• OT for Applied Mathematicians • 5.3.5 Derivative of W_p^p along curves of measures.

• Thm 5.24. Let $(\varrho_t^{(i)}, v_t^{(i)})$ for $i=1,2$. be two solutions of the continuity equation $\partial_t \varrho_t^{(i)} + \nabla\cdot(v_t^{(i)}, \varrho_t^{(i)}) = 0$ on a compact domain Ω, and suppose that $\varrho_t^{(i)} \ll \mathcal{L}^d$ for every t and that $\varrho^{(i)}$ are absolutely continuous curves in $W_p(\Omega)$. Then we have

$$\frac{d}{dt}\left(\frac{1}{p} W_p^p(\varrho_t^{(1)}, \varrho_t^{(2)})\right) = \int \nabla\varphi_t \cdot v_t^{(1)} \varrho_t^{(1)}\, dx + \int \nabla\psi_t \cdot v_t^{(2)} \varrho_t^{(2)}\, dx.$$

for a.e. t, where (φ_t, ψ_t) is any pair of Kantorovich potentials in the transport between $\varrho_t^{(1)}$ and $\varrho_t^{(2)}$ for the cost $\frac{1}{p}|x-y|^p$.

• Cor 5.25. Under the same assumptions of Thm 5.24, we have

$$\frac{d}{dt}\left(\frac{1}{p} W_p^p(\varrho_t^{(1)}, \varrho_t^{(2)})\right) = \int_\Omega (x - T(x)) \cdot \left(v_t^{(1)}(x) - v_t^{(2)}(T_t(x))\right) \varrho_t^{(1)}(x)\, dx.$$

where T_t is the optimal transport map from $\varrho_t^{(1)}$ to $\varrho_t^{(2)}$ for the cost $\frac{1}{p}|x-y|^p$.

- variational derivative

for a functional $F : (P_p, W_p) \longrightarrow \mathbb{R}$ and a probability measure ρ, if $F(\varepsilon\rho + (1-\varepsilon)\tilde{\rho}) < \infty$.

we define the variational derivative of F at ρ as function f s.t.

$$\frac{d}{d\varepsilon}\Big|_{\varepsilon=0} F(\rho + \varepsilon(\rho - \tilde{\rho})) = \int f \, d(\tilde{\rho} - \rho).$$

study of Wasserstein flows as it appears in the continuity equation

$$\partial_t \mu + \mathrm{div}(\mu \cdot (f\mu)) = 0.$$

where f is the variational derivative of F.

Further the variational derivative of $W_2(\mu, \nu)^2$ is $\varphi_{\mu,\nu}$: the kantorovich potential between the measures.

- Def 1.12 : The functions φ realizing the maximum in (3.1) are called Kantorovich potential for the transport from μ to ν. This is in fact a small abuse, because traditionally, this term was only in the case $c(x,y) = |x-y|$, but it is nowadays understood in the general cases as well.

(3.1) ⋯ Duality formula : $\min\left\{ \int_{X \times X} c(x,y) \, d\gamma : \gamma \in \Pi(\mu,\nu) \right\} = \max\left\{ \int_X u \, d(\mu - \nu) : u \in Lip_1 \right\}$.

$$\arg\min_{\rho} D_{KL}(\rho \,\|\, \pi_n)$$

$$= \arg\min_{\rho} \int \rho(x) \cdot \log\left(\frac{\rho(x)}{\pi_n(x)}\right) dx$$

$$= \arg\min_{\rho} \underbrace{\int \rho(x) \cdot \log(\rho(x))\, dx}_{\text{Fokker-Planck.}} + \int \underbrace{-\log(\pi_n(x))}_{V(x).} \cdot \rho(x)\, dx. \quad \text{potential.}$$

$$= \arg\min_{\rho} F(\rho)\ ; \quad F(\rho) = \underbrace{\int \rho(x) \cdot \log(\rho(x))\, dx}_{\text{entropy.}} + \underbrace{\int V(x)\, d\rho(x).}_{\text{potential energy}}$$

Fokker-Planck equation: $\partial_t(\rho) = \nabla \cdot (\rho \cdot \nabla V)$, potential $V(x) = -\log(\pi_n(x))$.

$$\frac{\partial \rho(x,t)}{\partial t} = \frac{\partial}{\partial x}\left[\frac{\partial V(x)}{\partial x}\rho(x,t)\right] + \frac{\partial^2 \rho(x,t)}{\partial x^2}.$$

The Variational Formulation of the Fokker-Planck Equation:

The Gibbs distribution ρ_s satisfies a variational principle: it minimizes over all probability density on $\mathbb{R}^n$ the free energy functional

$$F(\rho) = E(\rho) + \beta^{-1} S(\rho).$$

where $E(\rho) = \int_{\mathbb{R}^n} \Psi \rho\, dx$ plays a role of an (potential) energy functional, and

$S(\rho) = \int_{\mathbb{R}^n} \rho \log \rho\, dx$ is the negative of the Gibbs-Boltzmann entropy functional.

Fokker - Planck free energy

$$F_{FP}(\rho) = \mathcal{U}(\rho) - \beta^{-1} \cdot \mathcal{E}(\rho).$$

$$\begin{cases} \mathcal{U}(\rho) = \int_{\mathbb{R}^D} \Phi(x)\, d\rho(x) & \text{potential energy} \\ \mathcal{E}(\rho) = - \int_{\mathbb{R}^D} \log \frac{d\rho}{dx}(x)\, d\rho(x) & \text{entropy} \\ \beta > 0 \text{ is magnitude.} \end{cases}$$

estimate :

$$\hat{\mathcal{U}}(x_1, \dots, x_N) = \frac{1}{N} \sum_{n=1}^{N} \Phi(T(x_n))$$

$$\widehat{\Delta\mathcal{E}}(x_1, \dots, x_N) = \frac{1}{N} \sum_{n=1}^{N} \log |\det \nabla T(x_n)|.$$

$$\hat{W}_2^2(x_1, \dots, x_N) = \frac{1}{N} \sum_{n=1}^{N} \| T(x) - x \|_2^2.$$

Loss :

$$\hat{\mathcal{L}} = \frac{1}{2h} \hat{W}_2^2 + \hat{\mathcal{U}} - \beta^{-1} \cdot \widehat{\Delta\mathcal{E}}.$$

$\Phi : \mathbb{R}^D \to \mathbb{R}$ is the potential function. $\Phi(x) = -\log(\pi_n(x))$, π_n is Stationary Dist. (posterior).