Bias-Variance Tradeoff in Numerical Solution to the Poisson Equation

5th Workshop on Cognition and Control
Gainesville, Florida, Jan 13-14, 2017

Amirhossein Taghvaei
Joint work with P. G. Mehta and S. P. Meyn

Coordinated Science Laboratory
University of Illinois at Urbana-Champaign

Jan 13, 2017

ILLINOIS

Poisson equation: $$-\frac{1}{\rho(x)}\nabla\cdot(\rho(x)\nabla\phi(x)) = h(x) - \hat{h}$$
$$\int_{\mathbb{R}^d} \phi(x)\rho(x)\,\mathrm{d}x = 0$$

- $\rho : \mathbb{R}^d \to \mathbb{R}^+$ (prob. density)
- $h : \mathbb{R}^d \to \mathbb{R}$ (given function), $\hat{h} := \int h(x)\rho(x)\,\mathrm{d}x$
- $\phi : \mathbb{R}^d \to \mathbb{R}$ (solution)

Problem:

Given: $\{X^1, \ldots, X^N\} \overset{\text{i.i.d}}{\sim} \rho$

Find: $\{\nabla\phi(X^1), \ldots, \nabla\phi(X^N)\}$ (approximately)

Almost like a statistical learning problem

R. S. Laugesen, P. G. Mehta, S. P. Meyn, and M. Raginsky. Poisson Equation in Nonlinear Filtering. SICON, 2015

Poisson equation: $$-\frac{1}{\rho(x)}\nabla\cdot(\rho(x)\nabla\phi(x)) = h(x)-\hat{h}$$
$$\int_{\mathbb{R}^d}\phi(x)\rho(x)\,\mathrm{d}x = 0$$

- $\rho:\mathbb{R}^d\to\mathbb{R}^+$ (prob. density)
- $h:\mathbb{R}^d\to\mathbb{R}$ (given function), $\hat{h} := \int h(x)\rho(x)\,\mathrm{d}x$
- $\phi:\mathbb{R}^d\to\mathbb{R}$ (solution)

Problem:

Given: $\{X^1,\ldots,X^N\}\overset{\text{i.i.d}}{\sim}\rho$

Find: $\{\nabla\phi(X^1),\ldots,\nabla\phi(X^N)\}$ (approximately)

Almost like a statistical learning problem

R. S. Laugesen, P. G. Mehta, S. P. Meyn, and M. Raginsky. Poisson Equation in Nonlinear Filtering. SICON, 2015

Poisson equation: $$-\frac{1}{\rho(x)}\nabla\cdot(\rho(x)\nabla\phi(x)) = h(x)-\hat{h}$$
$$\int_{\mathbb{R}^d}\phi(x)\rho(x)\,\mathrm{d}x = 0$$

- $\rho:\mathbb{R}^d\to\mathbb{R}^+$ (prob. density)
- $h:\mathbb{R}^d\to\mathbb{R}$ (given function), $\hat{h} := \int h(x)\rho(x)\,\mathrm{d}x$
- $\phi:\mathbb{R}^d\to\mathbb{R}$ (solution)

Problem:

Given: $\{X^1,\dots,X^N\}\overset{\text{i.i.d}}{\sim}\rho$

Find: $\{\nabla\phi(X^1),\dots,\nabla\phi(X^N)\}$ (approximately)

Almost like a statistical learning problem

R. S. Laugesen, P. G. Mehta, S. P. Meyn, and M. Raginsky. Poisson Equation in Nonlinear Filtering. SICON, 2015

Feedback Particle Filter

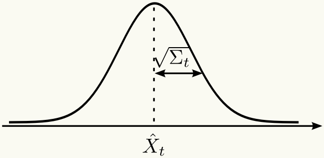
Generalization of the Kalman Filter

Kalman Filter:

$$\mathrm{d}X_t = AX_t\,\mathrm{d}t + \mathrm{d}B_t$$
$$\mathrm{d}Z_t = HX_t\,\mathrm{d}t + \mathrm{d}W_t$$

$\mathsf{P}(X_t|\mathcal{Z}_t) =$ Gaussian $N(\hat{X}_t, \Sigma_t)$,

$$\mathrm{d}\hat{X}_t = A\hat{X}_t\,\mathrm{d}t + \mathsf{K}_t(\,\mathrm{d}Z_t - H\hat{X}_t\,\mathrm{d}t)$$
$$\frac{\mathrm{d}\Sigma_t}{\mathrm{d}t} = \ldots \text{(Riccati equation)}$$



Feedback Particle Filter:

$$\mathrm{d}X_t = a(X_t)\,\mathrm{d}t + \mathrm{d}B_t$$
$$\mathrm{d}Z_t = h(X_t)\,\mathrm{d}t + \mathrm{d}W_t$$

$\mathsf{P}(X_t|\mathcal{Z}_t) \approx$ empirical dist. $\{X^1, \ldots, X^N\}$,

$$\mathrm{d}X_t^i = a(X_t^i)\,\mathrm{d}t + \mathrm{d}B_t^i + \mathsf{K}_t(X_t^i) \circ (\,\mathrm{d}Z_t - \frac{h(X_t^i) + \hat{h}_t}{2}\,\mathrm{d}t)$$

Challenge: Compute the gain function $\mathsf{K}_t := \nabla\phi$ from Poisson eq.

