On Nonconvex SVGD

Lukang Sun

March 25, 2021

1 Setup

$$F = \frac{1}{n}\sum_{i=1}^{n} f_i, \qquad \pi \sim e^{-F}$$

$$f_n = \frac{1}{n_0}\sum_{i\in I} f_i, \qquad \#I = n_0, \qquad \sigma_n \sim e^{-f_n}$$

$$\phi_t(x) = \left(I - t P_{\mu_n}\nabla\log\left(\frac{\mu_n}{\sigma_n}\right)\right)(x), \qquad \mu_t = (\phi_t)_{\#}\,\mu_n$$

$$\varphi(t) = \int \log\left(\frac{\mu_t}{\pi}\right) d\mu_t$$

2 PAGE

We define the update rule as:

$$S_{\mu_{n+1}}\nabla\log\left(\frac{\mu_{n+1}}{\sigma_{n+1}}\right) = \begin{cases} S_{\mu_{n+1}}\nabla\log\left(\frac{\mu_{n+1}}{\pi}\right) & p, \\ S_{\mu_n}\nabla\log\left(\frac{\mu_n}{\sigma_n}\right) + S_{\mu_{n+1}}\nabla\log\left(\frac{\mu_{n+1}}{\sigma_{I'}}\right) - S_{\mu_n}\nabla\log\left(\frac{\mu_n}{\sigma_{I'}}\right) & 1-p, \end{cases}$$

where $\sigma_{I'} \sim \exp^{-f_{I'}}$, $f_{I'} = \frac{1}{b'}\sum_{i\in I'} f_i$.

3 Main Proof

$$\phi(\gamma) \le \phi(0) + \gamma\left\langle S_{\mu_n}\nabla\log\frac{\mu_n}{\pi}, S_{mu_n}\nabla\log\left(\frac{\mu_n}{\sigma_n}\right)\right\rangle_H + \frac{C\cdot\gamma^2}{2}\left\langle S_{\mu_n}\nabla\log\frac{\mu_n}{\sigma_n}, S_{\mu_n}\nabla\log\left(\frac{\mu_n}{\sigma_n}\right)\right\rangle_H$$

then by similar calculation as in PAGE, we have,

$$\varphi(\gamma) \le \varphi(0) - \frac{\gamma}{2}\left\|S_{\mu_n}\nabla\log\left(\frac{\mu_n}{\pi}\right)\right\|_H^2 - \left(\frac{1}{2\gamma}-\frac{c}{2}\right)\cdot\gamma^2\left\|S_{\mu_n}\nabla\log\left(\frac{\mu_n}{\sigma_n}\right)\right\|_H^2$$

$$+ \frac{\gamma}{2}\left\|S_{\mu_n}\nabla\log\left(\frac{\mu_n}{\pi}\right) - S_{\mu_n}\log\left(\frac{\mu_n}{\sigma_n}\right)\right\|_H^2$$

Let

$$G_{n+1} := \mathbf{E}\left[\left\|S_{\mu_{n+1}}\nabla\log\left(\frac{\mu_{n+1}}{\pi}\right) - S_{\mu_{n+1}}\nabla\log\left(\frac{\mu_{n+1}}{\sigma_{n+1}}\right)\right\|_H^2\right]$$

Then a direct calculation now reveals that

$$G_{n+1} = p \cdot 0 + (1-p) \cdot \mathrm{E} \left\| S_{\mu_n} \nabla \log\left(\frac{\mu_n}{\sigma_n}\right) + S_{\mu_{n+1}} \nabla \log\left(\frac{\mu_{n+1}}{\sigma_{I'}}\right) - S_{\mu_n} \nabla \log\left(\frac{\mu_n}{\sigma_{I'}}\right) - S_{\mu_{n+1}} \nabla \log\left(\frac{\mu_{n+1}}{\pi}\right) \right\|_H^2$$

$$= (1-p) \cdot \mathrm{E} \left\| S_{\mu_n} \nabla \log\left(\frac{\mu_n}{\sigma_n}\right) + S_{\mu_{n+1}} \nabla \log\left(\frac{\mu_{n+1}}{\sigma_{I'}}\right) - S_{\mu_n} \nabla \log\left(\frac{\mu_n}{\sigma_{I'}}\right) - S_{\mu_{n+1}} \nabla \log\left(\frac{\mu_{n+1}}{\pi}\right) + S_{\mu_n} \nabla \log\left(\frac{\mu_n}{\pi}\right) \right\|_H^2$$

$$+ (1-p) \cdot \left\| S_{\mu_n} \nabla \log\left(\frac{\mu_n}{\pi}\right) - S_{\mu_n} \nabla \log\left(\frac{\mu_n}{\sigma_I}\right) \right\|_H^2$$

