On the weighted total variation distance between probability measures

June 26, 2021

Abstract

A short note on the weighted total variation distance for probability measures on $\mathbb{R}^d$. Everything is certainly well-known, but I could not find a proper reference.

Given $\phi : \mathbb{R}^d \to \mathbb{R}_+$ a non-negative measurable function, we define the weighted total variation distance between two probability measures ν and μ by

$$d_{TV}^{\phi}(\nu,\mu) := \int \phi(x)|\nu - \mu|(dx).$$

Theorem 1. *The weighted total variation distance satisfies*

$$d_{TV}^{\phi}(\nu,\mu) = \inf_{\mathcal{L}(X)=\nu,\mathcal{L}(Y)=\mu} \mathbb{E}(\phi(X)+\phi(Y))\mathbb{1}_{X\neq Y}.$$

In addition, the Kantorovich duality theorem holds

$$d_{TV}^{\phi}(\nu,\mu) = \sup_{g} \int g(x)(\nu-\mu)(dx),$$

where the supremum is taken over all the measurable functions $g : \mathbb{R}^d \to \mathbb{R}$ *s.t.*

$$\forall x, y, \quad |g(x) - g(y)| \leq (\phi(x)+\phi(y))\mathbb{1}_{x\neq y}. \tag{1}$$

Remark 2. *This result states that the weighted total variation distance is equal to the Wasserstein distance* W_1 *for the following peculiar semi-metric*

$$d(x,y) := (\phi(x)+\phi(y))\mathbb{1}_{x\neq y}.$$

See also [CD18, Ch. 5].

Remark 3. *The standard (non-weighted) total variation distance is obtained by choosing* $\phi \equiv 1/2$. *In that case, the result simplifies to*

$$d_{TV}^{1/2}(\nu,\mu) = \inf_{\mathcal{L}(X)=\nu,\mathcal{L}(Y)=\mu} \mathbb{P}(X\neq Y) = \sup_{A\in\mathcal{B}(\mathbb{R}^d)} \nu(A) - \mu(A),$$

where $\mathcal{B}(\mathbb{R}^d)$ *denotes the Borel sets of* $\mathbb{R}^d$.

Proof. Step 1. We prove that $d_{TV}^{\phi}(\nu,\mu) \le \sup_g \int g(x)(\nu-\mu)(dx)$.
Let m be a non-negative measure such that $\nu \ll m$ and $\mu \ll m$ (one can choose $m = \nu + \mu$). The Radon–Nikodym theorem provides non-negative measurable functions h_ν and h_μ such that

$$\nu(dx) = h_\nu(x)m(dx) \quad \text{and} \quad \mu(dx) = h_\mu(x)m(dx).$$

Let $B := \{h_\nu \ge h_\mu\}$ and $\bar{B} := \mathbb{R}^d \setminus B$. We choose $g(x) := \phi(x)\mathbb{1}_B(x) - \phi(x)\mathbb{1}_{\bar{B}}(x)$. It satisfies (1) and

$$\int \phi(x)|\nu-\mu|(dx) = \int g(x)(\nu-\mu)(dx).$$

Step 2. We prove that $\sup_g \int g(x)(\nu-\mu)(dx) \le \inf_{\mathcal{L}(X)=\nu,\mathcal{L}(Y)=\mu} \mathbb{E}(\phi(X)+\phi(Y))\mathbb{1}_{X\neq Y}$.
Let g such that (1) holds. Consider π any coupling between ν and μ, we have

$$\begin{aligned}\int g(x)\nu(dx) - \int g(y)\mu(dy) &= \int\int (g(x)-g(y))\pi(dx,dy)\\ &\le \int\int |g(x)-g(y)|\pi(dx,dy)\\ &\le \int\int (\phi(x)+\phi(y))\mathbb{1}_{x\neq y}\pi(dx,dy).\end{aligned}$$

Taking the infimum of the right-hand side over all the couplings gives the result.
Step 3. We prove that $\inf_{\mathcal{L}(X)=\nu,\mathcal{L}(Y)=\mu} \mathbb{E}(\phi(X)+\phi(Y))\mathbb{1}_{X\neq Y} \le d_{TV}^{\phi}(\nu,\mu)$.
We use the following explicit coupling taken from [Vil09]

$$\pi(dx,dy) := (\nu\wedge\mu)(dx)\delta_x(dy) + \frac{1}{a}(\nu-\mu)_+(dx)(\nu-\mu)_-(dy),$$

where $a := \int(\nu-\mu)_+ = \int(\nu-\mu)_-$ and $\nu\wedge\mu := \nu - (\nu-\mu)_+$. Let $(X,Y)\sim\pi$. We have

$$\mathbb{E}\phi(X) = \mathbb{E}\phi(X)\mathbb{1}_{X=Y} + \mathbb{E}\phi(X)\mathbb{1}_{X\neq Y},$$

and denoting by $\alpha := \int \phi(x)(\nu-\mu)_+(dx)$, it holds that

$$\begin{aligned}\mathbb{E}\phi(X)\mathbb{1}_{X=Y} &\ge \int \phi(x)(\nu\wedge\mu)(dx)\\ &= \int \phi(x)\nu(dx) - \int \phi(x)(\nu-\mu)_+(dx)\\ &= \mathbb{E}\phi(X) - \alpha.\end{aligned}$$

So $\mathbb{E}\phi(X)\mathbb{1}_{X\neq Y} \le \alpha$. Similarly, let $\beta := \int \phi(x)(\nu-\mu)_-(dx)$, a similar argument shows that

$$\mathbb{E}\phi(Y)\mathbb{1}_{X\neq Y} \le \beta.$$

Altogether, we have

$$\mathbb{E}(\phi(X)+\phi(Y))\mathbb{1}_{X\neq Y} \le \alpha+\beta = d_{TV}^{\phi}(\nu,\mu).$$

Combining the three steps ends the proof. □

References

[CD18] R. Carmona and F. Delarue. *Probabilistic theory of mean field games with applications. I*. Vol. 83. Probability Theory and Stochastic Modelling. Mean field FBSDEs, control, and games. Springer, Cham, 2018, pp. xxv+713.

[Vil09] C. Villani. *Optimal transport*. Vol. 338. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Old and new. Springer-Verlag, Berlin, 2009, pp. xxii+973. DOI: 10.1007/978-3-540-71050-9.