Weierstrass Institute for
Applied Analysis and Stochastics

Leibniz Association

SDE based regression for random PDEs

Christian Bayer

Felix Anker, Martin Eigel, Marcel Ladkau, Johannes Neumann, John Schoenmakers

Mohrenstrasse 39 · 10117 Berlin · Germany · Tel. +49 30 20372 0 · www.wias-berlin.de

January 6, 2016

Outline

Idea

$$-\nabla \cdot (\kappa(x)\nabla u(x)) = f(x), \quad x \in D \subset \mathbb{R}^d$$
$$u(x) = g(x), \quad x \in \partial D$$

- Assuming κ, f, g are deterministic, the Feynman-Kac formula gives a collection of random variable $\phi^x = \phi^x(\kappa, f, g)$, $x \in D$, with

$$\forall x \in D: \ u(x) = E\left[\phi^x\right].$$

- If κ, f, g are random, obtain Φ^x, $x \in D$, with

$$u(x) = E\left[\Phi^x | \kappa, f, g\right], \quad x \in D$$

- Hence,

$$E[u(x)] = E\left[\Phi^x\right], \quad x \in D,$$
$$\mathrm{var}\left[u(x)\right] \leq \mathrm{var}\left[\Phi^x\right], \quad x \in D$$

- In general, need spatial resolution of $v(x) \equiv E[u(x)]$, $x \in D$. Several possibilities: interpolation or (local or global) Monte Carlo regression.

Idea

$$-\nabla \cdot (\kappa(x)\nabla u(x)) = f(x), \quad x \in D \subset \mathbb{R}^d$$
$$u(x) = g(x), \quad x \in \partial D$$

- Assuming κ, f, g are deterministic, the Feynman-Kac formula gives a collection of random variable $\phi^x = \phi^x(\kappa, f, g)$, $x \in D$, with

$$\forall x \in D: \; u(x) = E\left[\phi^x\right].$$

- If κ, f, g are random, obtain Φ^x, $x \in D$, with

$$u(x) = E\left[\Phi^x | \kappa, f, g\right], \quad x \in D$$

- Hence,

$$E[u(x)] = E\left[\Phi^x\right], \quad x \in D,$$
$$\text{var}\left[u(x)\right] \leq \text{var}\left[\Phi^x\right], \quad x \in D$$

- In general, need spatial resolution of $v(x) \equiv E[u(x)]$, $x \in D$. Several possibilities: interpolation or (local or global) Monte Carlo regression.

Idea

$$-\nabla \cdot (\kappa(x)\nabla u(x)) = f(x), \quad x \in D \subset \mathbb{R}^d$$
$$u(x) = g(x), \quad x \in \partial D$$

- Assuming κ, f, g are deterministic, the Feynman-Kac formula gives a collection of random variable $\phi^x = \phi^x(\kappa, f, g)$, $x \in D$, with

$$\forall x \in D: \; u(x) = E\left[\phi^x\right].$$

- If κ, f, g are random, obtain Φ^x, $x \in D$, with

$$u(x) = E\left[\Phi^x | \kappa, f, g\right], \quad x \in D$$

- Hence,

$$E[u(x)] = E\left[\Phi^x\right], \quad x \in D,$$
$$\operatorname{var}\left[u(x)\right] \leq \operatorname{var}\left[\Phi^x\right], \quad x \in D$$

- In general, need spatial resolution of $v(x) \equiv E[u(x)]$, $x \in D$. Several possibilities: interpolation or (local or global) Monte Carlo regression.

Idea

$$-\nabla \cdot (\kappa(x)\nabla u(x)) = f(x), \quad x \in D \subset \mathbb{R}^d$$
$$u(x) = g(x), \quad x \in \partial D$$

- Assuming κ, f, g are deterministic, the Feynman-Kac formula gives a collection of random variable $\phi^x = \phi^x(\kappa, f, g)$, $x \in D$, with
$$\forall x \in D: \; u(x) = E\left[\phi^x\right].$$
- If κ, f, g are random, obtain Φ^x, $x \in D$, with
$$u(x) = E\left[\Phi^x | \kappa, f, g\right], \quad x \in D$$
- Hence,
$$E[u(x)] = E\left[\Phi^x\right], \quad x \in D,$$
$$\operatorname{var}\left[u(x)\right] \leq \operatorname{var}\left[\Phi^x\right], \quad x \in D$$
- In general, need spatial resolution of $v(x) \equiv E[u(x)]$, $x \in D$. Several possibilities: interpolation or (local or global) Monte Carlo regression.

Idea

$$-\nabla \cdot (\kappa(x)\nabla u(x)) = f(x), \quad x \in D \subset \mathbb{R}^d$$
$$u(x) = g(x), \quad x \in \partial D$$

- Assuming κ, f, g are deterministic, the Feynman-Kac formula gives a collection of random variable $\phi^x = \phi^x(\kappa, f, g)$, $x \in D$, with

$$\forall x \in D: \; u(x) = E\left[\phi^x\right].$$

- If κ, f, g are random, obtain Φ^x, $x \in D$, with

$$u(x) = E\left[\Phi^x | \kappa, f, g\right], \quad x \in D$$

- Hence,

$$E[u(x)] = E\left[\Phi^x\right], \quad x \in D,$$
$$\mathrm{var}\left[u(x)\right] \leq \mathrm{var}\left[\Phi^x\right], \quad x \in D$$

- In general, need spatial resolution of $v(x) \equiv E[u(x)]$, $x \in D$. Several possibilities: interpolation or (local or global) Monte Carlo regression.

