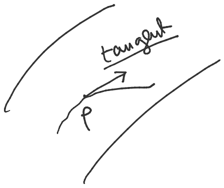
Wasserstein gradient flow 3

Tuesday, January 17, 2023 12:03 PM

Recall $\mathcal{P}_2(\mathbb{R}^d)$ equipped with $\boxed{W_2 \text{ metric}}$

infinite-dim Riemannian manifold.



Smooth curves: Continuity eqn ✓

$$\dot{\rho}_t + \nabla\cdot(v_t \rho_t) = 0 \qquad v_t \in L^2(\rho_t).$$

Every particle is moving by the velocity field (v_t).

n particles

$X_t(i) \leftarrow$ ith particle.

$$\underline{\dot{X}_t(i) = v_t(X_t(i))}$$

$$\rho_t = \frac{1}{n}\sum_{i=1}^n \delta_{X_t(i)}$$

$$\frac{d}{dt}\int f(x)\rho_t(x)\,dx = \frac{d}{dt}\left[\frac{1}{n}\sum_{i=1}^n f(X_t(i))\right]$$

$$= \frac{1}{n}\sum_{i=1}^n \nabla f(X_t(i))\cdot \dot{X}_t(i)$$

$$= \frac{1}{n}\sum_{i=1}^n \nabla f(X_t(i))\cdot v_t(X_t(i))$$

$$= \int \nabla f(y)\, v_t(y)\, \rho_t(y)\,dy$$

$$= -\int f(y)\,\nabla\cdot(v_t(y)\rho_t(y))\,dy$$

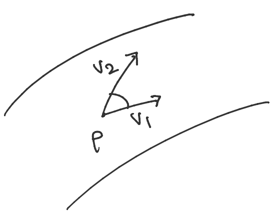
$$\boxed{\frac{d}{dt}\int f(x)\rho_t(x)\,dx} = \int f(x)\,\dot{\rho}_t(x)\,dx$$

first variation

$$\boxed{\dot{X}_t = v_t(X_t)}$$

$$\mathrm{Tan}_\rho = \overline{\{\nabla g,\ g \in C_c^\infty\}}^{L^2(\rho)}$$

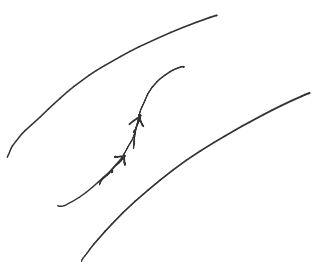
Metric Tensor



$v_1, v_2 \in \mathrm{Tan}_\rho$

$$\langle v_1, v_2\rangle = \int v_1(x)\, v_2(x)\, \rho(x)\,dx.$$

Length of curves



$(\rho_t,\ 0 \le t \le 1)$

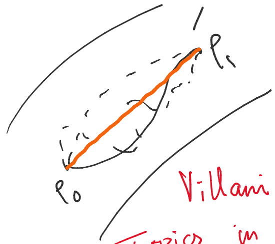
$$\dot{\rho}_t = -\nabla\cdot(v_t\rho_t) \qquad v_t \in \mathrm{Tan}_{\rho_t}$$

$$\mathrm{Length}((\rho_t)) = \int_0^1 \|v_t\|_{L^2(\rho_t)}\,dt$$

Geodesic

$W_2(\rho_0, \rho_1) =$ minimum length of a smooth curve joining ρ_0 and ρ_1

$W_2(\rho_0, \rho_1)$ curve joining ρ_0 and ρ_1



Villani Topics in OT Section 8.1.1

Benamou – Brenier Theorem

geodesic $\|v_t\|^2 = W_2^2(\rho_0,\rho_1)$

$$W_2^2(\rho_0, \rho_1) = \boxed{\min} \int_0^1 \|v_t\|^2_{L^2(\rho_t)}\, dt$$

all (ρ_t), $0 \le t \le 1$, $\quad \dot\rho_t + \nabla\cdot(v_t \rho_t) = 0$

$v_t \in L^2(\rho_t)$

Where is this minimized? $\boxed{\text{Constant velocity curve.}}$

McCann interpolation (Brenier)