T. Yang, P. G. Mehta, and S. P. Meyn. feedback particle filter, *TAC*, 2013
T. Yang, R. S. Laugesen, P. G. Mehta, and S. P. Meyn. Multivariable feedback particle filter, *Automatica*, 2016

Feedback Particle Filter

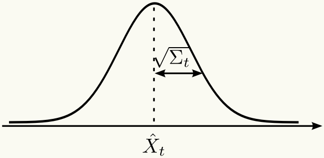
Generalization of the Kalman Filter

Kalman Filter:

$$\mathrm{d}X_t = AX_t\,\mathrm{d}t + \mathrm{d}B_t$$
$$\mathrm{d}Z_t = HX_t\,\mathrm{d}t + \mathrm{d}W_t$$

$\mathsf{P}(X_t|\mathcal{Z}_t) =$ Gaussian $N(\hat{X}_t, \Sigma_t)$,

$$\mathrm{d}\hat{X}_t = A\hat{X}_t\,\mathrm{d}t + \mathsf{K}_t(\mathrm{d}Z_t - H\hat{X}_t\,\mathrm{d}t)$$
$$\frac{\mathrm{d}\Sigma_t}{\mathrm{d}t} = \ldots \text{(Riccati equation)}$$

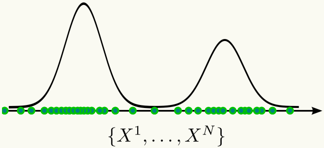


Feedback Particle Filter:

$$\mathrm{d}X_t = a(X_t)\,\mathrm{d}t + \mathrm{d}B_t$$
$$\mathrm{d}Z_t = h(X_t)\,\mathrm{d}t + \mathrm{d}W_t$$

$\mathsf{P}(X_t|\mathcal{Z}_t) \approx$ empirical dist. $\{X^1, \ldots, X^N\}$,

$$\mathrm{d}X_t^i = a(X_t^i)\,\mathrm{d}t + \mathrm{d}B_t^i$$
$$+ \mathsf{K}_t(X_t^i) \circ (\mathrm{d}Z_t - \frac{h(X_t^i) + \hat{h}_t}{2}\,\mathrm{d}t)$$



Challenge: Compute the gain function $\mathsf{K}_t := \nabla\phi$ from Poisson eq.

T. Yang, P. G. Mehta, and S. P. Meyn. feedback particle filter, *TAC*, 2013

T. Yang, R. S. Laugesen, P. G. Mehta, and S. P. Meyn. Multivariable feedback particle filter, *Automatica*, 2016

Feedback Particle Filter

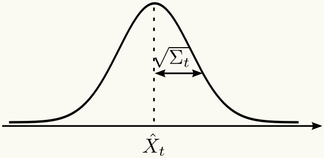
Generalization of the Kalman Filter

Kalman Filter:

$$\mathrm{d}X_t = AX_t\,\mathrm{d}t + \mathrm{d}B_t$$
$$\mathrm{d}Z_t = HX_t\,\mathrm{d}t + \mathrm{d}W_t$$

$\mathsf{P}(X_t|\mathcal{Z}_t) =$ Gaussian $N(\hat{X}_t, \Sigma_t)$,

$$\mathrm{d}\hat{X}_t = A\hat{X}_t\,\mathrm{d}t + \mathsf{K}_t(\,\mathrm{d}Z_t - H\hat{X}_t\,\mathrm{d}t)$$
$$\frac{\mathrm{d}\Sigma_t}{\mathrm{d}t} = \ldots \text{(Riccati equation)}$$

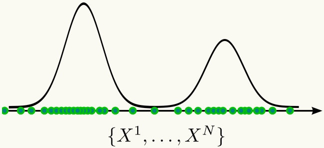


Feedback Particle Filter:

$$\mathrm{d}X_t = a(X_t)\,\mathrm{d}t + \mathrm{d}B_t$$
$$\mathrm{d}Z_t = h(X_t)\,\mathrm{d}t + \mathrm{d}W_t$$

$\mathsf{P}(X_t|\mathcal{Z}_t) \approx$ empirical dist. $\{X^1, \ldots, X^N\}$,

$$\mathrm{d}X_t^i = a(X_t^i)\,\mathrm{d}t + \mathrm{d}B_t^i$$
$$+\,\mathsf{K}_t(X_t^i) \circ (\,\mathrm{d}Z_t - \frac{h(X_t^i) + \hat{h}_t}{2}\,\mathrm{d}t)$$

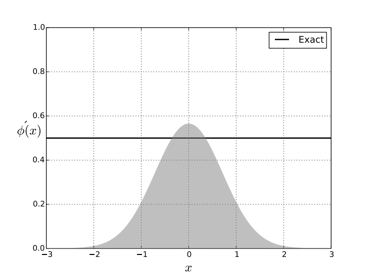


Challenge: Compute the gain function $\mathsf{K}_t := \nabla\phi$ from Poisson eq.

T. Yang, P. G. Mehta, and S. P. Meyn. feedback particle filter, *TAC*, 2013

T. Yang, R. S. Laugesen, P. G. Mehta, and S. P. Meyn. Multivariable feedback particle filter, *Automatica*, 2016

Gaussian distribution linear h

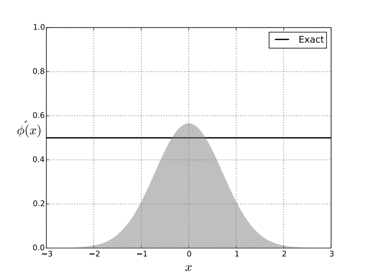


$\nabla\phi(x) = $ constant (Kalman gain)

Bimodal distribution linear h

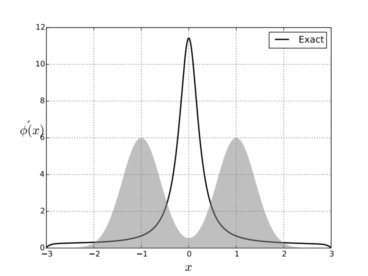
$\nabla\phi(x) = \ldots$ (Nonlinear gain)