Outline

Feynman-Kac formula for a parabolic Cauchy problem

$$\partial_t u(t,x) - Lu(t,x) = f(x), \quad x \in \mathbb{R}^d,\ t \geq 0,$$
$$u(0,x) = g(x)$$

all coefficients deterministic,

$$Lf(x) = \sum_{i=1}^{n} b_i(x) \frac{\partial}{\partial x_i} f(x) + \frac{1}{2} \sum_{i,j=1}^{n} a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} f(x)$$

Theorem (Feynman-Kac formula)

Let W be a d-dimensional Brownian motion, $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ such that $a = \sigma^\top \sigma$ and let $X = X^x$ solve

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \quad X_0 = x \in \mathbb{R}^d,$$

and $Z_t = \int_0^t f(X_s)ds$. Then

$$u(t,x) = E\left[g(X_t^x) + Z_t^x\right], \quad x \in \mathbb{R}^d,\ t \geq 0.$$

$$\partial_t u(t,x) - Lu(t,x) = f(x), \quad x \in \mathbb{R}^d,\ t \geq 0,$$
$$u(0,x) = g(x)$$

all coefficients deterministic,

$$Lf(x) = \sum_{i=1}^{n} b_i(x) \frac{\partial}{\partial x_i} f(x) + \frac{1}{2} \sum_{i,j=1}^{n} a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} f(x)$$

Theorem (Feynman-Kac formula)

Let W be a d-dimensional Brownian motion, $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ such that $a = \sigma^\top \sigma$ and let $X = X^x$ solve

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \quad X_0 = x \in \mathbb{R}^d,$$

and $Z_t = \int_0^t f(X_s)ds$. Then

$$u(t,x) = E\left[g(X_t^x) + Z_t^x\right], \quad x \in \mathbb{R}^d,\ t \geq 0.$$

$$-Lu(x) = f(x), \quad x \in D$$
$$u(x) = g(x), \quad x \in \partial D$$

Theorem (Feynman-Kac formula)

Let $X = X^x$ *and* $\tau = \tau^x$ *be defined by*

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \quad X_0 = x \in \mathbb{R}^n$$
$$\tau := \inf \left\{ t \geq 0 \,\middle|\, X_t^x \in D^c \right\}.$$

Further, let $Z_t^x = \int_0^t f(X_s^x)ds$, *then*

$$u(x) = E\left[g\left(X_\tau^x\right) + Z_\tau^x\right], \quad x \in D.$$

- Similar representation available for Neumann problem, involving reflected diffusion.

Feynman-Kac formula for an elliptic problem

$$-Lu(x) = f(x), \quad x \in D$$
$$u(x) = g(x), \quad x \in \partial D$$

Theorem (Feynman-Kac formula)

Let $X = X^x$ *and* $\tau = \tau^x$ *be defined by*

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \quad X_0 = x \in \mathbb{R}^n$$
$$\tau := \inf \left\{ t \geq 0 \,\middle|\, X_t^x \in D^c \right\}.$$

Further, let $Z_t^x = \int_0^t f(X_s^x)ds$*, then*

$$u(x) = E\left[g\left(X_\tau^x\right) + Z_\tau^x\right], \quad x \in D.$$

- Similar representation available for Neumann problem, involving reflected diffusion.

Stochastic representation for random coefficients

$$-Lu(x) = f(x), \quad x \in D$$
$$u(x) = g(x), \quad x \in \partial D$$

Theorem (Feynman-Kac formula)

Consider again a standard Brownian motion W independent of L, g, f and define $X = X^x, \tau = \tau^x$ by

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \quad X_0 = x \in D$$
$$\tau := \inf\left\{ t \geq 0 \,\middle|\, X_t^x \in D^c \right\}.$$

Further, let $Z_t^x = \int_0^t f(X_s^x)ds$, then

$$u(x) = E\left[g\left(X_\tau^x\right) + Z_\tau^x \,\middle|\, L, f, g \right], \quad x \in D$$
$$E\left[u(x)\right] = E\left[g\left(X_\tau^x\right) + Z_\tau^x \right], \quad x \in D.$$

- $\mathrm{var}\left[g\left(X_\tau^x\right) + Z_\tau^x \right] = \mathrm{var}[u(x)] + E\left[\mathrm{var}\left[g\left(X_\tau^x\right) + Z_\tau^x \,|L, f, g \right]\right]$

$$-Lu(x) = f(x), \quad x \in D$$
$$u(x) = g(x), \quad x \in \partial D$$

Theorem (Feynman-Kac formula)

Consider again a standard Brownian motion W independent of L, g, f and define $X = X^x, \tau = \tau^x$ by