4 Key Inequality

We need to estimate

$$\left\| S_{\mu_{n+1}} \nabla \log\left(\frac{\mu_{n+1}}{\pi}\right) - S_{\mu_n} \nabla \log\left(\frac{\mu_n}{\pi}\right) \right\|_H^2$$

and

$$\left\| S_{\mu_{n+1}} \nabla \log\left(\frac{\mu_{n+1}}{\sigma_{I'}}\right) - S_{\mu_n} \nabla \log\left(\frac{\mu_n}{\sigma_{I'}}\right) \right\|_H^2$$

$S_{\mu_{n+1}} \nabla \log\left(\frac{\mu_{n+1}}{\pi}\right) - S_{\mu_n} \nabla \log\left(\frac{\mu_n}{\pi}\right)$
$= \int k(y, \cdot) \nabla \log(\mu_{n+1}(y)) \mu_{n+1}(y) dy - \int k(\phi_\gamma(x), \cdot) \nabla \log(\mu_n(x)) \mu_n(x) dx$
$+ \int k(\phi_\gamma(x), \cdot) \nabla \log(\pi(x)) \mu_n(x) dx - \int k(y, \cdot) \nabla \log(\pi(y)) \mu_{n+1}(y) dy + \int k(\phi_\gamma(x), \cdot) \nabla \log(\frac{\mu_n(x)}{\pi(x)}) \mu_n(x) dx -$
$\int k(x, \cdot) \nabla \log(\frac{\mu_n(x)}{\pi(x)}) \mu_n(x) dx = I + II + III$
$\|II\|_H^2 = \int \int k(\phi_\gamma(x), \phi_\gamma(x'))(\nabla \log(x) - \nabla \log(\pi(\phi_\gamma(x))))(\nabla \log(x') - \nabla \log(\pi(\phi_\gamma(x')))) \mu(x) \mu(x') dx dx'$
since F is L-smooth, so
$|\nabla \log(x') - \nabla \log(\pi(\phi_\gamma(x')))| \le L \cdot |x - \phi_\gamma(x)| = L \cdot \gamma \cdot |P_{\mu_\gamma} \nabla \log(\frac{\mu_n(x)}{\sigma_n(x)})| \le C \cdot \gamma \cdot \left\| S_{\mu_n} \nabla \log(\frac{\mu_n}{\sigma_n}) \right\|_H$
then by Cauchy-Schwartz inequality and μ_n is weekly compact(will see this in the end), we derive

$$\|II\|_H^2 \le C \cdot \gamma^2 \cdot \left\| S_{\mu_n} \nabla \log(\frac{\mu_n}{\sigma_n}) \right\|_H^2$$

$\|I\|_H^2 = \left\| \int k(y, \cdot) \nabla \log(\mu_{n+1}(y)) \mu_{n+1}(y) dy - \int k(\phi_\gamma(x), \cdot) \nabla \log(\mu_n(x)) \mu_n(x) dx \right\|_H^2$
$= \int \int k(y, y') \nabla \log(\mu_{n+1}(y)) \nabla \log(\mu_{n+1}(y')) \mu_{n+1}(y) \mu_{n+1}(y') dy dy'$
$- \int \int k(y, \phi_\gamma(x)) \nabla \log(\mu_{n+1}(y)) \nabla \log(\mu_n(x')) \mu_{n+1}(y) \mu_n(x) dy dx'$
$- \int \int k(\phi_\gamma(x), y') \nabla \log(\mu_{n+1}(y')) \nabla \log(\mu_n(x)) \mu_{n+1}(y') \mu_n(x) dy' dx$
$+ \int \int k(\phi_\gamma(x), \phi_\gamma(x')) \nabla \log(\mu_n(x)) \nabla \log(\mu_n(x')) \mu_n(x) \mu_n(x') dx dx'$
$= \int \int tr \nabla_{1,2}^2 k(y, y') \mu_{n+1}(y) \mu_{n+1}(y') dy dy'$
$- \int \int tr(\nabla_{1,2}^2 k(y, \phi_\gamma(x')) J \phi_\gamma(x')) \mu_{n+1}(y) \mu_n(x') dy dx'$
$- \int \int tr(\nabla_{1,2}^2 k(\phi_\gamma(x), y') J \phi_\gamma(x)) \mu_n(x) \mu_{n+1}(y') dx dy'$
$+ \int \int tr(\nabla_{1,2}^2 K(\phi_\gamma(x), \phi_\gamma(x')) J \phi_\gamma(x) J \phi_\gamma(x')) \mu_n(x) \mu_n(x') dx dx'$
$= \int \int tr(\nabla_{1,2}^2 K(\phi_\gamma(x), \phi_\gamma(x'))(I - J\phi_\gamma(x))(I - J\phi_\gamma(x'))) \mu_n(x) \mu_n(x') dx dx'$

