Regularized Stein Variational Gradient Descent

-Talk in BIRS Stein's method workshop

Ye He, University of California Davis

joint work with Krishnakumar Balasubramanian, Bharath K Sriperumbudur and Jianfeng Lu

Apr.12, 2022

Motivation

LMC $\longleftrightarrow$ SVGD

Content

The talk includes discussion of 4 parts:

- Prelimilary on RKHS
- A regularized formulation of SVGD
- Some analysis on the regularized SVGD
- Future work

Prelimilary on RKHS

Introduction to the Reproducing kernel Hilbert Space(RKHS)

[Steinwart, SVM]

An RKHS, $\mathcal{H}$, is a functional space over $\mathbb{R}^d$ s.t.

$$\delta_x : \mathcal{H} \to \mathbb{R} \quad \text{is continuous for } \forall x \in \mathbb{R}^d.$$
$$f \mapsto f(x)$$

A reproducing kernel, $k : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ is symmetric, positive s.t.

$$f(x) = \langle f, k(\cdot, x) \rangle_{\mathcal{H}}$$

Remark: $\mathcal{H}$ and k are uniquely determined by each other.

$\Rightarrow \mathcal{H}_k$.

Prelimilary on RKHS

The integral operator and its adjoint

$\forall$ prob. measure μ, assume

$$\|k\|_{L^2(\mu)} := \left(\int_{\mathbb{R}^d} k(x,x)\,\mu(dx)\right)^{\frac{1}{2}} < +\infty .$$

$i_{k,\mu} : \mathcal{H}_k \to L_2(\mu)$ is an inclusion operator.

$$i_{k,\mu}^* : L_2(\mu) \to \mathcal{H}_k$$

$$f \mapsto \int_{\mathbb{R}^d} f(x) k(\cdot, x)\,\mu(dx)$$

- $i_{k,\mu}$ and $i_{k,\mu}^*$ are adjoint to each other.

$\forall f \in L_2(\mu), g \in \mathcal{H}_k$,

$$\langle f, i_{k,\mu} g \rangle_{L_2(\mu)} = \int_{\mathbb{R}^d} f(x) g(x)\,\mu(dx) = \int_{\mathbb{R}^d} f(x) \langle g, k(\cdot, x)\rangle_{\mathcal{H}_k}\,\mu(dx)$$

$$= \langle i_{k,\mu}^* f, g \rangle_{\mathcal{H}_k} .$$

Prelimilary on RKHS

Properties of the operators

$$T_{k,\mu} := i_{k,\mu} i_{k,\mu}^* : L_2(\mu) \to L_2(\mu)$$

$$S_{k,\mu} := i_{k,\mu}^* i_{k,\mu} : \mathcal{H}_k \to \mathcal{H}_k$$

- Both self-adjoint, positive, compact.
- Share same nonzero eigenvalues and eigenvectors.
- When $k \in C(\mathbb{R}^d, \mathbb{R}^d)$ and $\mathrm{supp}(\mu) = \mathbb{R}^d$, $S_{k,\mu}$ is strictly positive.
- When k is integrably strictly positive definite (ISPD), $T_{k,\mu}$ strictly positive.

$$\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} k(x,y)\,\rho(dx)\,\rho(dy) > 0, \quad \forall \text{ finite Borel signed measure } \rho.$$

Prelimilary on RKHS

Interpolation spaces between RKHS and L^2

- $\mathcal{H}_K \subset L_2(\mu)$ is dense
  - $\nearrow$ $\mathcal{H}_K \subset L_2(\mu)$ b/c $i_{K,\mu}$ is an inclusion.
  - $\searrow$ ISPD $\Rightarrow$ $i_{K,\mu}^*$ is injective.