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \quad X_0 = x \in D$$
$$\tau := \inf\left\{ t \geq 0 \,\middle|\, X_t^x \in D^c \right\}.$$

Further, let $Z_t^x = \int_0^t f(X_s^x)ds$, then

$$u(x) = E\left[g\left(X_\tau^x\right) + Z_\tau^x \,\middle|\, L, f, g\right], \quad x \in D$$
$$E\left[u(x)\right] = E\left[g\left(X_\tau^x\right) + Z_\tau^x\right], \quad x \in D.$$

- $\operatorname{var}\left[g\left(X_\tau^x\right) + Z_\tau^x\right] = \operatorname{var}[u(x)] + E\left[\operatorname{var}\left[g\left(X_\tau^x\right) + Z_\tau^x \,|L, f, g\right]\right]$

$$-Lu(x) = f(x), \quad x \in D$$
$$u(x) = g(x), \quad x \in \partial D$$

Theorem (Feynman-Kac formula)

Consider again a standard Brownian motion W independent of L, g, f and define $X = X^x$, $\tau = \tau^x$ by

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \quad X_0 = x \in D$$
$$\tau := \inf\left\{ t \geq 0 \,\middle|\, X_t^x \in D^c \right\}.$$

Further, let $Z_t^x = \int_0^t f(X_s^x)ds$, then

$$u(x) = E\left[g\left(X_\tau^x\right) + Z_\tau^x \,\middle|\, L, f, g \right], \quad x \in D$$
$$E\left[u(x)\right] = E\left[g\left(X_\tau^x\right) + Z_\tau^x \right], \quad x \in D.$$

- $\mathrm{var}\left[g\left(X_\tau^x\right) + Z_\tau^x \right] = \mathrm{var}[u(x)] + E\left[\mathrm{var}\left[g\left(X_\tau^x\right) + Z_\tau^x \,|L, f, g \right]\right]$

Implementation

$$-\nabla \cdot (\kappa(x)\nabla u(x)) = f(x)$$

corresponds to

$$dX_t = \nabla\kappa(X_t)dt + \sqrt{2\kappa(X_t)}dW_t$$

- Numerical solution by the Euler-Maruyama method:

$$\overline{X}_{t+\Delta t} = \overline{X}_t + \nabla\kappa\left(\overline{X}_t\right)\Delta t + \sqrt{2\kappa(\overline{X}_t)}\Delta W, \quad \Delta W_t \sim \mathcal{N}(0, \Delta t I_n)$$

- Weak error generally $\mathcal{O}(\Delta t)$, however for stopped diffusion X_τ only $\mathcal{O}(\sqrt{\Delta t})$
- Adaptive time-stepping based on distance to the boundary ∂D improves error to $\mathcal{O}(\Delta t)$ again
- Fully parallel computations.

$$-\nabla \cdot (\kappa(x)\nabla u(x)) = f(x)$$

corresponds to

$$dX_t = \nabla\kappa(X_t)dt + \sqrt{2\kappa(X_t)}dW_t$$

- Numerical solution by the Euler-Maruyama method:

$$\overline{X}_{t+\Delta t} = \overline{X}_t + \nabla\kappa\left(\overline{X}_t\right)\Delta t + \sqrt{2\kappa(\overline{X}_t)\Delta W_,} \quad \Delta W_t \sim \mathcal{N}(0, \Delta t I_n)$$

- Weak error generally $\mathcal{O}(\Delta t)$, however for stopped diffusion X_τ only $\mathcal{O}(\sqrt{\Delta t})$
- Adaptive time-stepping based on distance to the boundary ∂D improves error to $\mathcal{O}(\Delta t)$ again
- Fully parallel computations.

Outline

Goal

Compute $x \mapsto v(x) \equiv E[\Phi^x]$, $x \in D$, $\Phi^x = g(X^x_\tau) + Z^x_\tau$.

1. *Deterministic techniques:* Given a grid of $x_i \in D$ and *approximate* values $\overline{v}(x_i)$, $i = 1, \dots, N$, compute $x \mapsto v(x)$ by
   - interpolation
   - regression
2. *Stochastic techniques:* Given random points $x_i \in D$ and corresponding samples $\Phi_i^{x_i}$, $i = 1, \dots, N$, compute $x \mapsto v(x)$ by
   - global regression: minimize $\frac{1}{N} \sum_{i=1}^N \left(\Phi_i^{x_i} - \overline{v}(x) \right)^2$ over a finite-dimensional space $\overline{v} \in V$
   - local regression: $v(x)$ approximated by a weighted average—weighted by distance of x to x_i—of $\Phi_i^{x_i}$

The regression problem

Goal

Compute $x \mapsto v(x) \equiv E\left[\Phi^x\right]$, $x \in D$, $\Phi^x = g\left(X_\tau^x\right) + Z_\tau^x$.

