Wasserstein gradient flow 2

Tuesday, January 10, 2023 12:04 PM

$\mathcal{P}_2(\mathbb{R}^d)$ - set of Borel prob measures on $\mathbb{R}^d$ with finite second moments

$W_2(P, Q)$ - Wasserstein - 2 metric.

[See Villani - Topics in OT, Chapter 8]

Suppose $\mathcal{F} : \mathcal{P}_2(\mathbb{R}^d) \longrightarrow \mathbb{R}$

What do I mean by gradient of this function ?

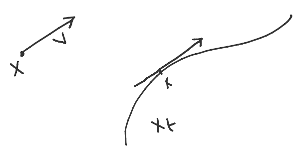
What about Euclidean spaces ?

$F : \mathbb{R}^d \longrightarrow \mathbb{R}$ $\quad x \longmapsto \nabla F(x)$

How to interpret this when $\mathbb{R}^d$ is seen as a Riemannian manifold ?

$x \in \mathbb{R}^d$ $\quad$ Tangent space $= \{ v \in \mathbb{R}^d$, velocities of curves as it passes through $x \}$

Think of $\nabla F(x)$ as an element in the tangent space.



Which element ? Suppose a C^1 curve starts at x.

$x_0 = x$ $\quad \dot{x}_t = v_t$

$v_0 = v$

Consider $t \longmapsto F(x_t)$

Hence, there is a notion of directional derivative

$$\frac{d}{dt}\Big|_{t=0} F(x_t) \underset{\text{rule}}{\overset{\text{chain}}{=}} \nabla F(x) \cdot \dot{x}_t\Big|_{t=0} = \boxed{\nabla F(x) \cdot v}$$

Definition $\nabla F(x)$ is that element in the tangent space at x such that

$$\boxed{\frac{d}{dt}\Big|_{t=0} F(x_t) = \nabla F(x) \cdot v}$$

where $\left. \begin{array}{l} \dot{x}_0 = v \\ x_0 = x \end{array} \right\}$

Wasserstein - 2 space is a "infinite - dimensional Riemannian manifold."

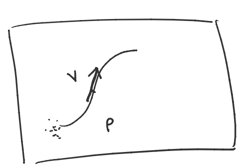
If $\rho \in \mathcal{P}_2(\mathbb{R}^d)$ $\quad \mathrm{Tan}_\rho = \overline{\{ \nabla g, \; g \in C_c^\infty(\mathbb{R}^d) \}}^{L^2(\rho)}$

Let $\underset{\underset{L^2(\rho)}{\uparrow}}{v} \in \mathrm{Tan}\,\rho$. How can I create a smooth curve with velocity v at ρ ?

Continuity eqn. $\quad \partial_t \rho_t + \nabla \cdot (v_t \rho_t) = 0$

$\left. \begin{array}{l} \rho_0 = \rho \\ v_0 = v \end{array} \right\}$

$v_t(x)$



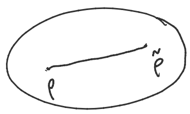
Let $\mathcal{F} : \mathcal{P}_2(\mathbb{R}^d) \longrightarrow \mathbb{R}$

$\downarrow$ is a convex set

How can I construct a directional derivative of $\mathcal{F}$?

$\left(\rho_\varepsilon = (1-\varepsilon)\rho + \varepsilon \tilde{\rho}, \quad 0 \le \varepsilon \le 1 \right) \subseteq \mathcal{P}_2(\mathbb{R}^d)$

is a convex set



$\left(\rho_\varepsilon = (1-\varepsilon)\rho + \varepsilon \tilde{\rho}, \quad 0 \le \varepsilon \le 1 \right)$

Consider $\varepsilon \longmapsto \mathcal{F}(\rho_\varepsilon)$ and compute $\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \mathcal{F}(\rho_\varepsilon)$

Definition First variation Given a function $\mathcal{F} : \mathcal{P}_2 \longrightarrow \mathbb{R}$ (regularity assumed)

$\frac{\delta \mathcal{F}}{\delta \rho}$ is that function which satisfies

$$\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \mathcal{F}(\rho_\varepsilon) = \boxed{\int \frac{\delta F}{\delta \rho}(x)\, \dot{\rho}_0(x)\, dx}$$

$\dot{\rho}_t + \nabla\cdot(v_t \rho_t) = 0.$

$\left.\frac{d}{dt}\right|_{t=0} \mathcal{F}(\rho_t)$

How to define Wasserstein gradient from first variation?