- $T_{K,\mu}: L_2(\mu) \longrightarrow L_2(\mu)$

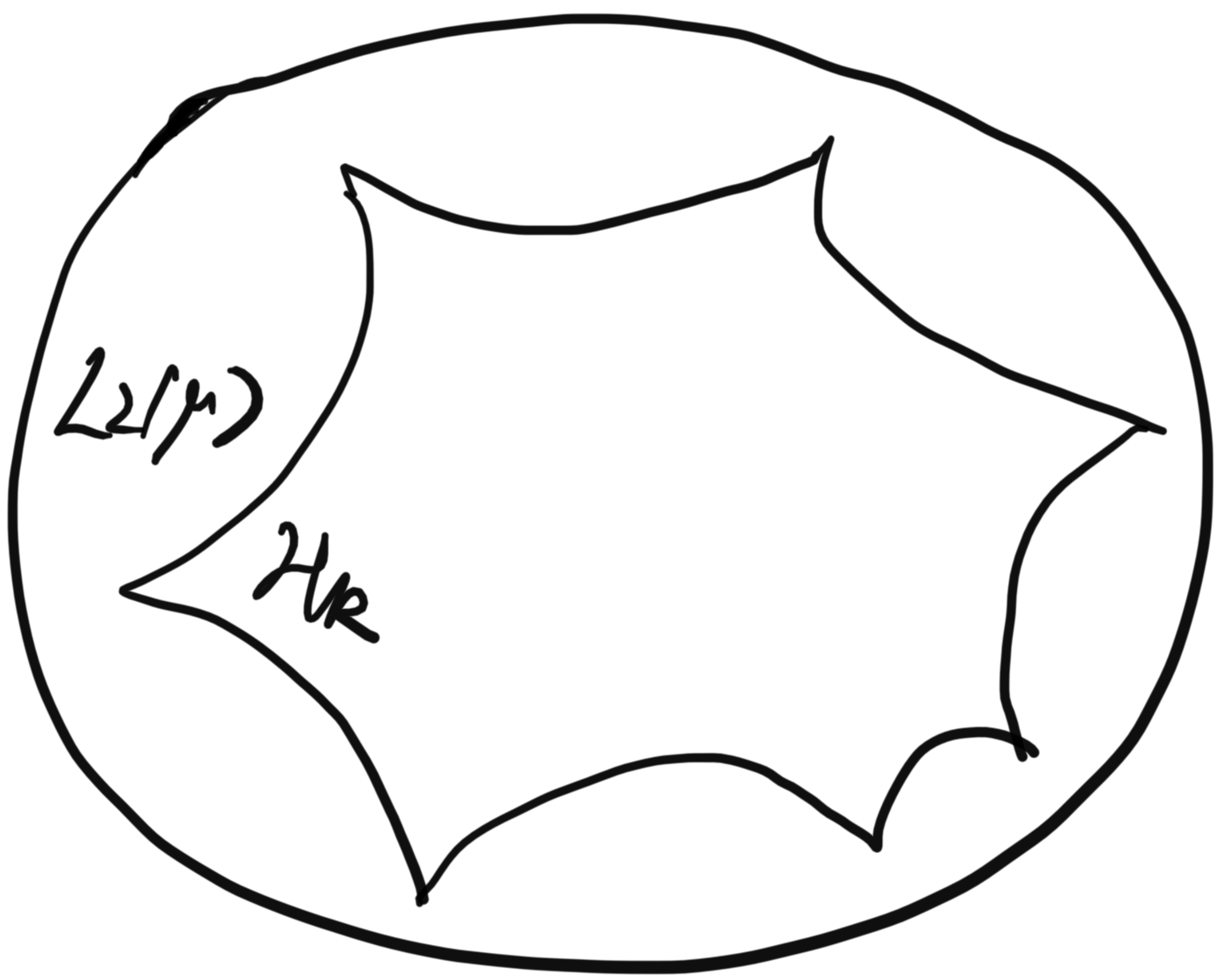
  $T_{K,\mu} = \sum_{i=1}^{\infty} \lambda_i \langle \ , e_i \rangle_{L_2(\mu)} e_i$ with $\lambda_1 \geq \cdots \geq \lambda_i \geq \cdots > 0$, $\{e_i\}_{i=1}^{\infty}$ O.N.B.

  $\forall f \in L_2(\mu)$, $\exists \gamma \in (0, \tfrac{1}{2}]$ s.t.

  $f = T_{K,\mu}^{\gamma} h$ with $h \in L_2(\mu)$

  - $\gamma = 0$, trivial.
  - $\gamma = \tfrac{1}{2}$, $f = T_{K,\mu}^{\frac{1}{2}} h = \sum_{i=1}^{\infty} \sqrt{\lambda_i} \langle h, e_i \rangle_{L_2(\mu)} e_i$

    $\in \mathcal{H}_K$.



Stein variational formulation

<u>Goal</u>: to generate samples following $\pi \propto e^{-V}$.