1. *Deterministic techniques:* Given a grid of $x_i \in D$ and *approximate* values $\overline{v}(x_i)$, $i = 1, \ldots, N$, compute $x \mapsto v(x)$ by
   - interpolation
   - regression
2. *Stochastic techniques:* Given random points $x_i \in D$ and corresponding samples $\Phi_i^{x_i}$, $i = 1, \ldots, N$, compute $x \mapsto v(x)$ by
   - global regression: minimize $\frac{1}{N}\sum_{i=1}^{N}\left(\Phi_i^{x_i} - \overline{v}(x)\right)^2$ over a finite-dimensional space $\overline{v} \in V$
   - local regression: $v(x)$ approximated by a weighted average—weighted by distance of x to x_i—of $\Phi_i^{x_i}$

Goal

Compute $x \mapsto v(x) \equiv E\left[\Phi^x\right]$, $x \in D$, $\Phi^x = g\left(X^x_\tau\right) + Z^x_\tau$.

1. *Deterministic techniques:* Given a grid of $x_i \in D$ and *approximate* values $\overline{v}(x_i)$, $i = 1, \ldots, N$, compute $x \mapsto v(x)$ by
   - interpolation
   - regression
2. *Stochastic techniques:* Given random points $x_i \in D$ and corresponding samples $\Phi_i^{x_i}$, $i = 1, \ldots, N$, compute $x \mapsto v(x)$ by
   - global regression: minimize $\frac{1}{N}\sum_{i=1}^{N}\left(\Phi_i^{x_i} - \overline{v}(x)\right)^2$ over a finite-dimensional space $\overline{v} \in V$
   - local regression: $v(x)$ approximated by a weighted average—weighted by distance of x to x_i—of $\Phi_i^{x_i}$

Goal

Compute $x \mapsto v(x) \equiv E\left[\Phi^x\right]$, $x \in D$, $\Phi^x = g\left(X_\tau^x\right) + Z_\tau^x$.

1. *Deterministic techniques:* Given a grid of $x_i \in D$ and *approximate* values $\overline{v}(x_i)$, $i = 1, \ldots, N$, compute $x \mapsto v(x)$ by
   - interpolation
   - regression
2. *Stochastic techniques:* Given random points $x_i \in D$ and corresponding samples $\Phi_i^{x_i}$, $i = 1, \ldots, N$, compute $x \mapsto v(x)$ by
   - global regression: minimize $\frac{1}{N} \sum_{i=1}^N \left(\Phi_i^{x_i} - \overline{v}(x)\right)^2$ over a finite-dimensional space $\overline{v} \in V$
   - local regression: $v(x)$ approximated by a weighted average—weighted by distance of x to x_i—of $\Phi_i^{x_i}$

Goal

Compute $x \mapsto v(x) \equiv E[\Phi^x]$, $x \in D$, $\Phi^x = g(X^x_\tau) + Z^x_\tau$.

1. *Deterministic techniques:* Given a grid of $x_i \in D$ and *approximate* values $\overline{v}(x_i)$, $i = 1, \ldots, N$, compute $x \mapsto v(x)$ by
   - interpolation
   - regression
2. *Stochastic techniques:* Given random points $x_i \in D$ and corresponding samples ${\Phi_i}^{x_i}$, $i = 1, \ldots, N$, compute $x \mapsto v(x)$ by
   - global regression: minimize $\frac{1}{N} \sum_{i=1}^N \left(\Phi_i^{x_i} - \overline{v}(x)\right)^2$ over a finite-dimensional space $\overline{v} \in V$
   - local regression: $v(x)$ approximated by a weighted average—weighted by distance of x to x_i—of $\Phi_i^{x_i}$

Remark

Stochastic techniques can be used with approximate values $\overline{v}(x_i)$, too.

Goal

Compute $x \mapsto v(x) \equiv E\left[\Phi^x\right]$, $x \in D$.

- Basis functions $\psi_1, \ldots, \psi_K : D \to \mathbb{R}$ (orthonormal w.r.t. μ)
- A probability measure μ on D
- Generate ind. samples $x_1, \ldots, x_N$ from μ, and $\Phi_1^{x_1}, \ldots, \Phi_N^{x_N}$
- Here, x_i are *independent* of κ, f, g and W.

$$\widehat{\gamma} := \arg\min_{\gamma \in \mathbb{R}^K} \frac{1}{N} \sum_{i=1}^{N} \left(\Phi_i^{x_i} - \sum_{k=1}^{K} \gamma_k \psi_k(x_i) \right)^2, \quad \widehat{v}(x) := \sum_{k=1}^{K} \widehat{\gamma}_k \psi_k(x)$$

Remark

$$\widehat{v}(x) \xrightarrow{N\to\infty} \sum_{k=1}^{K} \langle v, \psi_k \rangle_{L^2(D,\mu)} \psi_k(x) \text{ in } L^2(\Omega \times D, P \otimes \mu).$$

Goal

Compute $x \mapsto v(x) \equiv E\left[\Phi^x\right]$, $x \in D$.

