Wasserstein Gradient Flows of Moreau Envelopes of f-Divergences in Reproducing Kernel Hilbert Spaces

joint work with

Sebastian Neumayer, TU Chemnitz

Gabriele Steidl, TU Berlin

Nicolaj Rux, TU Berlin

Learning and Optimization 2024

Luminy, 17.06.2024

Goal. Recover $\nu \in \mathcal{P}(\mathbb{R}^d)$ by minimizing f-divergence $D_{f,\nu}$ to ν, such as $\mathrm{KL}(\cdot \mid \nu)$ (e.g. VI, GANs).

Problem. Only samples $\rightsquigarrow$ empirical measures, but

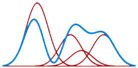
$$\mu \not\ll \nu \implies D_{f,\nu}(\mu) = \infty.$$

weak convergence

Our Solution. Regularize $D_{f,\nu} \colon \mathcal{M}(\mathbb{R}^d) \to [0, \infty]$.

1. "Kernel trick"

$$m \colon \mu \mapsto m_\mu := \int_{\mathbb{R}^d} K(x, \cdot) \, \mathrm{d}\mu(x) \in \mathcal{H}_K$$



pointwise convergence

$$\text{“}D_{f,\nu} \circ m^{-1}\text{”} = G_{f,\nu} \colon \mathcal{H}_K \to [0, \infty]$$

2. Moreau envelope regularization

$${}^{\lambda}G_{f,\nu}(m_\mu) = \min_{\sigma \in \mathcal{M}_+(\mathbb{R}^d)} D_{f,\nu}(\sigma) + \frac{1}{2\lambda} \|m_\sigma - m_\mu\|_{\mathcal{H}_K}^2, \qquad \lambda > 0$$

We prove existence & uniqueness of W_2 gradient flows of $\mu \mapsto {}^{\lambda}G_{f,\nu}(m_\mu)$.

Simulate particle flows $= W_2$ gradient flows starting at empirical measure