- <u>optimization formulation</u>:

$$x \longmapsto x + \varepsilon \phi(x) := T(x), \quad x \in \mathbb{R}^d, x \sim \rho, \phi: \mathbb{R}^d \to \mathbb{R}^d$$

- to find the optimal ϕ^* s.t. $KL(T_\# \rho | \pi)$ decays fastest.

In [Q.Liu et al 2016],

$$\phi^* = \underset{\phi \in \mathcal{F}}{\operatorname{argmax}} \left\{ \mathbb{E}_{x \sim \rho}[S_\pi \phi(x)], \text{ constraints on } \phi \right\}$$

where $S_\pi \phi(x) := -\nabla V(x) \cdot \phi(x) + \nabla \cdot \phi(x)$

↑ Stein operator.

Stein variational formulation

-Langevin dynamics

Choose $\mathcal{F} = L_2^d(\rho)$, constraint: $\|\phi\|_{L_2^d(\rho)} \leq 1$.

We get

$$\phi^* \propto -\nabla V(x) - \frac{\nabla \rho(x)}{\rho(x)} = -\nabla \log \frac{\rho}{\pi}.$$

and

$$x \longmapsto T(x) = x - \varepsilon \nabla \log \frac{\rho}{\pi}$$

Let $\varepsilon \to 0^+$,

$$\partial_t \rho = \nabla \cdot \left(\rho \nabla \log \frac{\rho}{\pi} \right)$$

Fokker Planck equation to Langevin dynamics.

Stein variational formulation
-SVGD

Choose $\mathcal{F} = \mathcal{H}_k^d$, constraint : $\|\phi\|_{\mathcal{H}_k^d} \leq 1$

In [Q. Liu et 2016],

$$\phi^* \propto - i_{k,\rho}^* \nabla \log \frac{\rho}{\pi}$$

and

$$x \longmapsto T(x) = x - \varepsilon\, i_{k,\rho}^* \nabla \log \frac{\rho}{\pi}(x)$$

Let $\varepsilon \to 0^+$, in [Lu et al 2018],

$$\partial_t \rho = \nabla \cdot \left(\rho\, i_{k,\rho}^* \nabla \log \frac{\rho}{\pi} \right).$$

Stein variational formulation

Comparison between Langevin and SVGD

Langevin : $\partial_t \rho = \nabla \cdot \left(\rho \nabla \log \frac{\rho}{\pi} \right)$

SVGD : $\partial_t \rho = \nabla \cdot \left(\rho \, i_{k,\rho}^{*} \nabla \log \frac{\rho}{\pi} \right)$

- Observe that,

$$\nabla \log \frac{\rho}{\pi} = I \left(\nabla \log \frac{\rho}{\pi} \right) \quad \text{with} \quad I : L_2^d(\rho) \to L_2^d(\rho), \ f \mapsto f.$$

$$i_{k,\rho}^{*} \nabla \log \frac{\rho}{\pi} = i_{k,\rho} i_{k,\rho}^{*} \nabla \log \frac{\rho}{\pi} = T_{k,\rho} \nabla \log \frac{\rho}{\pi} \in L_2^d(\rho)$$

- To interpolate between Langevin and SVGD

$I : L_2^d(\rho) \to L_2^d(\rho)$

$T_{k,\rho} : L_2^d(\rho) \to L_2^d(\rho)$

$$\boxed{(T_{k,\rho} + \nu I)^{-1} T_{k,\rho} : L_2^d(\rho) \to L_2^d(\rho), \ \nu > 0}$$

Stein variational formulation

-regularized SVGD

Choose $\mathcal{F} = \mathcal{H}_k^d$, constraint: $\nu\|\phi\|^2_{\mathcal{H}_k^d} + \|\phi\|^2_{L_2^d(\rho)} \leq 1$ [Krishnakumar et al 2017]

- $\nu\|\phi\|^2_{\mathcal{H}_k^d} + \|\phi\|^2_{L_2^d} = \langle \phi, (i_{k,\rho}^* i_{k,\rho} + \nu I_k)^{-1} i_{k,\rho}^* \phi \rangle_{\mathcal{H}_k^d}$
- $\mathbb{E}_{x\sim\rho}[S_\pi \phi(x)] = \langle \phi, i_{k,\rho}^* \nabla \log \frac{\rho}{\pi} \rangle_{\mathcal{H}_k^d}$