- Basis functions $\psi_1, \ldots, \psi_K : D \to \mathbb{R}$ (orthonormal w.r.t. μ)
- A probability measure μ on D
- Generate ind. samples $x_1, \ldots, x_N$ from μ, and $\Phi_1^{x_1}, \ldots, \Phi_N^{x_N}$
- Here, x_i are *independent* of κ, f, g and W.

$$\widehat{\gamma} := \arg\min_{\gamma \in \mathbb{R}^K} \frac{1}{N} \sum_{i=1}^{N} \left(\Phi_i^{x_i} - \sum_{k=1}^{K} \gamma_k \psi_k(x_i) \right)^2, \quad \widehat{v}(x) := \sum_{k=1}^{K} \widehat{\gamma}_k \psi_k(x)$$

Remark

$$\widehat{v}(x) \xrightarrow{N \to \infty} \sum_{k=1}^{K} \langle v, \psi_k \rangle_{L^2(D,\mu)} \psi_k(x) \text{ in } L^2\left(\Omega \times D, P \otimes \mu\right).$$

Goal

Compute $x \mapsto v(x) \equiv E[\Phi^x]$, $x \in D$.

- Basis functions $\psi_1, \ldots, \psi_K : D \to \mathbb{R}$ (orthonormal w.r.t. μ)
- A probability measure μ on D
- Generate ind. samples $x_1, \ldots, x_N$ from μ, and $\Phi_1^{x_1}, \ldots, \Phi_N^{x_N}$
- Here, x_i are *independent* of κ, f, g and W.

$$\widehat{\gamma} := \underset{\gamma \in \mathbb{R}^K}{\arg\min} \frac{1}{N} \sum_{i=1}^{N} \left(\Phi_i^{x_i} - \sum_{k=1}^{K} \gamma_k \psi_k(x_i) \right)^2, \quad \widehat{v}(x) := \sum_{k=1}^{K} \widehat{\gamma}_k \psi_k(x)$$

Remark

$$\widehat{v}(x) \xrightarrow{N \to \infty} \sum_{k=1}^{K} \langle v, \psi_k \rangle_{L^2(D,\mu)} \psi_k(x) \text{ in } L^2(\Omega \times D, P \otimes \mu).$$

Goal

Compute $x \mapsto v(x) \equiv E\left[\Phi^x\right]$, $x \in D$.

- Basis functions $\psi_1, \ldots, \psi_K : D \to \mathbb{R}$ (orthonormal w.r.t. μ)
- A probability measure μ on D
- Generate ind. samples $x_1, \ldots, x_N$ from μ, and $\Phi_1^{x_1}, \ldots, \Phi_N^{x_N}$
- Here, x_i are *independent* of κ, f, g and W.

$$\widehat{\gamma} := \underset{\gamma \in \mathbb{R}^K}{\arg\min} \frac{1}{N} \sum_{i=1}^{N} \left(\Phi_i^{x_i} - \sum_{k=1}^{K} \gamma_k \psi_k(x_i) \right)^2, \quad \widehat{v}(x) := \sum_{k=1}^{K} \widehat{\gamma}_k \psi_k(x)$$

Remark

$$\widehat{v}(x) \xrightarrow{N \to \infty} \sum_{k=1}^{K} \langle v, \psi_k \rangle_{L^2(D,\mu)} \psi_k(x) \text{ in } L^2\left(\Omega \times D, P \otimes \mu\right).$$

Semi-stochastic regression

- Let $\mathcal{Y} := \left(\Phi_1^{x_1}, \dots, \Phi_N^{x_N}\right) \in \mathbb{R}^N$, $\mathcal{M} := (\psi_k(x_i))_{i=1,\dots,N,k=1,\dots,K} \in \mathbb{R}^{N\times K}$,
$$\widehat{\gamma} = \left(\mathcal{M}^\top \mathcal{M}\right)^{-1} \mathcal{M}^\top \mathcal{Y}$$
- Inversion of the matrix $\mathcal{M}^\top\mathcal{M}$—rather solving the linear system—may be ill-conditioned. But
$$\frac{1}{N}\left(\mathcal{M}^\top\mathcal{M}\right)_{k,l} = \frac{1}{N}\sum_{i=1}^N \psi_k(x_i)\psi_l(x_i)$$
$$\xrightarrow{N\to\infty} \int_D \psi_k(x)\psi_l(x)\mu(dx) =: (\mathcal{G})_{k,l}$$
- $\mathcal{G} \in \mathbb{R}^{K\times K}$ computed efficiently. Orthonormal case: $\mathcal{G} = I_K$.

Definition (Semi-stochastic regression coefficients)

$$\overline{\gamma} := \frac{1}{N}\mathcal{G}^{-1}\mathcal{M}^\top\mathcal{Y}$$

- $\overline{\gamma}$ is *no* solution of the regression problem!

- Let $\mathcal{Y} := \left(\Phi_1^{x_1}, \dots, \Phi_N^{x_N}\right) \in \mathbb{R}^N$, $\mathcal{M} := (\psi_k(x_i))_{i=1,\dots,N,k=1,\dots,K} \in \mathbb{R}^{N\times K}$,
$$\widehat{\gamma} = \left(\mathcal{M}^\top \mathcal{M}\right)^{-1} \mathcal{M}^\top \mathcal{Y}$$
- Inversion of the matrix $\mathcal{M}^\top \mathcal{M}$—rather solving the linear system—may be ill-conditioned. But
$$\frac{1}{N}\left(\mathcal{M}^\top \mathcal{M}\right)_{k,l} = \frac{1}{N}\sum_{i=1}^{N} \psi_k(x_i)\psi_l(x_i)$$
$$\xrightarrow{N\to\infty} \int_D \psi_k(x)\psi_l(x)\mu(dx) =: (\mathcal{G})_{k.l}$$
- $\mathcal{G} \in \mathbb{R}^{K\times K}$ computed efficiently. Orthonormal case: $\mathcal{G} = I_K$.

