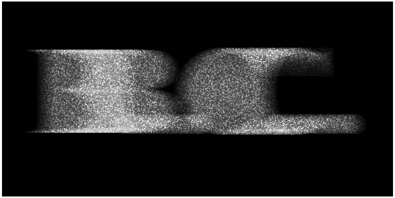
Solving Fredholm Integral Equations through Wasserstein Gradient Flows

Francesca R. Crucinio
with Adam M. Johansen & Arnaud Doucet

25 June 2020

Fredholm Integral Equations of the First Kind (FIE)

$$\mu(dy) = \int_{\mathbb{X}} K(x, y)\rho(dx)$$



constant speed horizontal motion

Fredholm Integral Equations of the First Kind (FIE)

- Applications: PDEs, indirect density estimation, signal reconstruction, causal inference, ...
- Inverse ill posed problems
- Solution methods often require discretisation/strong assumptions on f

FIE - Solution method

$$\mu(dy) = \int_{\mathbb{X}} \rho(dx) K(x, y) \qquad \forall y \in \mathbb{Y},$$

Take ρ, μ probability measures and K a Markov transition density

$$\rho^\star := \operatorname{argmin} \mathrm{KL}(\mu, \rho K) - \alpha\, \mathrm{ent}(\rho)$$

for some $\alpha > 0$.

FIE - Solution method

$$\mu(dy) = \int_{\mathbb{X}} \rho(dx) K(x,y) \qquad \forall y \in \mathbb{Y},$$

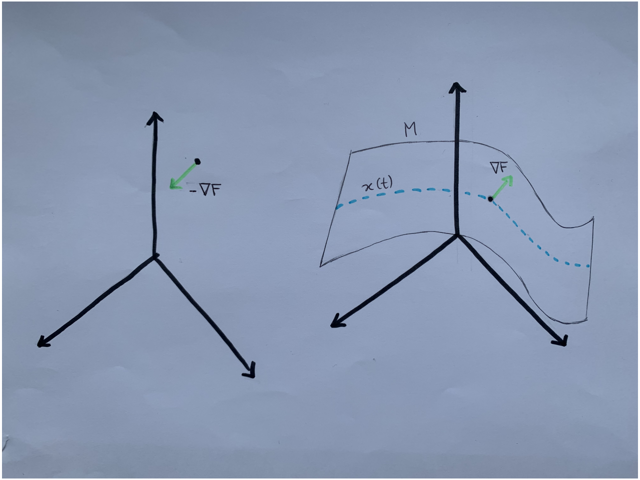
Take ρ, μ probability measures and K a Markov transition density

$$\rho^\star := \operatorname{argmin} \mathrm{KL}(\mu, \rho K) - \alpha \, \mathrm{ent}(\rho)$$

for some $\alpha > 0$.
How?

$$\text{minimisation} \longrightarrow \text{PDE} \longrightarrow \text{SDE}$$

Gradient Flows



$$x'(t) = -\nabla F(x(t))$$

Wasserstein Gradient Flows (I)

Take the manifold

$$\mathcal{P}_2^{ac}(\mathbb{R}^d) = \left\{ \mu \in \mathcal{P}(\mathbb{R}^d) : \int \|x\|_2^2 \, d\mu(x) < \infty, \mu << \mathcal{L} \right\}$$

with distance

$$W_2(\mu, \nu) := \left(\inf_{\pi \in \Pi(\mu,\nu)} \int \|x - y\|_2^2 \, d\pi(x, y) \right)^{1/2}$$

Wasserstein Gradient Flows (II)

Curves in $\left(\mathcal{P}_2^{ac}(\mathbb{R}^d), W_2\right)$ are geodetics

$$\mu_s = ((1-s)Id + st_\mu^\nu)_\# \mu$$

where t_μ^ν is the unique transport map between μ and ν and $T_\# \mu$ denotes the push-forward measure

$$T_\# \mu(A) = \mu(T^{-1}(A)).$$

Wasserstein Gradient Flows (III)

