take $\mu = \delta_1$, $\nu = \frac{1}{2}\delta_0 + \frac{1}{2}\delta_2$: but there exists no γ with $\nu = \gamma\#\mu$, since γ cannot separate masses.

A probability measure π on $\mathbb{R}^d \times \mathbb{R}^d$, one can interpret it as a method to transfer a measure. μ is sent to ν if:

$$|\mu| \int_{\mathbb{R}^d} d\pi(x,\cdot) = d\mu(x), \qquad |\nu| \int_{\mathbb{R}^d} d\pi(\cdot,y) = d\nu(y).$$

Monge-Kantorovich problem. Wasserstein distance:

$$W_p(\mu,\nu) = \left(|\mu| \min_{\pi\in\Pi(\mu,\nu)} \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^p \, d\pi(x,y) \right)^{1/p}.$$

where $\Pi(\mu,\nu)$ is the set of transference plans from μ to ν.

Wasserstein distance

Take $\mu = \delta_1$, $\nu = \frac{1}{2}\delta_0 + \frac{1}{2}\delta_2$: but there exists no γ with $\nu = \gamma \# \mu$, since γ cannot separate masses.

A probability measure π on $\mathbb{R}^d \times \mathbb{R}^d$, one can interpret it as a method to transfer a measure. μ is sent to ν if:

$$|\mu| \int_{\mathbb{R}^d} d\pi(x, \cdot) = d\mu(x), \qquad |\nu| \int_{\mathbb{R}^d} d\pi(\cdot, y) = d\nu(y).$$

Monge-Kantorovich problem. Wasserstein distance:

$$W_p(\mu, \nu) = \left(|\mu| \min_{\pi \in \Pi(\mu,\nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^p \, d\pi(x, y) \right)^{1/p}.$$

where $\Pi(\mu, \nu)$ is the set of transference plans from μ to ν.

The function $v : \mathcal{M} \to C^1(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$ satisfies

- $v[\mu]$ is uniformly Lipschitz and uniformly bounded, i.e. there exist L, M not depending on μ, such that for all $\mu \in \mathcal{M}, x, y \in \mathbb{R}^n$,

$$|v[\mu](x) - v[\mu](y)| \leq L|x-y| \qquad |v[\mu](x)| \leq M.$$

- v is a Lipschitz function, i.e. there exists K such that

$$\|v[\mu] - v[\nu]\|_{C^0} \leq KW_p(\mu, \nu).$$

Under such assumption there exists a unique solution to the transport equation:

$$\begin{cases} \partial_t \mu + \nabla \cdot (v[\mu]\, \mu) = 0 \\ \mu(0) = \mu_0 \end{cases}.$$

Replacing W with L^1 (total variation) we loose uniqueness.

The function $v : \mathcal{M} \to C^1(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$ satisfies

- $v[\mu]$ is uniformly Lipschitz and uniformly bounded, i.e. there exist L, M not depending on μ, such that for all $\mu \in \mathcal{M}, x, y \in \mathbb{R}^n$,

$$|v[\mu](x) - v[\mu](y)| \leq L|x-y| \qquad |v[\mu](x)| \leq M.$$

- v is a Lipschitz function, i.e. there exists K such that

$$\|v[\mu] - v[\nu]\|_{C^0} \leq KW_p(\mu, \nu).$$

Under such assumption there exists a unique solution to the transport equation:

$$\begin{cases} \partial_t \mu + \nabla \cdot (v[\mu]\ \mu) = 0 \\ \mu(0) = \mu_0 \end{cases}.$$

Replacing W with L^1 (total variation) we loose uniqueness.

The function $v : \mathcal{M} \to C^1(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$ satisfies

- $v[\mu]$ is uniformly Lipschitz and uniformly bounded, i.e. there exist L, M not depending on μ, such that for all $\mu \in \mathcal{M}, x, y \in \mathbb{R}^n$,

$$|v[\mu](x) - v[\mu](y)| \leq L|x - y| \qquad |v[\mu](x)| \leq M.$$

- v is a Lipschitz function, i.e. there exists K such that

$$\|v[\mu] - v[\nu]\|_{C^0} \leq KW_p(\mu, \nu).$$

Under such assumption there exists a unique solution to the transport equation:

$$\begin{cases} \partial_t \mu + \nabla \cdot (v[\mu]\ \mu) = 0 \\ \mu(0) = \mu_0 \end{cases}.$$

Replacing W with L^1 (total variation) we loose uniqueness.