We get

$$\phi^* \propto (i_{k,\rho}^* i_{k,\rho} + \nu I_k)^{-1} i_{k,\rho}^* \nabla \log \frac{\rho}{\pi} \in \mathcal{H}_k^d \subset L_2^d(\rho)$$

$$\phi^* \propto i_{k,\rho} (i_{k,\rho}^* i_{k,\rho} + \nu I_k)^{-1} i_{k,\rho}^* \nabla \log \frac{\rho}{\pi} \in L_2^d(\rho)$$

<u>Claim</u>: $i_{k,\rho} (i_{k,\rho}^* i_{k,\rho} + \nu I_k)^{-1} i_{k,\rho}^* = (i_{k,\rho} i_{k,\rho}^* + \nu I)^{-1} i_{k,\rho} i_{k,\rho}^*$

$$\Rightarrow \phi^* \propto (T_{k,\rho} + \nu I)^{-1} T_{k,\rho} \nabla \log \frac{\rho}{\pi}.$$

Stein variational formulation

-regularized SVGD

- Population limit:

$$\rho^{n+1} = (I_d - h_{n+1}(\iota_{k,\rho^n}\iota_{k,\rho^n}^* + \nu I_d)^{-1}\iota_{k,\rho^n}\iota_{k,\rho^n}^*\nabla \log \frac{\rho^n}{\pi})_{\#\rho^n}. \tag{1}$$

- Mean-field PDE:

$$\begin{cases} \partial_t \rho = \nabla \cdot \left(\rho \, (\iota_{k,\rho}\iota_{k,\rho}^* + \nu I_d)^{-1}\iota_{k,\rho}\iota_{k,\rho}^*(\nabla \log \frac{\rho}{\pi})\right), \\ \rho(0,\cdot) = \rho_0(\cdot). \end{cases} \tag{2}$$

- Finite particle system: $(X_n^i)_{i=1}^N$ are the N-particles at step n. $\bar{X}_n := [X_n^1, \cdots, X_n^N]^T$. For all $f : \mathbb{R}^d \to \mathbb{R}^d$, $L_n f := [f(X_n^1), \cdots, f(X_n^N)]^T$.

$$\bar{X}_{n+1} = \bar{X}_n - h_{n+1}(\frac{1}{N}K_n + \nu I_N)^{-1}(\frac{1}{N}K_n L_n(\nabla V) - \frac{1}{N}\sum_{j=1}^N L_n(\nabla k(X_n^j, \cdot))). \tag{3}$$

where $K_n \in \mathbb{R}^{N\times N}$ is the Gram matrix with $(K_n)_{i,j} = k(X_n^i, X_N^j)$.

Analysis on the regularized SVGD

the mean-field PDE

$$\begin{cases} \partial_t \rho = \nabla\cdot\left(\rho\,(T_{k,\rho}+\nu I)^{-1} T_{k,\rho}\nabla \log \frac{\rho}{\pi}\right) & (1) \\ \rho(0,x) = \rho_0(x). \end{cases}$$

- Existence and uniqueness of weak solution
- Stability in W_p
- KL-divergence decay / Fisher information convergence

Analysis on the regularized SVGD

the mean-field PDE

Existence and uniqueness of weak solution

- weak solution, $\rho \in C([0,\infty), \mathcal{P}_V)$ s.t. for all $\varphi \in C_0^\infty([0,\infty)\times\mathbb{R}^d)$,

$$\int_0^{+\infty}\int_{\mathbb{R}^d} \partial_t\varphi(t,x) - \langle \nabla\varphi(t,x), (T_{k,\rho}+\nu I)^{-1} T_{k,\rho}\nabla\log\frac{\rho}{\pi}(x)\rangle \rho(t,dx)\,dt + \int_{\mathbb{R}^d}\varphi(0,x)\rho_0(dx) = 0$$

where $\mathcal{P}_V = \{ \rho \in \mathcal{P},\ \|\rho\|_{\mathcal{P}_V} := \int_{\mathbb{R}^d} 1+V(x)\,\rho(dx) < +\infty \}$