For $F : M \subset \mathbb{R}^d \mapsto \mathbb{R}$ the gradient flow equation is

$$x'(t) = -\nabla F(x(t)).$$

For $F : \mathcal{P}_2^{ac}(\mathbb{R}^d) \mapsto \mathbb{R}$ we have[1]

$$\partial_t \rho_t = -\nabla \cdot \left(\rho_t \frac{\delta F}{\delta \rho_t} \right) \tag{1}$$

with $\frac{\delta F}{\delta \rho_t} := \frac{d}{d\epsilon} F(\rho + \epsilon \chi)_{|\epsilon=0}$.

[1] Jordan, Kinderlehrer, Otto (1998)

When is (1) well-behaved?

We need:

- F continuous, $F(\rho) < +\infty$ for some ρ
- F $\lambda-$convex: exists $\lambda \geq 0$ s.t.

$$F\left(((1-s)Id + st_{\mu}^{\nu})_{\#}\mu\right) \leq (1-s)F(\mu) + sF(\nu) - \frac{\lambda}{2}s(1-s)W_2^2(\nu,\mu)$$

- F is coercive: exists $r > 0$ s.t.

$$\inf\left\{F(\rho) : \rho \in \mathcal{P}_2^{ac}(\mathbb{R}^d),\ \int \|x\|_2^2 \, d\rho(x) \leq r\right\} > -\infty$$

Then...

The gradient flow PDE (1) has a unique solution for a given initial condition ρ_0.

For two initial conditions ρ_0^1, ρ_0^2 we have the estimate

$$W_2(\rho_t^1, \rho_t^2) \leq e^{-\lambda t} W_2(\rho_0^1, \rho_0^2).$$

Gradient Flow for FIE - Assumptions

$$\mu(dy) = \int_{\mathbb{X}} \rho(dx) K(x, y) \qquad \forall y \in \mathbb{Y},$$

1. $\mathbb{X}, \mathbb{Y}$ subsets of some Euclidean spaces with Borel σ-algebras;
2. μ, ρ densities with finite second moment;
3. K bounded and bounded away from 0, Lipschitz continuous with Lipschitz continuous gradient and λ-concave in x:

$$x \mapsto K(x, y) + \frac{\lambda}{2} \|x\|_2^2$$

is concave for some $\lambda > 0$ for all $y \in \mathbb{Y}$.

Gradient Flow for FIE

$$F(\rho) := \mathrm{KL}(\mu, \rho K) - \alpha\, \mathrm{ent}(\rho)$$

is continuous, coercive and λ geodesically convex with $\lambda = 0$ and has first variation

$$\frac{\delta F}{\delta \rho}(x) = -\int \mu(dy)\, \frac{K(x,y)}{\rho K(y)} + \alpha\left(1 + \log \rho(x)\right).$$

⇨ the gradient flow exist and is unique for each initial condition ρ_0

Gradient Flow for FIE - PDE

$$
\begin{aligned}
\partial_t \rho_t &= \nabla \cdot \left(\rho_t \nabla \frac{\delta F}{\delta \rho_t} \right) \\
&= \nabla \cdot \left(\rho_t \left[- \int \mu\,(dy)\, \frac{\nabla K(x,y)}{\rho K(y)} + \alpha \nabla \log \rho\,(x) \right] \right) \\
&= -\nabla \cdot \left(\rho_t \int \mu\,(dy)\, \frac{\nabla K(x,y)}{\rho_t K(y)} \right) + \alpha \triangle \rho_t
\end{aligned}
$$

is a Fokker-Plank equation with corresponding SDE...

