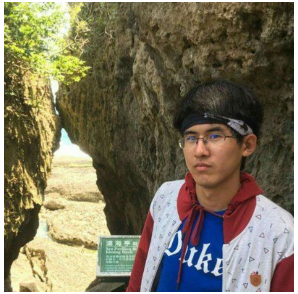
A Spectral Approach to Gradient Estimation for Implicit Distributions

Jiaxin Shi

Tsinghua University

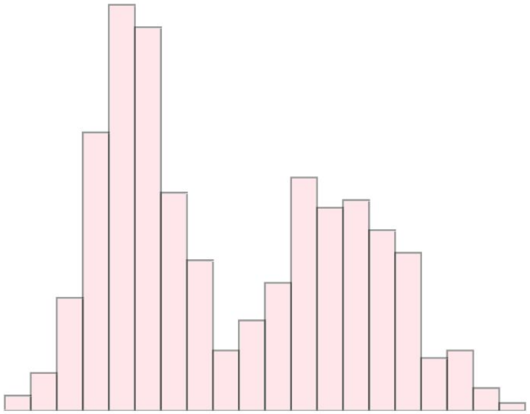


Shengyang Sun

Jun Zhu

Dealing with Intractable densities - A fundamental question:

- Can we directly estimate the gradient function from samples?

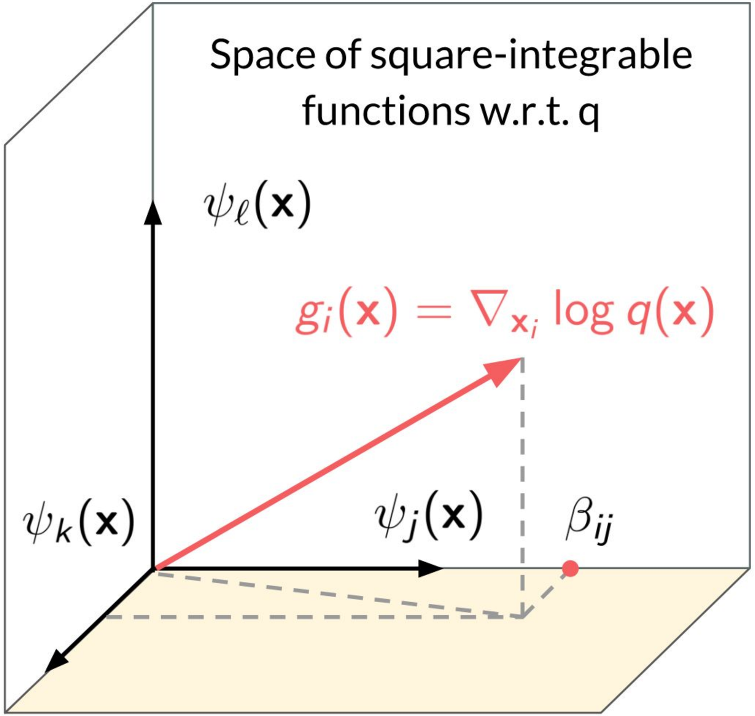


$x^{1:M} \sim q$

$$g(\mathbf{x}) = \nabla_{\mathbf{x}} \log q(\mathbf{x})$$

A Spectral Approach to Gradient Estimation for Implicit Distributions (ICML 2018)

Jiaxin Shi, Shengyang Sun, Jun Zhu



We proved that (under mild assumptions)

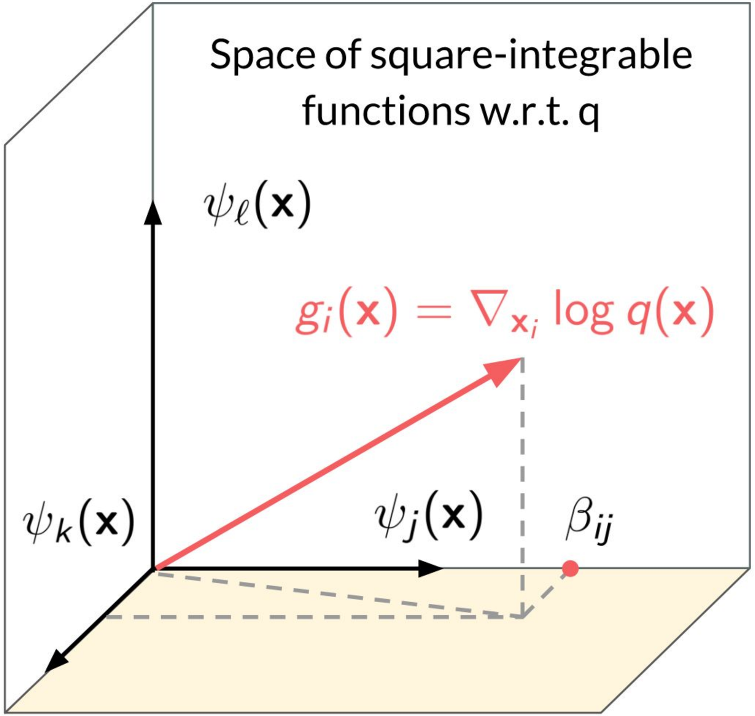
$$\nabla_{x_i} \log q(\mathbf{x}) = -\sum_{j=1}^{\infty} \left[\mathbb{E}_q \nabla_{x_i} \psi_j(\mathbf{x})\right] \psi_j(\mathbf{x})$$