- Theorem. Under assumptions on k and V, for $\forall\ \rho_0 \in \mathcal{P}_V$, there exists a unique weak solution $\rho \in C([0,\infty), \mathcal{P}_V)$ to the PDE (1). Moreover for all $t \geq 0$,

$$\|\rho(t,\cdot)\|_{\mathcal{P}_V} \leq \|\rho_0\|_{\mathcal{P}_V} \exp\left(C t^{\frac{1}{2}} \right).$$

$\uparrow$ depending on k, V, ρ_0, ν

Analysis on the regularized SVGD

the mean-field PDE

Idea of proof:

(1) Characteristic gradient flow induced by (1):

$$\begin{cases} \frac{d}{dt}\Phi(t,x,\rho_0) = -\left(T_{k,\rho_t}+\nu I\right)^{-1} T_{k,\rho_t} \nabla \log\frac{\rho_t}{\pi}\left(\Phi(t,x,\rho_0)\right) \\ \rho_t = \left(\Phi(t,\cdot,\rho_0)\right)\#\rho_0 \\ \Phi(0,x,\rho_0) = x \end{cases} \qquad (2)$$

$$\Rightarrow \Phi(t,x,\rho_0) = x - \int_0^t \left(T_{k,\rho_s}+\nu I\right)^{-1} T_{k,\rho_s} \nabla \log\frac{\rho_s}{\pi}\left(\Phi(s,x,\rho_0)\right) ds.$$

$$:= \mathcal{F}[\Phi](t,x)$$

(2) Show $\mathcal{F}$ is a contraction from Y to Y. $\rightsquigarrow$ some reasonable Banach space.

(3) local $\to$ global.

Stability in W_p

- Theorem. Under assumptions on k and V, let $\rho_1, \rho_2 \in \mathcal{P}$ be two initial conditions to PDE (1) s.t. $\|\rho_i\|_p := \int_{\mathbb{R}^d} |x|^p \rho_i(dx) < +\infty$, $i=1,2$, $p \in (1, +\infty)$. Let $\rho_1(t,x)$ and $\rho_2(t,x)$ be the two corresponding weak solutions, we have

$$\sup_{t\in[0,T]} W_p(\rho_1(t,\cdot), \rho_2(t,\cdot)) \le C \, W_p(\rho_1, \rho_2)$$

$\uparrow$

depending on $k, V, \rho_1, \rho_2, T, \nu$

Analysis on the regularized SVGD

the mean-field PDE

KL-divergence Decay

Let $\rho_t(x)$ be the weak solution to PDE (1).

- $$\frac{d}{dt}\mathrm{KL}(\rho_t|\pi) = -\left\langle i^*_{k,\rho_t}\nabla\log\frac{\rho_t}{\pi},\ \left(i^*_{k,\rho_t} i_{k,\rho_t} + \nu I_k\right)^{-1} i^*_{k,\rho_t}\nabla\log\frac{\rho_t}{\pi}\right\rangle_{\mathcal{H}_k^d}$$
$$:= I_{\nu,\mathrm{Stein}}(\rho_t|\pi) \leq 0$$

- **Properties of $I_{\nu,\mathrm{Stein}}(\mu|\pi)$**

(1) $$I_{\nu,\mathrm{Stein}}(\mu|\pi) = \sum_{i=1}^{\infty}\frac{\lambda_i}{\lambda_i+\nu}\left|\left\langle\nabla\log\frac{\mu}{\pi}, e_i\right\rangle_{L_2(\mu)}\right|^2$$

where

$$\lambda_1 \geq \cdots \geq \lambda_i \geq \cdots > 0$$

$\{e_i\}_{i=1}^{\infty}$ O.N.B of $L_2(\mu)$.

Analysis on the regularized SVGD

the mean-field PDE

Properties of $I_{\nu,\mathrm{Stein}}(\mu|\pi)$:

(2) Approximation to the Fisher information.