Gaussian distribution
linear h



$$\nabla\phi(x) = \text{constant} \quad \text{(Kalman gain)}$$

Bimodal distribution
linear h



$$\nabla\phi(x) = \ldots \text{(Nonlinear gain)}$$

Literature Review
Poisson equation and weighted Laplacian

Poisson equation: $$-\frac{1}{\rho}\nabla\cdot(\rho\nabla\phi) = h - \hat{h}$$

Weighted Laplacian: $$\Delta_\rho\phi := \frac{1}{\rho}\nabla\cdot(\rho\nabla\phi) = \Delta\phi + \nabla\log\rho\cdot\nabla\phi$$

PDE

- Markov Diffusion operators [D. Bakry, et. al. 2013]
- Heat kernels [A. Grigoryan, 2009]

Stochastic analysis

- Simulation and optimization theory for Markov models [S. Meyn, R. Tweedie, 2012]

Statistical learning

- Nonlinear dimensionality reduction [M. Belkin, 2003]
- Diffusion maps [R. Coifman, S. Lafon, 2006]
- Spectral clustering [M. Hein, et. al. 2006]

Three Formulations of the Poisson Equation

P) Weak formulation: (Galerkin)

$$\langle \nabla\phi, \nabla\psi \rangle = \langle h - \hat{h}, \psi \rangle, \quad \forall \psi \in H^1(\mathbb{R}^d, \rho)$$

where $\langle f, g \rangle := \int f(x)g(x)\rho(x)\,\mathrm{d}x$

S) Semigroup formulation: (kernel-based)

$$\phi = P\phi + \tilde{h}$$

where $P := e^{\epsilon \Delta_\rho}$ and $\tilde{h} := \int_0^t e^{s\Delta_\rho}(h - \hat{h})\,\mathrm{d}s$

3) Variational formulation: (Neural net ?)

$$\min_{\phi \in H_0^1(\mathbb{R}^d, \rho)} \mathsf{E}\left[\frac{1}{2}|\nabla\phi(X)|^2 - \phi(X)(h(X) - \hat{h})\right]$$

where $X \sim \rho$

Three Formulations of the Poisson Equation

P) Weak formulation: (Galerkin)

$$\langle \nabla\phi, \nabla\psi \rangle = \langle h - \hat{h}, \psi \rangle, \quad \forall \psi \in H^1(\mathbb{R}^d, \rho)$$

where $\langle f, g \rangle := \int f(x) g(x) \rho(x) \, \mathrm{d}x$

S) Semigroup formulation: (kernel-based)

$$\phi = P\phi + \tilde{h}$$

where $P := e^{\epsilon \Delta_\rho}$ and $\tilde{h} := \int_0^t e^{s\Delta_\rho}(h - \hat{h}) \, \mathrm{d}s$

3) Variational formulation: (Neural net ?)

$$\min_{\phi \in H_0^1(\mathbb{R}^d, \rho)} \mathsf{E}\left[\frac{1}{2}|\nabla\phi(X)|^2 - \phi(X)(h(X) - \hat{h})\right]$$

where $X \sim \rho$

Three Formulations of the Poisson Equation

P) Weak formulation: (Galerkin)

$$\langle \nabla\phi, \nabla\psi \rangle = \langle h - \hat{h}, \psi \rangle, \quad \forall \psi \in H^1(\mathbb{R}^d, \rho)$$

where $\langle f, g \rangle := \int f(x) g(x) \rho(x) \, \mathrm{d}x$

S) Semigroup formulation: (kernel-based)

$$\phi = P\phi + \tilde{h}$$

where $P := e^{\epsilon \Delta_\rho}$ and $\tilde{h} := \int_0^t e^{s\Delta_\rho}(h - \hat{h}) \, \mathrm{d}s$

3) Variational formulation: (Neural net ?)

$$\min_{\phi \in H_0^1(\mathbb{R}^d, \rho)} \mathsf{E}\left[\frac{1}{2}|\nabla\phi(X)|^2 - \phi(X)(h(X) - \hat{h})\right]$$

where $X \sim \rho$

Three Formulations of the Poisson Equation

P) Weak formulation: (Galerkin)

$$\langle \nabla\phi, \nabla\psi \rangle = \langle h - \hat{h}, \psi \rangle, \quad \forall \psi \in H^1(\mathbb{R}^d, \rho)$$

where $\langle f, g \rangle := \int f(x)g(x)\rho(x)\,\mathrm{d}x$

S) Semigroup formulation: (kernel-based)

$$\phi = P\phi + \tilde{h}$$

where $P := e^{\epsilon\Delta_\rho}$ and $\tilde{h} := \int_0^t e^{s\Delta_\rho}(h - \hat{h})\,\mathrm{d}s$

3) Variational formulation: (Neural net ?)

$$\min_{\phi \in H_0^1(\mathbb{R}^d, \rho)} \mathsf{E}\left[\frac{1}{2}|\nabla\phi(X)|^2 - \phi(X)(h(X) - \hat{h})\right]$$

where $X \sim \rho$

P) Weak formulation: (Galerkin)

$$\langle \nabla\phi, \nabla\psi \rangle = \langle h - \hat{h}, \psi \rangle, \quad \forall \psi \in H^1(\mathbb{R}^d, \rho)$$

where $\langle f, g \rangle := \int f(x)g(x)\rho(x)\,\mathrm{d}x$

S) Semigroup formulation: (kernel-based)

$$\phi = P\phi + \tilde{h}$$

where $P := e^{\epsilon\Delta_\rho}$ and $\tilde{h} := \int_0^t e^{s\Delta_\rho}(h - \hat{h})\,\mathrm{d}s$

3) Variational formulation: (Neural net ?)