This orthonormal basis can be constructed by spectral decomposition of a kernel

$$\int k(\mathbf{x}, \mathbf{y}) \psi_j(\mathbf{y}) q(\mathbf{y}) d\mathbf{y} = \mu_j \psi_j(\mathbf{x})$$

A Spectral Approach to Gradient Estimation for Implicit Distributions (ICML 2018)

Jiaxin Shi, Shengyang Sun, Jun Zhu



Monte Carlo Nyström approximation

$$\nabla_{x_i} \log q(\mathbf{x}) = -\sum_{j=1}^{J} \left[\mathbb{E}_q \nabla_{x_i} \bar{\psi}_j(\mathbf{x})\right] \bar{\psi}_j(\mathbf{x})$$

Spectral Stein Gradient Estimator

A Spectral Approach to Gradient Estimation for Implicit Distributions (ICML 2018)

Jiaxin Shi, Shengyang Sun, Jun Zhu

Theorem (Error Bound) Given mild assumptions, the error

$$\int |\hat{g}_i(\mathbf{x}) - g_i(\mathbf{x})|^2 q(\mathbf{x})\, d\mathbf{x}$$

Our estimator True gradient

is bounded by

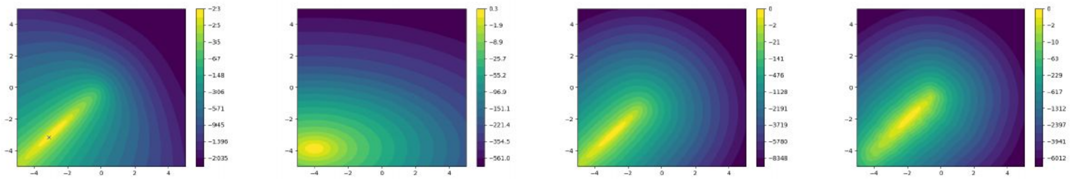
$$J^2 \left(O_p\left(\frac{1}{M}\right) + O_p\left(\frac{C}{\mu_J \Delta_J^2 M}\right)\right) + J O_p\left(\frac{C}{\mu_J \Delta_J^2 M}\right) + \|g_i\|_{\mathcal{H}}^2 O(\mu_J),$$

Estimation error Approximation error

where $\Delta_J = \min_{1 \le j \le J} |\mu_j - \mu_{j+1}|$

A Spectral Approach to Gradient Estimation for Implicit Distributions (ICML 2018)

Jiaxin Shi, Shengyang Sun, Jun Zhu

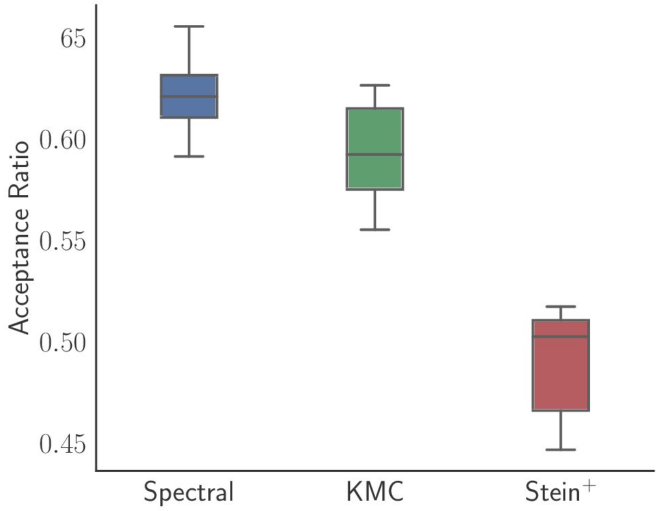


(b) True posterior (c) VI (factorized) (d) HMC (e) KIVI

Variational Inference with Implicit Distributions

(b) Implicit VAE, w/o entropy (c) Implicit VAE, Spectral

Improving Deep Generative Models



Gradient-free Hamiltonian Monte Carlo

A Spectral Approach to Gradient Estimation for Implicit Distributions (ICML 2018)

Jiaxin Shi, Shengyang Sun, Jun Zhu

Thanks

$$\nabla_{x_i} \log q(\mathbf{x}) = -\sum_{j=1}^{\infty} \left[\mathbb{E}_q \nabla_{x_i} \psi_j(\mathbf{x}) \right] \psi_j(\mathbf{x})$$