$v \in \mathrm{Tan}_\rho$ $\qquad \dot{\rho}_t + \nabla\cdot(v_t \rho_t) = 0 \qquad v_0 = v, \quad \rho_0 = \rho$

Let $v^* = \nabla_{W_2} \mathcal{F}(\rho) \in \mathrm{Tan}_\rho$

$$\left.\frac{d}{dt}\right|_{t=0} \mathcal{F}(\rho_t) = \boxed{\int v^*(x)\, v(x)\, \rho(x)\, dx}$$

$v \in \mathrm{Tan}_\rho$

$$\boxed{v^* = \nabla_x \left(\frac{\delta F}{\delta \rho} \right)}$$

See Villani Topics in OT Page 251

$$\| \qquad \int \frac{\delta F}{\delta \rho}(x)\, \dot{\rho}_0(x)\, dx = -\int \frac{\delta F}{\delta \rho}(x)\, \nabla\cdot(v\rho)\, dx$$

$$\overset{IP}{=} \int \nabla_x \left(\frac{\delta F}{\delta \rho} \right)(x)\, v(x)\, \rho(x)\, dx$$

Definition (Lemma in AGS 10.4.1)

$$\boxed{\nabla_{W_2} \mathcal{F}(\rho) = \nabla_x \left(\frac{\delta F}{\delta \rho} \right)}$$

Example 1

$$\mathcal{F}(\rho) = \int_{\mathbb{R}} g(x)\, \rho(x)\, dx \qquad g - \text{bdd function } C^1$$

$\frac{\delta \mathcal{F}}{\delta \rho} \qquad \rho_\varepsilon = (1-\varepsilon)\rho + \varepsilon \tilde{\rho}$

$$\mathcal{F}(\rho_\varepsilon) = \int g(x) \left[(1-\varepsilon)\rho(x) + \varepsilon \tilde{\rho}(x) \right] dx$$

$$= \int g(x)\, \rho(x) + \varepsilon \int g(x) \underbrace{\left(\tilde{\rho}(x) - \rho(x) \right)}_{-\nabla\cdot(v_t(x) \rho_t)} dx$$

$$\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \mathcal{F}(\rho_\varepsilon) = \int g(x)\, \dot{\rho}_0(x)\, dx$$

$$\boxed{\frac{\delta \mathcal{F}}{\delta \rho} = g(x)}$$

$$\nabla_W \mathcal{F}(\rho) = \nabla \left(\frac{\delta \mathcal{F}}{\delta \rho} \right)$$

$\|$

$\langle \quad \mathcal{F}(\ldots) \quad v \rangle$

$$\nabla_{W_2} \mathcal{F}(\rho) = \nabla\left(\frac{\delta \mathcal{F}}{\delta \rho}\right) = \nabla g(x)$$

$$\Big|_{\varepsilon=0} = \left\langle \nabla_{W_2}\mathcal{F}(\rho), v \right\rangle$$

$$\boxed{\frac{\delta \mathcal{F}}{\delta \rho} = g(x)}$$

$f(y) = y \log y$

$f'(y) = \log y + 1$

Example 2

$$\mathcal{F}(\rho) = \mathrm{Ent}(\rho) = \begin{cases} \int \rho(x)\log \rho(x)\, dx = \int f(\rho(x))\,dx \\ \infty \end{cases}$$

$$\rho_\varepsilon = (1-\varepsilon)\rho + \varepsilon \tilde{\rho} \qquad \mathcal{F}(\rho_\varepsilon) = \int f\left((1-\varepsilon)\rho(x) + \varepsilon \tilde{\rho}(x)\right) dx$$

$$\boxed{\varepsilon \longmapsto \mathcal{F}(\rho_\varepsilon)}$$

$$\frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \mathcal{F}(\rho_\varepsilon) = \int f'\left((1-\varepsilon)\rho + \varepsilon\tilde{\rho}\right)(\tilde{\rho}-\rho)\,dx \Big|_{\varepsilon=0}$$

$$= \int f'(\rho)(\tilde{\rho}-\rho)\,dx$$

$$\frac{\delta \mathcal{F}}{\delta \rho} = f'(\rho) = \log \rho(x) + 1$$

$$\nabla_{W_2}\mathcal{F}(\rho) = \nabla\left(\frac{\delta \mathcal{F}}{\delta \rho}\right) = \nabla \log \rho(x)$$

Gradient flows

Euclidean $\quad \boxed{\dot{x}_t = -\nabla F(x_t)} \qquad x_0 = x$

Wasserstein space $\quad \nabla_{W_2}\mathcal{F}(\rho)(x) \in L^2(\rho)$

$$\boxed{\partial_t \rho_t + \nabla\cdot(v_t \rho_t) = 0 \qquad v_t = \nabla_{W_2}\mathcal{F}(\rho_t)}$$

Example

$$\mathcal{F}(\rho) = \int \rho(x)\log\rho(x)\,dx \qquad v_t = \nabla_{W_2}\mathcal{F}(\rho_t) = \nabla \log \rho_t = \frac{\nabla \rho_t}{\rho_t}$$

$$\partial_t \rho_t = -\nabla\cdot(v_t\rho_t) = -\nabla\cdot\left(\frac{\nabla\rho_t}{\cancel{\rho_t}}\cdot\cancel{\rho_t}\right)$$

$$= \Delta(\rho_t)$$

$$\partial_t \rho_t = \Delta(\rho_t) \longleftarrow \text{Heat equation.}$$