Definition (Semi-stochastic regression coefficients)

$$\overline{\gamma} := \frac{1}{N}\mathcal{G}^{-1}\mathcal{M}^\top \mathcal{Y}$$

- $\overline{\gamma}$ is *no* solution of the regression problem!

Semi-stochastic regression

- Let $\mathcal{Y} := \left(\Phi_1^{x_1}, \ldots, \Phi_N^{x_N}\right) \in \mathbb{R}^N$, $\mathcal{M} := (\psi_k(x_i))_{i=1,\ldots,N,k=1,\ldots,K} \in \mathbb{R}^{N\times K}$,
$$\widehat{\gamma} = \left(\mathcal{M}^\top \mathcal{M}\right)^{-1} \mathcal{M}^\top \mathcal{Y}$$
- Inversion of the matrix $\mathcal{M}^\top\mathcal{M}$—rather solving the linear system—may be ill-conditioned. But
$$\frac{1}{N}\left(\mathcal{M}^\top\mathcal{M}\right)_{k,l} = \frac{1}{N}\sum_{i=1}^{N} \psi_k(x_i)\psi_l(x_i)$$
$$\xrightarrow{N\to\infty} \int_D \psi_k(x)\psi_l(x)\mu(dx) =: (\mathcal{G})_{k.l}$$
- $\mathcal{G} \in \mathbb{R}^{K\times K}$ computed efficiently. Orthonormal case: $\mathcal{G} = I_K$.

Definition (Semi-stochastic regression coefficients)

$$\overline{\gamma} := \frac{1}{N}\mathcal{G}^{-1}\mathcal{M}^\top\mathcal{Y}$$

- $\overline{\gamma}$ is *no* solution of the regression problem!

- Let $\mathcal{Y} := \left(\Phi_1^{x_1}, \ldots, \Phi_N^{x_N}\right) \in \mathbb{R}^N$, $\mathcal{M} := (\psi_k(x_i))_{i=1,\ldots,N,k=1,\ldots,K} \in \mathbb{R}^{N\times K}$,
$$\widehat{\gamma} = \left(\mathcal{M}^\top \mathcal{M}\right)^{-1} \mathcal{M}^\top \mathcal{Y}$$
- Inversion of the matrix $\mathcal{M}^\top\mathcal{M}$—rather solving the linear system—may be ill-conditioned. But
$$\frac{1}{N}\left(\mathcal{M}^\top \mathcal{M}\right)_{k,l} = \frac{1}{N}\sum_{i=1}^{N} \psi_k(x_i)\psi_l(x_i)$$
$$\xrightarrow{N\to\infty} \int_D \psi_k(x)\psi_l(x)\mu(dx) =: (\mathcal{G})_{k.l}$$
- $\mathcal{G} \in \mathbb{R}^{K\times K}$ computed efficiently. Orthonormal case: $\mathcal{G} = I_K$.

Definition (Semi-stochastic regression coefficients)

$$\overline{\gamma} := \frac{1}{N}\mathcal{G}^{-1}\mathcal{M}^\top \mathcal{Y}$$

- $\overline{\gamma}$ is *no* solution of the regression problem!

Semi-stochastic regression

- Let $\mathcal{Y} := \left(\Phi_1^{x_1}, \dots, \Phi_N^{x_N}\right) \in \mathbb{R}^N$, $\mathcal{M} := (\psi_k(x_i))_{i=1,\dots,N,k=1,\dots,K} \in \mathbb{R}^{N\times K}$,
$$\widehat{\gamma} = \left(\mathcal{M}^\top \mathcal{M}\right)^{-1} \mathcal{M}^\top \mathcal{Y}$$
- Inversion of the matrix $\mathcal{M}^\top\mathcal{M}$—rather solving the linear system—may be ill-conditioned. But
$$\frac{1}{N}\left(\mathcal{M}^\top\mathcal{M}\right)_{k,l} = \frac{1}{N}\sum_{i=1}^N \psi_k(x_i)\psi_l(x_i)$$
$$\xrightarrow{N\to\infty} \int_D \psi_k(x)\psi_l(x)\mu(dx) =: (\mathcal{G})_{k.l}$$
- $\mathcal{G} \in \mathbb{R}^{K\times K}$ computed efficiently. Orthonormal case: $\mathcal{G} = I_K$.

Definition (Semi-stochastic regression coefficients)

$$\overline{\gamma} := \frac{1}{N}\mathcal{G}^{-1}\mathcal{M}^\top\mathcal{Y}$$

- $\overline{\gamma}$ is *no* solution of the regression problem!