Generalized Wasserstein distance

Three different admissible actions on μ, ν: either add/remove mass to μ, or add/remove mass to ν or transport mass from μ to ν.

$\mathcal{M}$ space of Borel measures with finite mass.
Given $a, b \in (0, \infty)$ and $p \geq 1$, the generalized Wasserstein distance is given by

$$W_P^{a,b}(\mu, \nu) = \inf_{\substack{\tilde{\mu}, \tilde{\nu} \in \mathcal{M}^p \\ |\tilde{\mu}| = |\tilde{\nu}|}} \left(a^p |\mu - \tilde{\mu}| + a|\nu - \tilde{\nu}|^p + b^p W_p(\tilde{\mu}, \tilde{\nu}) \right)^{1/p}.$$

Proposition

The operator $W_p^{a,b}$ is a distance. Moreover, one can restrict the computation to $\tilde{\mu} \leq \mu, \tilde{\nu} \leq \nu$ and the infimum is always attained.

Generalized Wasserstein distance

Three different admissible actions on μ, ν: either add/remove mass to μ, or add/remove mass to ν or transport mass from μ to ν.

$\mathcal{M}$ space of Borel measures with finite mass.
Given $a, b \in (0, \infty)$ and $p \geq 1$, the generalized Wasserstein distance is given by

$$W_P^{a,b}(\mu,\nu) = \inf_{\substack{\tilde{\mu},\tilde{\nu} \in \mathcal{M}^p \\ |\tilde{\mu}| = |\tilde{\nu}|}} \left(a^p|\mu - \tilde{\mu}| + a|\nu - \tilde{\nu}|^p + b^p W_p(\tilde{\mu},\tilde{\nu})\right)^{1/p}.$$

Proposition

The operator $W_p^{a,b}$ is a distance. Moreover, one can restrict the computation to $\tilde{\mu} \leq \mu, \tilde{\nu} \leq \nu$ and the infimum is always attained.

Generalized Wasserstein distance

Three different admissible actions on μ,ν: either add/remove mass to μ, or add/remove mass to ν or transport mass from μ to ν.

$\mathcal{M}$ space of Borel measures with finite mass.
Given $a,b\in(0,\infty)$ and $p\geq 1$, the generalized Wasserstein distance is given by

$$W_P^{a,b}(\mu,\nu)=\inf_{\substack{\tilde\mu,\tilde\nu\in\mathcal{M}^p\\ |\tilde\mu|=|\tilde\nu|}}\left(a^p|\mu-\tilde\mu|+a|\nu-\tilde\nu|^p+b^pW_p(\tilde\mu,\tilde\nu)\right)^{1/p}.$$

Proposition

The operator $W_p^{a,b}$ is a distance. Moreover, one can restrict the computation to $\tilde\mu\leq\mu,\tilde\nu\leq\nu$ and the infimum is always attained.

Properties of the generalized Wasserstein distance

Theorem

Let $\mu_n, \mu \in \mathcal{M}$. Then

$$W_p^{a,b}(\mu_n, \mu) \to 0 \quad \text{is equivalent to} \quad \mu_n \rightharpoonup \mu \ \text{ and } \ \mu_n \text{ is tight.}$$

Moreover, M is complete with respect to $W_p^{a,b}$.

Proposition

$W_p^{a,b}$ is equivalent to the Levy-Prokhorov distance for probability measures:

$$d_{LP}(\mu, \nu) := \inf\{\alpha > 0 : \text{ for any closed } A, \mu(A) \leq \nu(A^\alpha) + \alpha\}.$$

Proposition

We have (see Kantorovich-Rubinstein duality):

$$W_1^{1,1}(\mu, \nu) = \sup\left\{ \int_{\mathbb{R}^d} f\, d(\mu - \nu) \,\middle|\, \|f\|_{C^0} \leq 1, \|f\|_{Lip} \leq 1 \right\}.$$

Properties of the generalized Wasserstein distance

Theorem

Let $\mu_n, \mu \in \mathcal{M}$. *Then*

$$W_p^{a,b}(\mu_n, \mu) \to 0 \qquad \text{is equivalent to} \qquad \mu_n \rightharpoonup \mu \ \text{ and } \ \mu_n \text{ is tight.}$$

Moreover, M is complete with respect to $W_p^{a,b}$.