- $$I(\mu|\pi) := \int_{\mathbb{R}^d} \left|\nabla \log \frac{\mu}{\pi}(x)\right|^2 \mu(dx)$$
$$= \sum_{i=1}^{\infty} \left| \langle \nabla \log \frac{\mu}{\pi}, e_i \rangle_{L_2(\mu)} \right|^2 \geq I_{\nu,\mathrm{Stein}}(\mu|\pi)$$

- $$I(\mu|\pi) = I_{\nu,\mathrm{Stein}}(\mu|\pi) + \sum_{i=1}^{\infty} \frac{\nu}{\lambda_i+\nu} \left| \langle \nabla \log \frac{\mu}{\pi}, e_i \rangle_{L_2(\mu)} \right|^2$$

Recall: $\nabla \log \frac{\mu}{\pi} \in L_2(\mu)$, $\exists\, \gamma \in (0, \frac{1}{2}]$ s.t. $\nabla \log \frac{\mu}{\pi} = T_{k,\mu}^{\gamma} h$, $h \in L_2(\mu)$.

$$\Rightarrow I(\mu|\pi) = I_{\nu,\mathrm{Stein}}(\mu|\pi) + \sum_{i=1}^{\infty} \frac{\nu \lambda_i^{2\gamma}}{\lambda_i+\nu} \left| \langle h, e_i \rangle_{L_2(\mu)} \right|^2$$
$$\leq I_{\nu,\mathrm{Stein}}(\mu|\pi) + \nu^{2\gamma} \|h\|_{L_2(\mu)}^2$$

$$\Rightarrow \boxed{I(\mu|\pi) - \nu^{2\gamma} \|h\|_{L_2(\mu)}^2 \leq I_{\nu,\mathrm{Stein}}(\mu|\pi) \leq I(\mu|\pi).}$$

Decay of KL-divergence

Assume $\pi \sim LSI(\lambda)$, i.e. $KL(\mu|\pi) \le \frac{1}{2\lambda} I(\mu|\pi)$, we have

$$\frac{d}{dt} KL(\rho_t|\pi) \le -I(\rho_t|\pi) + \nu^{2\gamma t} R(t)^2 \qquad R(t) = \| T_{k,\rho_t}^{-\gamma t} \nabla \log \frac{\rho_t}{\pi} \|_{L_2(\rho_t)} < +\infty$$

$$\le -2\lambda KL(\rho_t|\pi) + \nu^{2\gamma t} R(t)^2$$

$$\Rightarrow \boxed{KL(\rho_t|\pi) \le e^{-2\lambda t} KL(\rho_0|\pi) + \int_0^t \nu^{2\gamma s} R(s)^2 e^{2\lambda(s-t)} ds}$$

Convergence of Fisher information

$$\boxed{\int_0^{+\infty} I(\rho_t|\pi)\, dt \le KL(\rho_0|\pi) + \int_0^{+\infty} \nu^{2\gamma t} R(t)^2 dt}$$

Analysis on the regularized SVGD

Comparison to Langevin dynamics

Let $\mu_t(x)$ be the solution to the FPE to Langevin dynamics with initial condition $\mu_0(x)$.

Assume $\pi \sim LSI(\lambda)$ and we choose μ_0 s.t. $\mu_t \sim LSI(\lambda)$ for $\forall t \geq 0$, then we get

$$\frac{d}{dt} KL(\rho_t | \mu_t) \leq -\frac{3\lambda}{2} KL(\rho_t | \mu_t) + \nu^{2\gamma t} R(t)^2$$

$$\Rightarrow \boxed{KL(\rho_t | \mu_t) \leq e^{-\frac{3\lambda t}{2}} KL(\rho_0 | \mu_0) + \int_0^t \nu^{2\gamma s} R(s)^2 e^{\frac{3\lambda}{2}(s-t)} ds}$$

Analysis on the regularized SVGD

Decay of KL-divergence along the population

Recall:

$$\rho^{n+1} = (I_d - h_{n+1}(\iota_{k,\rho^n}\iota^*_{k,\rho^n} + \nu_{n+1} I_d)^{-1}\iota_{k,\rho^n}\iota^*_{k,\rho^n} \nabla \log \frac{\rho^n}{\pi})_{\#\rho^n}.$$

Assumption A1

(1) *There exists $B > 0$ such that for all $x \in \mathbb{R}^d$: $\|\nabla k(x, \cdot)\|_{\mathcal{H}_k^d} \leq B$.*

(2) *The potential function $V : \mathbb{R}^d \to \mathbb{R}$ is twice continuous differentiable and gradient Lipschitz with parameter L.*

(3) *Along the population limit, $I(\rho^n|\pi) < \infty$ for all fixed $n \geq 0$.*

Remark: similar assumptions in [Anna, Adil et al 2021].