We are interested in the convergence of $\overline{v}$ to v_K with

$$\overline{v}(x) := \sum_{k=1}^{K} \overline{\gamma}_k \psi_k(x), \quad v_K(x) := \sum_{k=1}^{K} \langle v, \psi_k \rangle_{L^2(D,\mu)} \psi_k(x)$$

Theorem

Let $\lambda_{\min} > 0$ *(e.g.* $\lambda_{\min} = 1$*) be the smallest eigenvalue of* $\mathcal{G}$ *and let* $\mathcal{V} > 0$ *s.t.*

- $\int_D \psi_k(x)^2 v(x)^2 \mu(dx) \leq \mathcal{V}$,
- $\int_D \psi_k(x)^2 \operatorname{var}[\Phi^x] \mu(dx) \leq \mathcal{V}$,

$k = 1, \ldots, K$. *Then*

$$\int_D E\left[|\overline{v}(x) - v_K(x)|^2 \right] \mu(dx) \leq \frac{4\mathcal{V}}{\lambda_{\min}} \frac{K}{N}.$$

- **Projection error to the set of basis functions:** $\|v - v_K\| \sim e(K)$
- **Regression error:**

$$\|v_K - \overline{v}\| \leq \varepsilon \text{ at cost } \sim N \times K \text{ with } N \sim K\varepsilon^{-2}$$

- **Time discretization error of the SDE:** Given a (possibly random) time grid t_i, approximate X, Z, τ by $\widetilde{X}, \widetilde{Z}, \widetilde{\tau}$. Using *adaptive* algorithms:

$$\left\|\overline{v} - \widetilde{\overline{v}}\right\| \leq \varepsilon \text{ at cost } \sim \varepsilon^{-1},$$

$\widetilde{\overline{v}}$: result of regression based on $\widetilde{\Phi^x} := g\left(\widetilde{X}^x_{\widetilde{\tau}}\right) + \widetilde{Z}^x_{\widetilde{\tau}}$

Total cost for error tolerance ε

$$C_1\left(e^{-1}(\varepsilon)\right)^2 \varepsilon^{-2} + C_2 e^{-1}(\varepsilon)\varepsilon^{-3}, \quad C_2 \gg C_1$$

This is independent of d—unless via e.

- **Projection error to the set of basis functions:** $\|v - v_K\| \sim e(K)$
- **Regression error:**

$$\|v_K - \overline{v}\| \leq \varepsilon \text{ at cost } \sim N \times K \text{ with } N \sim K\varepsilon^{-2}$$

- **Time discretization error of the SDE:** Given a (possibly random) time grid t_i, approximate X, Z, τ by $\widetilde{X}, \widetilde{Z}, \widetilde{\tau}$. Using *adaptive* algorithms:

$$\left\|\overline{v} - \widetilde{\overline{v}}\right\| \leq \varepsilon \text{ at cost } \sim \varepsilon^{-1},$$

$\widetilde{\overline{v}}$: result of regression based on $\widetilde{\Phi^x} := g\left(\widetilde{X}^x_{\widetilde{\tau}}\right) + \widetilde{Z}^x_{\widetilde{\tau}}$

Total cost for error tolerance ε

$$C_1 \left(e^{-1}(\varepsilon)\right)^2 \varepsilon^{-2} + C_2 e^{-1}(\varepsilon)\varepsilon^{-3}, \quad C_2 \gg C_1$$

This is independent of d—unless via e.

- **Projection error to the set of basis functions:** $\|v - v_K\| \sim e(K)$
- **Regression error:**

$$\|v_K - \overline{v}\| \leq \varepsilon \text{ at cost } \sim N \times K \text{ with } N \sim K\varepsilon^{-2}$$

- **Time discretization error of the SDE:** Given a (possibly random) time grid t_i, approximate X, Z, τ by $\widetilde{X}, \widetilde{Z}, \widetilde{\tau}$. Using *adaptive* algorithms:

$$\left\|\overline{v} - \widetilde{\overline{v}}\right\| \leq \varepsilon \text{ at cost } \sim \varepsilon^{-1},$$

$\widetilde{\overline{v}}$: result of regression based on $\widetilde{\Phi^x} := g\left(\widetilde{X}^x_{\widetilde{\tau}}\right) + \widetilde{Z}^x_{\widetilde{\tau}}$

Total cost for error tolerance ε

$$C_1\left(e^{-1}(\varepsilon)\right)^2 \varepsilon^{-2} + C_2 e^{-1}(\varepsilon)\varepsilon^{-3}, \quad C_2 \gg C_1$$

This is independent of d—unless via e.