assume $|\nabla_{1,2}^2 K(\cdot, \cdot)| \le C$, then $tr(\nabla_{1,2}^2 K(\phi_\gamma(x), \phi_\gamma(x'))(I - J\phi_\gamma(x))(I - J\phi_\gamma(x'))) \le C \cdot \gamma^2 \cdot \left\| S_{\mu_n} \nabla \log(\frac{\mu_n}{\sigma_n}) \right\|_H^2$

so $\|I\|_H^2 \le C \cdot \gamma^2 \cdot \left\| S_{\mu_n} \nabla \log(\frac{\mu_n}{\sigma_n}) \right\|_H^2$

$\|III\|_H^2 = \int \int (k(\phi_\gamma(x), \phi_\gamma(x')) - k(\phi_\gamma(x), x') - k(x, \phi_\gamma(x')) + k(x, x')) \nabla \log(\frac{\mu_n(x)}{\pi(x)}) \nabla \log(\frac{\mu_n(x')}{\pi(x')})$
$\mu_n(x) \mu_n(x') dx dx'$.
denote $F(x, x') := k(\phi_\gamma(x), \phi_\gamma(x')) - k(\phi_\gamma(x), x') - k(x, \phi_\gamma(x')) + k(x, x')$
then by integration by part, we have
$\|III\|_H^2 = \int \int (tr(\nabla_{1,2}^2 F(x, x')) + \nabla_2 F(x, x') \nabla \log(\pi(x)) + \nabla_1 F(x, x') \nabla \log(\pi(x')) + F(x, x') \nabla \log(\pi(x)) \nabla \log(\pi(x'))) \mu_n(x) \mu$
We need to know the order of $F(x, x'), \nabla_1 F(x, x'), \nabla_2 F(x, x'), \nabla_{1,2}^2 F(x, x')$

in the following we assume supremum norm the second,third and fourth order derivatives of k are finite, and we miss $O(|\phi_\gamma(x) - x|^2)$ and $O(|\phi_\gamma(x') - x'|^2)$ in the derivation.

$|F(x,x')| = | < \phi_\gamma(x' - x'), \nabla_2 k(\phi_\gamma(x), x') > + < x' - \phi_\gamma(x'), \nabla_2 k(x, x') > |$

$\leq |\nabla_{1,2}^2 k(x, x')(\phi_\gamma(x) - x)(\phi_\gamma(x') - x')| \leq C \cdot \gamma^2 \cdot \left\| S_{\mu_n} \nabla \log(\frac{\mu_n}{\sigma_n}) \right\|_H^2$

$|\nabla_1 F(x,x')| = |\nabla_1 k(\phi_\gamma(x), \phi_\gamma(x')) J\phi_\gamma(x) - \nabla_1 k(\phi_\gamma(x), x') J\phi_\gamma(x) - \nabla_1 k(x, \phi_\gamma(x')) + \nabla_1 k(x, x')|$

$= |\nabla_{1,2}^2 k(\phi_\gamma(x), x')(\phi_\gamma(x' - x')) J\phi_\gamma(x) - \nabla_{1,2}^2 k(x, x')(\phi_\gamma(x') - x')|$

$= |\nabla_{11,2}^3 k(x, x')(\phi_\gamma(x') - x')(\phi_\gamma(x) - x) + \nabla_{1,2}^2 k(\phi_\gamma(x), x)(\phi_\gamma(x') - x')(J\phi_\gamma(x) - I)| \leq C \cdot \gamma^2 \left\| S_{\mu_n} \nabla \log(\frac{\mu_n}{\sigma_n}) \right\|_H^2$

similarly $|\nabla_2 F(x,x')| \leq C \cdot \gamma^2 \left\| S_{\mu_n} \nabla \log(\frac{\mu_n}{\sigma_n}) \right\|_H^2$