Proposition

$W_p^{a,b}$ *is equivalent to the Levy-Prokhorov distance for probability measures:*

$$d_{LP}(\mu, \nu) := \inf\left\{\alpha > 0 : \ \text{for any closed } A, \mu(A) \le \nu(A^\alpha) + \alpha\right\}.$$

Proposition

We have (see Kantorovich-Rubinstein duality):

$$W_1^{1,1}(\mu, \nu) = \sup\left\{\int_{\mathbb{R}^d} f \, d(\mu - \nu) \,\middle|\, \|f\|_{C^0} \le 1, \|f\|_{Lip} \le 1\right\}.$$

Properties of the generalized Wasserstein distance

Theorem

Let $\mu_n, \mu \in \mathcal{M}$. *Then*

$$W_p^{a,b}(\mu_n,\mu) \to 0 \quad \text{is equivalent to} \quad \mu_n \rightharpoonup \mu \ \text{ and } \ \mu_n \text{ is tight.}$$

Moreover, M is complete with respect to $W_p^{a,b}$.

Proposition

$W_p^{a,b}$ *is equivalent to the Levy-Prokhorov distance for probability measures:*

$$d_{LP}(\mu,\nu) := \inf\left\{\alpha > 0 : \text{ for any closed } A, \mu(A) \le \nu(A^\alpha) + \alpha\right\}.$$

Proposition

We have (see Kantorovich-Rubinstein duality):

$$W_1^{1,1}(\mu,\nu) = \sup\left\{\int_{\mathbb{R}^d} f\, d(\mu-\nu) \mid \|f\|_{C^0} \le 1, \|f\|_{Lip} \le 1\right\}.$$

Benamou-Brenier formula

Consider the transport equation with source:

$$\partial_t \mu_t + \nabla \cdot (v_t \mu_t) = h_t. \tag{1}$$

Define:

$$\mathcal{B}[\mu, v, h] := a^2 \left(\int_0^1 dt \left(\int_{\mathbb{R}^d} d|h_t| \right) \right)^2 + b^2 \int_0^1 dt \left(\int_{\mathbb{R}^d} d\mu_t \, |v_t|^2 \right).$$

Proposition

We have

$$\left(W_2^{a,b} \right)^2 = \inf \left\{ \mathcal{B}[\mu, v, h] \;\middle|\; \begin{array}{c} \mu \text{ is a solution of (1) for } v, h \\ \text{and } \mu(0) = \mu_0, \mu(1) = \mu_1 \end{array} \right\}.$$

Benamou-Brenier formula

Consider the transport equation with source:

$$\partial_t \mu_t + \nabla \cdot (v_t \mu_t) = h_t. \tag{1}$$

Define:

$$\mathcal{B}[\mu, v, h] := a^2 \left(\int_0^1 dt \left(\int_{\mathbb{R}^d} d|h_t| \right) \right)^2 + b^2 \int_0^1 dt \left(\int_{\mathbb{R}^d} d\mu_t \, |v_t|^2 \right).$$

Proposition

We have

$$\left(W_2^{a,b}\right)^2 = \inf \left\{ \mathcal{B}[\mu, v, h] \;\middle|\; \begin{array}{c} \mu \text{ is a solution of (1) for } v, h \\ \text{and } \mu(0) = \mu_0, \mu(1) = \mu_1 \end{array} \right\}.$$

Benamou-Brenier formula

Consider the transport equation with source:

$$\partial_t \mu_t + \nabla \cdot (v_t \mu_t) = h_t. \tag{1}$$

Define:

$$\mathcal{B}[\mu, v, h] := a^2 \left(\int_0^1 dt \left(\int_{\mathbb{R}^d} d|h_t| \right) \right)^2 + b^2 \int_0^1 dt \left(\int_{\mathbb{R}^d} d\mu_t \, |v_t|^2 \right).$$

Proposition

We have

$$\left(W_2^{a,b}\right)^2 = \inf \left\{ \mathcal{B}[\mu, v, h] \;\middle|\; \begin{array}{c} \mu \text{ is a solution of (1) for } v, h \\ \text{and } \mu(0) = \mu_0, \mu(1) = \mu_1 \end{array} \right\}.$$