$$\min_{\phi \in H_0^1(\mathbb{R}^d,\rho)} \mathsf{E}\left[\frac{1}{2}|\nabla\phi(X)|^2 - \phi(X)(h(X) - \hat{h})\right]$$

where $X \sim \rho$

Three Formulations of the Poisson Equation

P) Weak formulation: (Galerkin)

$$\langle \nabla\phi, \nabla\psi \rangle = \langle h - \hat{h}, \psi \rangle, \quad \forall \psi \in H^1(\mathbb{R}^d, \rho)$$

where $\langle f, g \rangle := \int f(x) g(x) \rho(x) \, \mathrm{d}x$

S) Semigroup formulation: (kernel-based)

$$\phi = P\phi + \tilde{h}$$

where $P := e^{\epsilon \Delta_\rho}$ and $\tilde{h} := \int_0^t e^{s\Delta_\rho} (h - \hat{h}) \, \mathrm{d}s$

3) Variational formulation: (Neural net ?)

$$\min_{\phi \in H_0^1(\mathbb{R}^d, \rho)} \mathsf{E}\left[\frac{1}{2} |\nabla \phi(X)|^2 - \phi(X)(h(X) - \hat{h}) \right]$$

where $X \sim \rho$

Three Formulations of the Poisson Equation

P) Weak formulation: (Galerkin)

$$\langle \nabla\phi, \nabla\psi \rangle = \langle h - \hat{h}, \psi \rangle, \quad \forall \psi \in H^1(\mathbb{R}^d, \rho)$$

where $\langle f, g \rangle := \int f(x)g(x)\rho(x)\,\mathrm{d}x$

S) Semigroup formulation: (kernel-based)

$$\phi = P\phi + \tilde{h}$$

where $P := e^{\epsilon\Delta_\rho}$ and $\tilde{h} := \int_0^t e^{s\Delta_\rho}(h - \hat{h})\,\mathrm{d}s$

3) Variational formulation: (Neural net ?)

$$\min_{\phi \in H_0^1(\mathbb{R}^d,\rho)} \mathsf{E}\left[\frac{1}{2}|\nabla\phi(X)|^2 - \phi(X)(h(X) - \hat{h})\right]$$

where $X \sim \rho$

Strong form:

$$-\frac{1}{\rho(x)}\nabla\cdot(\rho(x)\nabla\phi(x)) = h(x) - \hat{h}$$

Weak form:

$$\langle\nabla\phi,\nabla\psi\rangle = \langle h-\hat{h},\psi\rangle, \quad \forall\psi\in H^1(\mathbb{R}^d,\rho)$$

where $\langle f,g\rangle := \int f(x)g(x)\rho(x)\,\mathrm{d}x$

Galerkin approximation:

$$\langle\nabla\phi^{(M)},\nabla\psi\rangle = \langle h-\hat{h},\psi\rangle, \quad \forall\psi\in S$$

where $S = \text{span}\{\psi_1,\ldots,\psi_M\}$

Empirical approximation

$$\frac{1}{N}\sum_{i=1}^{N}\nabla\phi^{(M)}(X^i)\cdot\nabla\psi(X^i) = \frac{1}{N}\sum_{i=1}^{N}(h(X^i)-\hat{h})\psi(X^i), \quad \forall\psi\in S$$

where $X^i \sim \rho$

Strong form:

$$-\frac{1}{\rho(x)}\nabla\cdot(\rho(x)\nabla\phi(x)) = h(x) - \hat{h}$$

Weak form:

$$\langle\nabla\phi, \nabla\psi\rangle = \langle h - \hat{h}, \psi\rangle, \quad \forall\psi \in H^1(\mathbb{R}^d, \rho)$$

where $\langle f, g\rangle := \int f(x)g(x)\rho(x)\,\mathrm{d}x$

Galerkin approximation:

$$\langle\nabla\phi^{(M)}, \nabla\psi\rangle = \langle h - \hat{h}, \psi\rangle, \quad \forall\psi \in S$$

where $S = \text{span}\{\psi_1, \ldots, \psi_M\}$

Empirical approximation

$$\frac{1}{N}\sum_{i=1}^{N}\nabla\phi^{(M)}(X^i)\cdot\nabla\psi(X^i) = \frac{1}{N}\sum_{i=1}^{N}(h(X^i) - \hat{h})\psi(X^i), \quad \forall\psi \in S$$

where $X^i \sim \rho$

Strong form:

$$-\frac{1}{\rho(x)}\nabla\cdot(\rho(x)\nabla\phi(x)) = h(x)-\hat{h}$$

Weak form:

$$\langle\nabla\phi,\nabla\psi\rangle = \langle h-\hat{h},\psi\rangle,\quad \forall\psi\in H^1(\mathbb{R}^d,\rho)$$

where $\langle f,g\rangle := \int f(x)g(x)\rho(x)\,\mathrm{d}x$

Galerkin approximation:

$$\langle\nabla\phi^{(M)},\nabla\psi\rangle = \langle h-\hat{h},\psi\rangle,\quad \forall\psi\in S$$

where $S = \text{span}\{\psi_1,\ldots,\psi_M\}$

Empirical approximation

$$\frac{1}{N}\sum_{i=1}^{N}\nabla\phi^{(M)}(X^i)\cdot\nabla\psi(X^i) = \frac{1}{N}\sum_{i=1}^{N}(h(X^i)-\hat{h})\psi(X^i),\quad \forall\psi\in S$$

where $X^i\sim\rho$

Strong form:

$$-\frac{1}{\rho(x)}\nabla\cdot(\rho(x)\nabla\phi(x)) = h(x) - \hat{h}$$

Weak form:

$$\langle\nabla\phi,\nabla\psi\rangle = \langle h-\hat{h},\psi\rangle,\quad \forall\psi\in H^1(\mathbb{R}^d,\rho)$$

where $\langle f,g\rangle := \int f(x)g(x)\rho(x)\,\mathrm{d}x$

Galerkin approximation:

$$\langle\nabla\phi^{(M)},\nabla\psi\rangle = \langle h-\hat{h},\psi\rangle,\quad \forall\psi\in S$$

where $S = \text{span}\{\psi_1,\ldots,\psi_M\}$

Empirical approximation

$$\frac{1}{N}\sum_{i=1}^{N}\nabla\phi^{(M)}(X^i)\cdot\nabla\psi(X^i) = \frac{1}{N}\sum_{i=1}^{N}(h(X^i)-\hat{h})\psi(X^i),\quad \forall\psi\in S$$

where $X^i\sim\rho$

- Select basis functions $\{\psi_1, \dots, \psi_M\}$
- Express the approximate solution as

$$\phi^{(M,N)}(x) = \sum_{m=1}^{M} c_m \psi_m(x)$$

- Obtain $c = (c_1, \dots, c_M)$ by solving

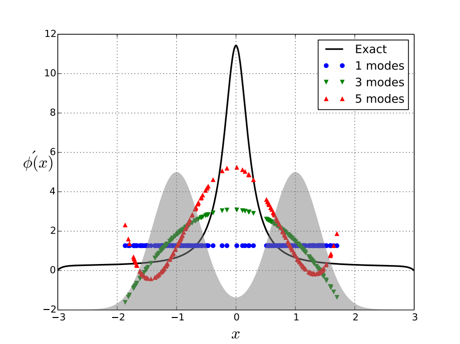
$$Ac = b$$

where

$$A_{ml} = \langle \nabla\psi_m, \nabla\psi_l \rangle \approx \frac{1}{N}\sum_{i=1}^{N} \nabla\psi_m(X^i) \cdot \nabla\psi_l(X^i)$$

$$b_m = \langle \psi_m, h \rangle \approx \frac{1}{N}\sum_{i=1}^{N} \psi_m(X^i) h(X^i) - \hat{h})$$

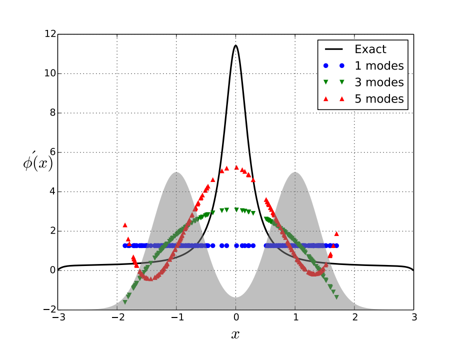
Galerkin Approximation
Numerical result



Issues:

- Choice of basis functions
- Singularity of A
- Computationally scales with $O(Nd^p)$

Galerkin Approximation
Numerical result

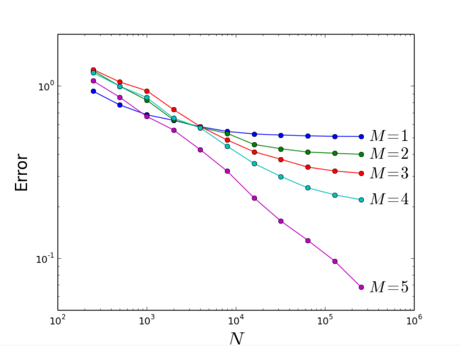


Issues:

- Choice of basis functions
- Singularity of A
- Computationally scales with $O(Nd^p)$

Special case: The basis functions are eigenfunctions of Δ_ρ

$$\underbrace{\mathsf{E}\left[\|\nabla\phi - \nabla\phi^{(M,N)}\|_{L^2}\right]}_{\text{Total error}} \leq \underbrace{\frac{1}{\sqrt{\lambda_M}}\|h - \Pi_S h\|_{L^2}}_{\text{Bias}} + \underbrace{\frac{1}{\sqrt{N}}\|h\|_\infty \sqrt{\sum_{m=1}^{M} \frac{1}{\lambda_m}}}_{\text{Variance}}$$



Poisson equation: $-\Delta_\rho \phi = h - \hat{h}$

Semigroup identity: $e^{\epsilon \Delta_\rho} = I + \int_0^\epsilon e^{s\Delta_\rho} \Delta_\rho \, \mathrm{d}s$

Semigroup formulation:

$$\phi = e^{\epsilon \Delta_\rho} \phi + \tilde{h}$$

where $\tilde{h} := \int_0^\epsilon e^{s\Delta_\rho} (h - \hat{h}) \, \mathrm{d}s$

Kernel representation:

$$\phi(x) = \int \tilde{k}_\epsilon(x,y) \phi(y) \rho(y) \, \mathrm{d}y + \tilde{h}(x)$$

Empirical approximation:

$$\phi(x) = \frac{1}{N} \sum_{i=1}^N \tilde{k}_\epsilon(x, X^i) \phi(X^i) + \tilde{h}(x)$$

But $\tilde{k}_\epsilon(x,y) = ?$

Poisson equation: $-\Delta_\rho \phi = h - \hat{h}$

Semigroup identity: $e^{\epsilon \Delta_\rho} = I + \int_0^\epsilon e^{s\Delta_\rho} \Delta_\rho \, \mathrm{d}s$

Semigroup formulation:

$$\phi = e^{\epsilon \Delta_\rho} \phi + \tilde{h}$$

where $\tilde{h} := \int_0^\epsilon e^{s\Delta_\rho} (h - \hat{h}) \, \mathrm{d}s$