$tr\nabla_{1,2}^2 F(x,x') = tr(J\phi_\gamma(x)\nabla_{1,2}^2 k(\phi_\gamma(x), \phi_\gamma(x')) J\phi_\gamma(x') - J\phi_\gamma(x)\nabla_{1,2}^2 k(\phi_\gamma(x), x') - \nabla_{1,2}^2 k(x, \phi_\gamma(x')) J\phi_\gamma(x') + \nabla_{1,2}^2 k(x, x'))$,

denote $J\phi_\gamma(x') = I - \gamma J P_{\mu_n}\nabla \log(\frac{\mu_n}{\sigma_n})(x') = I - \gamma P', J\phi_\gamma(x) = I - \gamma J P_{\mu_n}\nabla \log(\frac{\mu_n}{\sigma_n})(x) = I - \gamma P$

then

$tr\nabla_{1,2}^2 F(x,x') = tr((I - \gamma P)\nabla_{1,22}^3 k(\phi_\gamma(x), x')(\phi_\gamma(x') - x') - (I - \gamma P)\nabla_{1,2}^2 k(\phi_\gamma(x), \phi_\gamma(x'))(\gamma P')$

$- \nabla_{1,22}^3 k(x, x')(\phi_\gamma(x') - x') + (\gamma P')\nabla_{1,2}^2 k(x, \phi_\gamma(x')))$

$= tr(\nabla_{11,22}^4 k(x, x')(\phi_\gamma(x) - x)(\phi_\gamma(x') - x') - (\gamma P')\nabla_{11,2}^3 k(x, \phi_\gamma(x'))(\phi_\gamma(x) - x) - (\gamma P)\nabla_{1,22}^3 k(\phi_\gamma(x), x')(\phi_\gamma(x') - x') + (\gamma P)\nabla_{1,2}^2 k(x, x')(\gamma P'))$

$|tr\nabla_{1,2}^2 F(x,x')| \leq C \cdot \gamma^2 \left\| S_{\mu_n} \nabla \log(\frac{\mu_n}{\sigma_n}) \right\|_H^2$

then by weak compactness, we know,

$\|III\|_H^2 \leq C \cdot \gamma^2 \left\| S_{\mu_n} \nabla \log(\frac{\mu_n}{\sigma_n}) \right\|_H^2$

similarly, since the set of $\{\sigma_{I'}\}\}$ is finite and have similar property with π, so

$\left\| S_{\mu_{n+1}} \nabla \log(\frac{\mu_{n+1}}{\sigma_{I'}}) - S_{\mu_n} \nabla \log(\frac{\mu_n}{\sigma_{I'}}) \right\|_H^2 \leq C \cdot \gamma^2 \left\| S_{\mu_n} \nabla \log(\frac{\mu_n}{\sigma_n}) \right\|_H^2$

Finally,

$\mathbf{E} \left\| S_{\mu_{n+1}} \nabla \log(\frac{\mu_{n+1}}{\pi}) - S_{\mu_{n+1}} \nabla \log(\frac{\mu_{n+1}}{\sigma_{n+1}}) \right\|_H^2 \leq (1-p) \cdot \left\| S_{\mu_n} \nabla \log(\frac{\mu_n}{\pi}) - S_{\mu_n} \nabla \log(\frac{\mu_n}{\sigma_n}) \right\|_H^2 + C \cdot \gamma^2 \left\| S_{\mu_n} \nabla \log(\frac{\mu_n}{\sigma_n}) \right\|_H^2$

So if γ is small enough(depend on constant C), then

$\mathbf{E}(\Phi_{n+1}) \leq \mathbf{E}(\Phi_n) - \frac{\gamma}{2} \mathbf{E} \left\| S_{\mu_n} \nabla \log(\frac{\mu_n}{\pi}) \right\|_H^2$,

where

$\Phi_n := KL(\mu_n|\pi) + \frac{\gamma}{2p} \left\| S_{\mu_n} \nabla \log(\frac{\mu_n}{\pi}) - S_{\mu_n} \nabla \log(\frac{\mu_n}{\sigma_n}) \right\|_H^2$

(you see Φ_n is decreasing, so $KL(\mu_n|\pi) \leq C$, so $\{\mu_n\}$ is weakly compact)

Problem:is $\int |x|^2 \mu_n(x)dx$ weakly continuous?

If we cannot prove this, then we need more assumption on π, that is : there exists a constant $a > 0$ and a point x_0 such that $\int e^{ad(x_0,x)^2} \pi(dx) < +\infty$, then by a theorem from Vilani's book "optimal transport old and new" page 592. we have T_1 inequality $W_p(\mu, \nu) \leq \sqrt{\frac{2H_\nu(\mu)}{\lambda}}$, $p = 1, \lambda$ is some constant. So we have $W_1(\mu_n, \pi) \leq C$, by L-smoothness of F we have $\int |\nabla \log(\pi(x))|\mu_n(x)dx \leq C \cdot \int |x|\mu_n(x)dx \leq C$