Gradient Flow for FIE - SDE

$$dX_t = \int \mu(dy) \frac{\nabla K(X_t, y)}{\rho_t K(y)} dt + \sqrt{2\alpha} dW_t, \quad X_0 \sim \rho_0$$

- McKean-Vlasov SDE
- a strong solution exists and is unique
- requires discretisation in time and space

SDE - Implementation

- discretisation in space: take N copies of the SDE and use the resulting empirical measure instead of ρ_t

$$dX_t^{i,N} = \int \mu(dy) \frac{\nabla K(X_t^{i,N}, y)}{\rho_t^N K(y)} dt + \sqrt{2\alpha} dW_t^i, \quad \rho_t^N = \frac{1}{N} \sum_{i=1}^{N} \delta_{X_t^{i,N}}$$

- discretisation in time: Euler scheme

$$X_{n+1}^{i,N} = X_n^{i,N} + \int \mu(dy) \frac{\nabla K(X_n^{i,N}, y)}{\rho_n^N K(y)} \Delta t + \sqrt{2\alpha} \Delta W^i,$$

$$\rho_n^N = \frac{1}{N} \sum_{i=1}^{N} \delta_{X_n^{i,N}}$$

Toy Example

$$\mu(dy) = \int_{\mathbb{X}} \rho(dx) K(x, y),$$

with

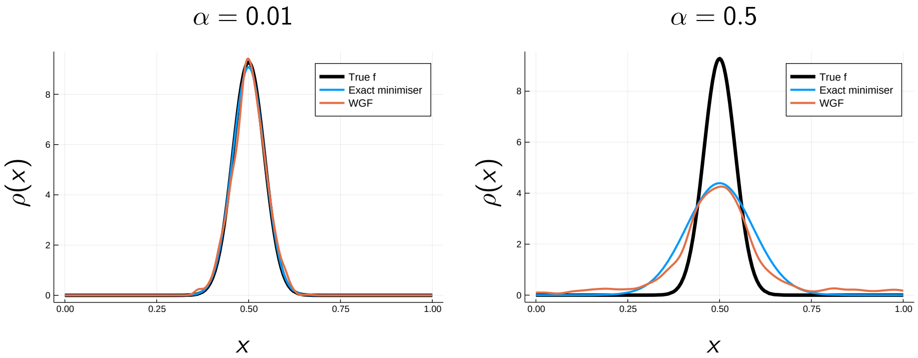
$$\begin{aligned} \mu(y) &= \mathcal{N}(y; m, \sigma_\mu^2 := \sigma_K^2 + \sigma_\rho^2) \\ K(x, y) &= \mathcal{N}(y; x, \sigma_K^2) \\ \rho(x) &= \mathcal{N}(x; m, \sigma_\rho^2). \end{aligned}$$

The unique minimiser of E is

$$\rho : \alpha(x) = \mathcal{N}(x; m, \sigma_\alpha^2)$$

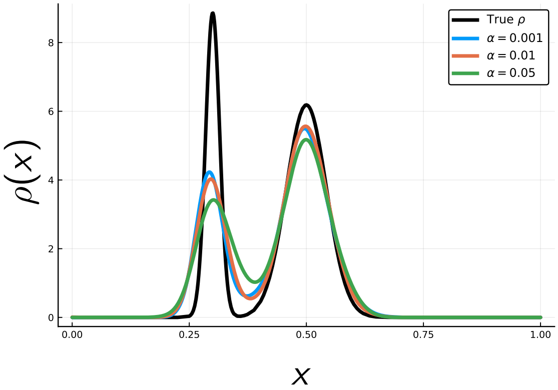
for $\alpha \in [0, 1)$.

Toy Example



$\alpha = 0.01$
$\alpha = 0.5$
True f
Exact minimiser
WGF
$\rho(x)$
x
0
2
4
6
8
0.00
0.25
0.50
0.75
1.00

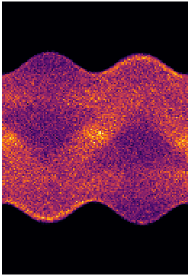
A second Toy Example



PET

The reconstruction of a cross-section of the brain (Figure (a)) from the data image provided by PET scanners (Figure (b)) is described by a 2D Fredholm integral equation of the first kind.

(a) 128-pixels Shepp-Logan phantom



(b) Sinogram + noise