Kernel representation:

$$\phi(x) = \int \tilde{k}_\epsilon(x,y)\phi(y)\rho(y) \, \mathrm{d}y + \tilde{h}(x)$$

Empirical approximation:

$$\phi(x) = \frac{1}{N} \sum_{i=1}^N \tilde{k}_\epsilon(x, X^i)\phi(X^i) + \tilde{h}(x)$$

But $\tilde{k}_\epsilon(x,y) = ?$

Poisson equation: $-\Delta_\rho \phi = h - \hat{h}$

Semigroup identity: $e^{\epsilon \Delta_\rho} = I + \int_0^\epsilon e^{s\Delta_\rho} \Delta_\rho \, \mathrm{d}s$

Semigroup formulation:

$$\phi = e^{\epsilon \Delta_\rho} \phi + \tilde{h}$$

where $\tilde{h} := \int_0^\epsilon e^{s\Delta_\rho} (h - \hat{h}) \, \mathrm{d}s$

Kernel representation:

$$\phi(x) = \int \tilde{k}_\epsilon(x,y) \phi(y) \rho(y) \, \mathrm{d}y + \tilde{h}(x)$$

Empirical approximation:

$$\phi(x) = \frac{1}{N} \sum_{i=1}^{N} \tilde{k}_\epsilon(x, X^i) \phi(X^i) + \tilde{h}(x)$$

But $\tilde{k}_\epsilon(x,y) = ?$

Poisson equation: $-\Delta_\rho \phi = h - \hat{h}$

Semigroup identity: $e^{\epsilon \Delta_\rho} = I + \int_0^\epsilon e^{s\Delta_\rho} \Delta_\rho \, \mathrm{d}s$

Semigroup formulation:

$$\phi = e^{\epsilon \Delta_\rho} \phi + \tilde{h}$$

where $\tilde{h} := \int_0^\epsilon e^{s\Delta_\rho} (h - \hat{h}) \, \mathrm{d}s$

Kernel representation:

$$\phi(x) = \int \tilde{k}_\epsilon(x,y) \phi(y) \rho(y) \, \mathrm{d}y + \tilde{h}(x)$$

Empirical approximation:

$$\phi(x) = \frac{1}{N} \sum_{i=1}^N \tilde{k}_\epsilon(x, X^i) \phi(X^i) + \tilde{h}(x)$$

But $\tilde{k}_\epsilon(x,y) = ?$

Poisson equation: $-\Delta_\rho \phi = h - \hat{h}$

Semigroup identity: $e^{\epsilon \Delta_\rho} = I + \int_0^\epsilon e^{s\Delta_\rho} \Delta_\rho \, \mathrm{d}s$

Semigroup formulation:

$$\phi = e^{\epsilon \Delta_\rho} \phi + \tilde{h}$$

where $\tilde{h} := \int_0^\epsilon e^{s\Delta_\rho} (h - \hat{h}) \, \mathrm{d}s$

Kernel representation:

$$\phi(x) = \int \tilde{k}_\epsilon(x,y)\phi(y)\rho(y) \, \mathrm{d}y + \tilde{h}(x)$$

Empirical approximation:

$$\phi(x) = \frac{1}{N} \sum_{i=1}^{N} \tilde{k}_\epsilon(x, X^i)\phi(X^i) + \tilde{h}(x)$$

But $\tilde{k}_\epsilon(x,y) = ?$

Kernel-based Approximation

Special case: $\rho = 1$

$$e^{\epsilon\Delta} f(x) = \int g_\epsilon(x,y) f(y)\,\mathrm{d}y. \quad \text{(for all } \epsilon > 0\text{)}$$

where g_ϵ is the Gaussian kernel.

In general:

$$e^{\epsilon\Delta_\rho} f(x) \approx \int \frac{1}{n_\epsilon(x)} \frac{g_\epsilon(x,y)}{\sqrt{\int g_\epsilon(y,z)\rho(z)\,\mathrm{d}z}} f(y)\rho(y)\,\mathrm{d}y := T_\epsilon f(x) \quad \text{(for } \epsilon \downarrow 0\text{)}$$

where n_ϵ is normalizing constant.

Empirical apprximation:

$$e^{\epsilon\Delta_\rho} f(x) \approx \sum_{j=1}^{N} \frac{1}{n_\epsilon^{(N)}(x)} \frac{g_\epsilon(x,X^j)}{\sqrt{\frac{1}{N}\sum_{l=1}^{N} g_\epsilon(X^j,X^l)}} f(X^j) := T_\epsilon^{(N)} f(x)$$

where $n_\epsilon^{(N)}$ is normalizing constant.

R. Coifman, S. Lafon, Diffusion maps, *Applied and computational harmonic analysis*, 2006,
M. Hein, J. Audibert, U. Von Luxburg, Convergence of graph Laplacians on random neighborhood graphs, *JLMR*, 2007

Kernel-based Approximation

Special case: $\rho = 1$

$$e^{\epsilon\Delta} f(x) = \int g_\epsilon(x,y) f(y)\,\mathrm{d}y. \quad \text{(for all } \epsilon > 0\text{)}$$

where g_ϵ is the Gaussian kernel.

In general:

$$e^{\epsilon\Delta_\rho} f(x) \approx \int \frac{1}{n_\epsilon(x)} \frac{g_\epsilon(x,y)}{\sqrt{\int g_\epsilon(y,z)\rho(z)\,\mathrm{d}z}} f(y)\rho(y)\,\mathrm{d}y := T_\epsilon f(x) \quad \text{(for } \epsilon \downarrow 0\text{)}$$

where n_ϵ is normalizing constant.