Analysis on the regularized SVGD

Decay of KL-divergence along the population

Theorem 1

Assume that π satisfies the log-Sobolev inequality with parameter λ. Under the A1, Let (ρ^n) be the population limit of regularized SVGD described in (1) *with initial condition $\rho^0 = \rho_0$ such that $KL(\rho_0|\pi) \leq R$. By choosing ν_{n+1} and the step-size h_{n+1} such that for all $n \geq 0$:*

$$\nu_{n+1} \leq \left(\frac{I(\rho^n|\pi)}{2 \left\| (\iota_{k,\rho^n} \iota^*_{k,\rho^n})^{-\gamma_n} \nabla \log \frac{\rho^n}{\pi} \right\|^2_{L^d_2(\rho^n)}} \right)^{\frac{1}{2\gamma_n}},$$

$$h_{n+1} < \min \left\{ L^{-1}, \sqrt{2} B^{-1} (\alpha - 1) \alpha^{-1} \nu_{n+1}^{\gamma_n} I(\rho^n|\pi)^{-\frac{1}{2}} \left(\sup_i \frac{\lambda_i^{(n)\,1+2\gamma_n}}{(\lambda_i^{(n)} + \nu_{n+1})^2} \right)^{-\frac{1}{2}}, \right.$$
$$\left. \frac{1}{2} B^{-2} \alpha^2 \nu_{n+1}^{2\gamma_n} (\sup_i \frac{\lambda_i^{(n)\,1+2\gamma_n}}{(\lambda_i^{(n)} + \nu_{n+1})^2})^{-1}, 4\lambda^{-1} \right\}.$$

where for each n, $\gamma_n \in (0, \frac{1}{2}]$ and $(\iota_{k,\rho^n} \iota^*_{k,\rho^n})^{-\gamma_n} \nabla \log \frac{\rho^n}{\pi} \in L^d_2(\rho^n)$. $\{\lambda_i^{(n)}\}$ is the sequence of positive eigenvalues of the operator $\iota_{k,\rho^n} \iota_{k,\rho^n}*$ in the order of decreasing values. α is some constant and $\alpha \in (1, 2)$. Then for all $n \geq 1$,

$$KL(\rho^n|\pi) \leq \Pi_{i=1}^n (1 - \frac{1}{4} \lambda h_i) R \tag{4}$$

Analysis on the regularized SVGD

Convergence of Fisher information along the population

Theorem 2

Under the A1, Let (ρ^n) be the population limit of regularized SVGD described in (1) with initial condition $\rho^0 = \rho_0$ such that $KL(\rho_0|\pi) \leq R$. By choosing ν_{n+1} and the step-size h_{n+1} such that for all $n \geq 0$:

$$h_{n+1} < \min\left\{ L^{-1}, \sqrt{2}B^{-1}(\alpha-1)\alpha^{-1}\nu_{n+1}^{\gamma_n} I(\rho^n|\pi)^{-\frac{1}{2}} \left(\sup_i \frac{{\lambda_i^{(n)}}^{1+2\gamma_n}}{(\lambda_i^{(n)}+\nu_{n+1})^2} \right)^{-\frac{1}{2}} \right\}.$$

where for each n, $\gamma_n \in (0, \frac{1}{2}]$, $(\iota_{k,\rho^n}\iota_{k,\rho^n}^*)^{-\gamma_n}\nabla \log \frac{\rho^n}{\pi} \in L_2^d(\rho^n)$ and $R_n := \left\|(\iota_{k,\rho^n}\iota_{k,\rho^n}^*)^{-\gamma_n}\nabla \log \frac{\rho^n}{\pi}\right\|_{L_2^d(\rho^n)}$.

$\{\lambda_i^{(n)}\}$ is the sequence of positive eigenvalues of the operator $\iota_{k,\rho^n}\iota_{k,\rho^n}*$ in the order of decreasing values. α is some constant and $\alpha \in (1,2)$. Then for all $n \geq 1$,

$$\sum_{n=0}^{\infty} \frac{h_{n+1}}{2} I(\rho^n|\pi) \leq \sum_{n=0}^{\infty} \nu_{n+1}^{2\gamma_n} h_{n+1} \left(1 + \frac{1}{2}\nu_{n+1}^{-\frac{1}{2}}\alpha^2 B^2 h_{n+1}\right) R_n^2 + R.$$

Future Work

- To analyze on the finite particle system