$\varphi - cx$

$\nabla\varphi : \rho_0 \longrightarrow \rho_1$

Suppose $\quad X(0) = x \qquad X(1) = \nabla\varphi(x)$

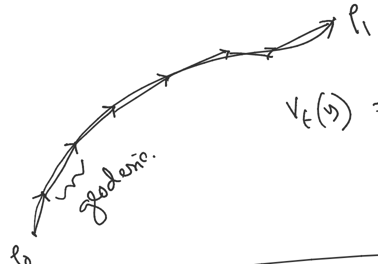
$X_t(x) = (1-t)\,x + t\nabla\varphi(x)$

$x = X_t^{-1}(y) \qquad \dot X_t(x) = \nabla\varphi(x) - x$

$\nabla\varphi(x)$

x

$\left(\rho_t = \rho_0 \# X_t\right) = \text{geodesic.}$



Brenier

$v_t(y) = \boxed{\nabla\varphi(x) - x}$ where $x = X_t^{-1}(y)$

$\dot\rho_t = -\nabla\cdot(v_t\rho_t)$ ✓

$v_t \in \mathrm{Tan}\,\rho_t$

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

$\left[\ \mathcal{F}(\rho) = \int F(x, \rho(x), \nabla\rho(x))\, dx \quad \text{where } F(x, z, p)\ \right]$

Natural: First variation

$\dot\rho_t + \nabla\cdot(v_t\rho_t) = 0 \qquad v_t \in \mathrm{Tan}_{\rho_t} \qquad \rho_0 = \rho_0$

$\mathcal{F}(\rho) = \int V(x)\rho(x)\,dx$, $\quad \frac{\delta\mathcal{F}}{\delta\rho} = V$

$\mathcal{F}(\rho) = \int \rho(x)\log\rho(x)\,dx$

$\frac{\delta\mathcal{F}}{\delta\rho} = \log\rho + 1$

$$\frac{d}{dt}\Big|_{t=0}\mathcal{F}(\rho_t) = \int g(x)\,\dot\rho_0(x)\,dx \quad ✓$$

$g(x) = \text{first variation} = \dfrac{\delta\mathcal{F}}{\delta\rho}$

Theorem $\quad \mathcal{F}(\rho) = \int F(x, \rho(x), \nabla\rho(x))\, dx$

AGS eqn. (10.4.2) $\quad \dfrac{\delta \mathcal{F}}{\delta \rho} = \dfrac{\partial F}{\partial z}\left(x, \rho(x), \nabla \rho(x)\right) - \nabla_p \cdot F\left(x, \rho(x), \nabla \rho(x)\right)$

Formula [Riemannian notion]

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

Villani Topics in OT Sec 8.2

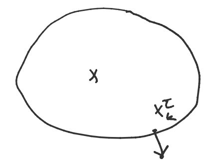
JKO Scheme

Euclidean gradient flow

$F : \mathbb{R}^d \to \mathbb{R}$

X_0

$$\underbrace{\dot{x}_t = -\nabla F(x_t)}_{\text{steepest descent.}}$$



Practice

Euler schemes

Explicit step size $\tau \approx 0$

$$x_0^\tau = x_0$$

$$\frac{x_{k+1}^\tau - x_k^\tau}{\tau} = -\nabla F(x_k^\tau)$$

Doesn't work very well!

$$x_{k+1}^\tau = x_k^\tau - \tau \nabla F(x_k^\tau).$$

Implicit Define

$$x_0^\tau = x_0$$

$$x_{k+1}^\tau = \operatorname*{argmin}_{x} \underbrace{\left[F(x) + \frac{1}{2\tau} \| x - x_k^\tau \|^2 \right]}_{\text{FOC} \quad = 0}$$

$$\nabla F(x) + \frac{1}{\tau}(x - x_k^\tau) = 0$$

$$\frac{x_{k+1}^\tau - x_k^\tau}{\tau} = -\nabla F(x_{k+1}^\tau)$$

Better convergence properties.