Empirical apprximation:

$$e^{\epsilon\Delta_\rho} f(x) \approx \sum_{j=1}^{N} \frac{1}{n_\epsilon^{(N)}(x)} \frac{g_\epsilon(x,X^j)}{\sqrt{\frac{1}{N}\sum_{l=1}^{N} g_\epsilon(X^j,X^l)}} f(X^j) := T_\epsilon^{(N)} f(x)$$

where $n_\epsilon^{(N)}$ is normalizing constant.

R. Coifman, S. Lafon, Diffusion maps, *Applied and computational harmonic analysis*, 2006,
M. Hein, J. Audibert, U. Von Luxburg, Convergence of graph Laplacians on random neighborhood graphs, *JLMR*, 2007

Kernel-based Approximation

Special case: $\rho = 1$

$$e^{\epsilon\Delta} f(x) = \int g_\epsilon(x,y) f(y)\,\mathrm{d}y. \quad \text{(for all } \epsilon > 0\text{)}$$

where g_ϵ is the Gaussian kernel.

In general:

$$e^{\epsilon\Delta_\rho} f(x) \approx \int \frac{1}{n_\epsilon(x)} \frac{g_\epsilon(x,y)}{\sqrt{\int g_\epsilon(y,z)\rho(z)\,\mathrm{d}z}} f(y)\rho(y)\,\mathrm{d}y := T_\epsilon f(x) \quad \text{(for } \epsilon \downarrow 0\text{)}$$

where n_ϵ is normalizing constant.

Empirical apprximation:

$$e^{\epsilon\Delta_\rho} f(x) \approx \sum_{j=1}^{N} \frac{1}{n_\epsilon^{(N)}(x)} \frac{g_\epsilon(x,X^j)}{\sqrt{\frac{1}{N}\sum_{l=1}^{N} g_\epsilon(X^j,X^l)}} f(X^j) := T_\epsilon^{(N)} f(x)$$

where $n_\epsilon^{(N)}$ is normalizing constant.

R. Coifman, S. Lafon, Diffusion maps, *Applied and computational harmonic analysis*, 2006,
M. Hein, J. Audibert, U. Von Luxburg, Convergence of graph Laplacians on random neighborhood graphs, *JLMR*, 2007

Exact solution:

$$\phi(x) = e^{\epsilon\Delta_\rho}\phi(x) + \tilde{h}(x)$$

where $\tilde{h} := \int_0^\epsilon e^{s\Delta_\rho}(h - \hat{h})\,\mathrm{d}s$

Approximation:

$$\phi_\epsilon^{(N)}(x) := T_\epsilon^{(N)}\phi_\epsilon^{(N)}(x) + \epsilon(h(x) - \hat{h}),$$

Numerics:

$$\Phi = \mathbf{T}\Phi + \epsilon(\mathbf{h} - \hat{h})$$

- $\Phi = (\Phi_\epsilon^{(N)}(X^1), \ldots, \Phi_\epsilon^{(N)}(X^N))$
- $\mathbf{T}_{ij} = \frac{1}{n_\epsilon(X^i)} \frac{g_\epsilon(X^i, X^j)}{\sqrt{\frac{1}{N}\sum_{l=1}^N g_\epsilon(X^i, X^l)}}$ (Markov matrix)

Gradeint:

$$\nabla\phi_\epsilon^{(N)}(x) = \nabla(T_\epsilon^{(N)}\phi_\epsilon^{(N)})(x) + \epsilon\nabla h(x)$$

Exact solution:

$$\phi(x) = e^{\epsilon\Delta_\rho}\phi(x) + \tilde{h}(x)$$

where $\tilde{h} := \int_0^\epsilon e^{s\Delta_\rho}(h - \hat{h})\,\mathrm{d}s$

Approximation:

$$\phi_\epsilon^{(N)}(x) := T_\epsilon^{(N)}\phi_\epsilon^{(N)}(x) + \epsilon(h(x) - \hat{h}),$$

Numerics:

$$\Phi = \mathbf{T}\Phi + \epsilon(\mathbf{h} - \hat{h})$$

- $\Phi = (\Phi_\epsilon^{(N)}(X^1), \dots, \Phi_\epsilon^{(N)}(X^N))$
- $\mathbf{T}_{ij} = \frac{1}{n_\epsilon(X^i)} \frac{g_\epsilon(X^i, X^j)}{\sqrt{\frac{1}{N}\sum_{l=1}^N g_\epsilon(X^i, X^l)}}$ (Markov matrix)

Gradeint:

$$\nabla\phi_\epsilon^{(N)}(x) = \nabla(T_\epsilon^{(N)}\phi_\epsilon^{(N)})(x) + \epsilon\nabla h(x)$$

Exact solution:

$$\phi(x) = e^{\epsilon\Delta_\rho}\phi(x) + \tilde{h}(x)$$

where $\tilde{h} := \int_0^\epsilon e^{s\Delta_\rho}(h - \hat{h})\,\mathrm{d}s$

Approximation:

$$\phi_\epsilon^{(N)}(x) := T_\epsilon^{(N)}\phi_\epsilon^{(N)}(x) + \epsilon(h(x) - \hat{h}),$$

Numerics:

$$\Phi = \mathbf{T}\Phi + \epsilon(\mathbf{h} - \hat{h})$$

- $\Phi = (\Phi_\epsilon^{(N)}(X^1), \ldots, \Phi_\epsilon^{(N)}(X^N))$
- $\mathbf{T}_{ij} = \dfrac{1}{n_\epsilon(X^i)} \dfrac{g_\epsilon(X^i, X^j)}{\sqrt{\frac{1}{N}\sum_{l=1}^N g_\epsilon(X^i, X^l)}}$ (Markov matrix)

Gradeint:

$$\nabla\phi_\epsilon^{(N)}(x) = \nabla(T_\epsilon^{(N)}\phi_\epsilon^{(N)})(x) + \epsilon\nabla h(x)$$

Exact solution:

$$\phi(x) = e^{\epsilon\Delta_\rho}\phi(x) + \tilde{h}(x)$$

where $\tilde{h} := \int_0^\epsilon e^{s\Delta_\rho}(h - \hat{h})\,\mathrm{d}s$

Approximation:

$$\phi_\epsilon^{(N)}(x) := T_\epsilon^{(N)}\phi_\epsilon^{(N)}(x) + \epsilon(h(x) - \hat{h}),$$

Numerics:

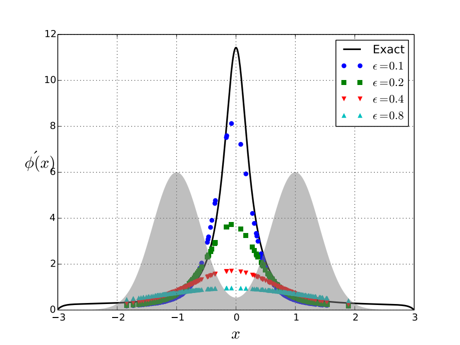
$$\Phi = \mathbf{T}\Phi + \epsilon(\mathbf{h} - \hat{h})$$

- $\Phi = (\Phi_\epsilon^{(N)}(X^1), \ldots, \Phi_\epsilon^{(N)}(X^N))$
- $\mathbf{T}_{ij} = \dfrac{1}{n_\epsilon(X^i)} \dfrac{g_\epsilon(X^i, X^j)}{\sqrt{\frac{1}{N}\sum_{l=1}^N g_\epsilon(X^i, X^l)}}$ (Markov matrix)

Gradeint:

$$\nabla\phi_\epsilon^{(N)}(x) = \nabla(T_\epsilon^{(N)}\phi_\epsilon^{(N)})(x) + \epsilon\nabla h(x)$$

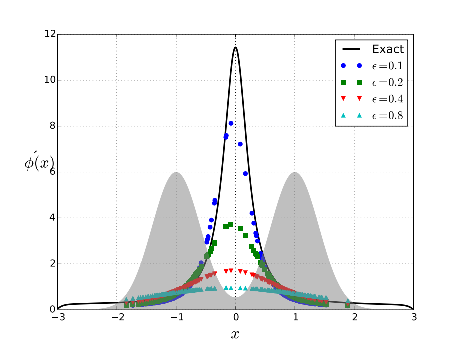
Kernel-based approximation
Numerical result



Properties

1. No singularity
2. Easy extension to Manifolds [C. Zhang, et. al. CDC 2015]
3. Better error bounds
4. Computational cost $O(N^2)$ (good in high dimensions)

Kernel-based approximation
Numerical result

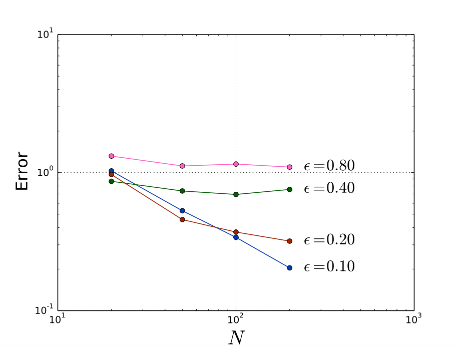


Properties

1. No singularity
2. Easy extension to Manifolds [C. Zhang, et. al. CDC 2015]
3. Better error bounds
4. Computational cost $O(N^2)$ (good in high dimensions)

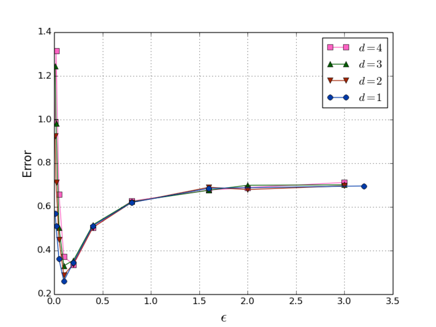
Special case: Bounded domain

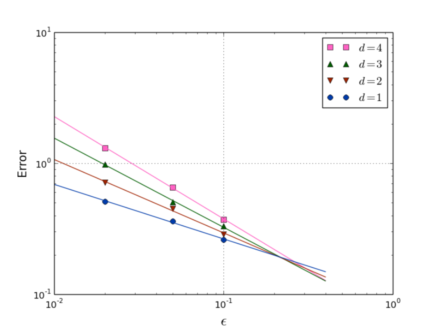
$$\underbrace{\mathsf{E}\left[\|\nabla\phi - \nabla\phi_\epsilon^{(N)}\|_2\right]}_{\text{Total error}} \le \underbrace{O(\epsilon)}_{\text{Bias}} + \underbrace{O(\frac{1}{\epsilon^{1+d/4}\sqrt{N}})}_{\text{Variance}}$$



Special case: Bounded domain

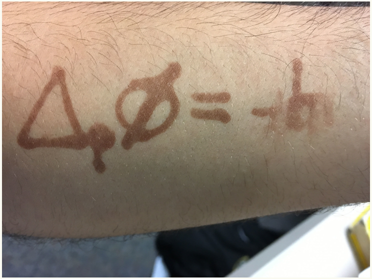
$$\underbrace{\mathsf{E}\left[\|\nabla\phi - \nabla\phi_\epsilon^{(N)}\|_2\right]}_{\text{Total error}} \leq \underbrace{O(\epsilon)}_{\text{Bias}} + \underbrace{O\Big(\frac{1}{\epsilon^{1+d/4}\sqrt{N}}\Big)}_{\text{Variance}}$$





Final slide

Thank you for your attention!



Poisson equation, almost everywhere