JKO Scheme = Implicit Euler in Wasserstein space.

ρ_0 $\quad$ Iteratively,

$\tau \approx 0$

$$\rho_{k+1}^\tau = \operatorname*{argmin}_{\rho} \left[\mathcal{F}(\rho) + \frac{1}{2\tau} W_2^2(\rho, \rho_k^\tau) \right]$$

Question? (τ) Interpolate

Question ?

$$\left(\rho_k^\tau, \quad k=0,1,\dots\right) \qquad \text{Interpolate.}$$

$$\rho_t^\tau = \rho_k^\tau, \quad k\tau < t \le (k+1)\tau$$

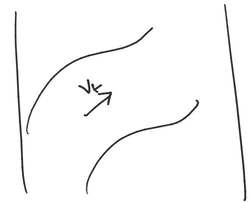
time $\quad \tau \quad 2\tau \quad 3\tau \quad \dots$

Does this curve $\left(\rho_t^\tau, t\geq 0\right)$ converge as $\tau \downarrow 0$?

Yes ! Wasserstein gradient flow !

Wasserstein gradient flow

$$\dot\rho_t + \nabla\cdot\left(v_t\,\rho_t\right) = 0$$



$$v_t = -\nabla_x\left(\frac{\delta \mathcal{F}}{\delta\rho}\right)$$

$$\dot\rho_t = \nabla\cdot\left(\nabla\left(\frac{\delta\mathcal{F}}{\delta\rho}\right)\cdot\rho_t\right).$$

Original

$$\mathcal{F}(\rho) = \mathrm{Ent}(\rho) = \int \rho(x)\log\rho(x)\,dx$$

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

$$\boxed{\dot\rho_t = \Delta\rho_t}$$

Heat equation

Go back to Implicit Euler

$$\rho_{k+1}^\tau = \underset{\rho}{\operatorname{argmin}}\Big[\underbrace{\mathcal{F}(\rho) + \frac{1}{2\tau}W_2^2\left(\rho,\rho_k^\tau\right)}_{\mathcal{G}(\rho)}\Big]$$

FOC : $\dfrac{\delta\mathcal{G}}{\delta\rho} = 0$

$$\frac{\delta\mathcal{G}(\rho)}{\delta\rho} = \frac{\delta\mathcal{F}}{\delta\rho} + \frac{1}{\tau}\underbrace{\frac{\delta}{\delta\rho}\boxed{\frac{1}{2}W^2\left(\rho,\rho_k^\tau\right)}}_{\varphi} \longrightarrow \text{duality} \quad \int\varphi(x)\rho(x)dx + \int\varphi^c(x)\rho_k^\tau(x)dx$$

φ – Kantorovich Potential transporting $\rho \longrightarrow \rho_k^\tau$.

$\boxed{\nabla\varphi} = x -$ Brenier

See Santambrogio Section 8.2

$$\frac{\delta\mathcal{F}}{\delta\rho} + \frac{1}{\tau}\varphi = 0$$

$\delta\rho$

$$\nabla\left(\frac{\delta \mathcal{F}}{\delta\rho}\right) + \frac{1}{\tau}\nabla\varphi = 0$$

$$\boxed{\frac{1}{\tau}\nabla\varphi} = -\nabla\left(\frac{\delta \mathcal{F}}{\delta\rho}\right)$$

$$\boxed{v \cong -\nabla\left(\frac{\delta \mathcal{F}}{\delta\rho}\right)}$$

ρ_k^τ

Gradient flow

$$\left.\begin{aligned} \dot{\rho}_t + \nabla\cdot(v_t\rho_t) &= 0 \\ v_t &= -\nabla\left(\frac{\delta \mathcal{F}}{\delta\rho}\right) \end{aligned}\right] \quad \text{Wass gradient flow.}$$

Many other examples PDE – Wass. grad. flow.

Summary

1. Repeated push-forwards of particles lead to smooth curves in W_2 space.
2. Some of these curves are grad. flows.
3. All these curves can be described by PDEs in the limit of many iteration.