- **Projection error to the set of basis functions:** $\|v - v_K\| \sim e(K)$
- **Regression error:**

$$\|v_K - \overline{v}\| \le \varepsilon \text{ at cost } \sim N \times K \text{ with } N \sim K\varepsilon^{-2}$$

- **Time discretization error of the SDE:** Given a (possibly random) time grid t_i, approximate X, Z, τ by $\widetilde{X}, \widetilde{Z}, \widetilde{\tau}$. Using *adaptive* algorithms:

$$\left\|\overline{v} - \widetilde{\overline{v}}\right\| \le \varepsilon \text{ at cost } \sim \varepsilon^{-1},$$

$\widetilde{\overline{v}}$: result of regression based on $\widetilde{\Phi^x} := g\left(\widetilde{X}^x_{\widetilde{\tau}}\right) + \widetilde{Z}^x_{\widetilde{\tau}}$

Total cost for error tolerance ε

$$C_1\left(e^{-1}(\varepsilon)\right)^2 \varepsilon^{-2} + C_2 e^{-1}(\varepsilon)\varepsilon^{-3}, \quad C_2 \gg C_1$$

This is independent of d—unless via e.

$$-\nabla \cdot (\kappa(x)\nabla u(x)) = 1, \quad x \in D := [0,1]^2$$
$$u(x) = \sin(\pi x_1) + \sin(\pi x_2), \quad x \in \partial D$$

Noise: finite-dimensional, based on uniform random variables

$$\kappa(x) = \kappa_m(x) = A \sum_{m=0}^{M} U_m m^{-\sigma} \cos(2\pi\beta_1(m)x_1)\cos(2\pi\beta_2(m)x_2) + \varepsilon,$$

$$U_m \sim \mathcal{U}([0,1]),$$
$$\beta_1(m) = m - k(m)(k(m)+1)/2,$$
$$\beta_2(m) = k(m) - \beta_1(m),$$
$$k(m) = \left\lfloor -1/2 + \sqrt{1/4 + 2m} \right\rfloor$$

Basis functions: global polynomials of degree 4 on D, $\mu = dx|_D$.

Two regimes for κ

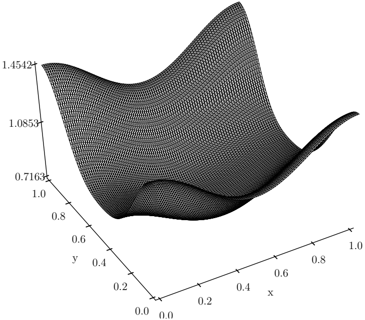


Figure: Sample from smooth κ

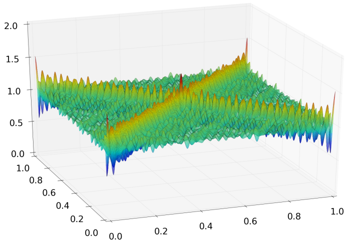
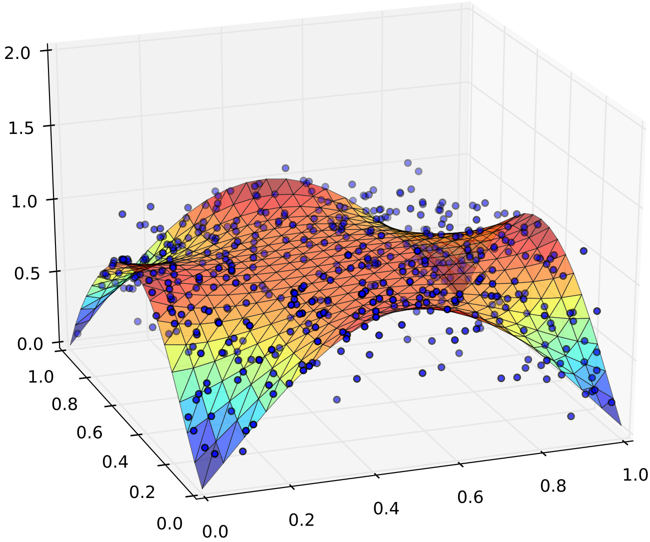
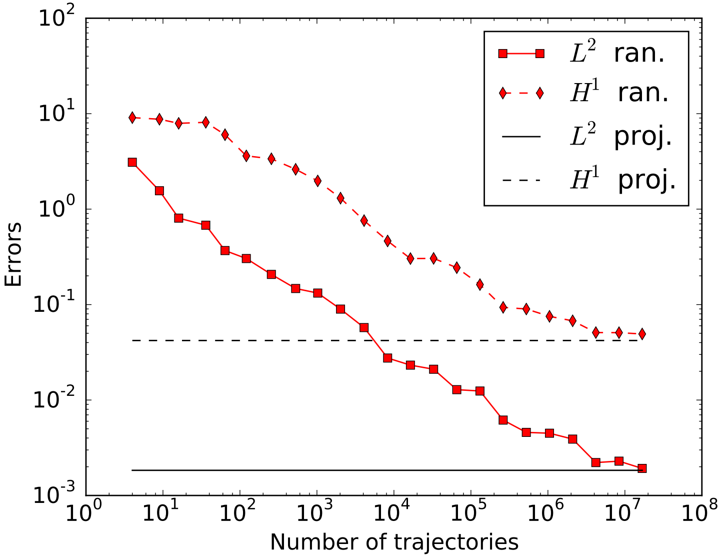


Figure: Sample from rough κ

The regression procedure





Errors
Number of trajectories
L^2 ran.
H^1 ran.
L^2 proj.
H^1 proj.
10^2
10^1
10^0
10^{-1}
10^{-2}
10^{-3}
10^0
10^1
10^2
10^3
10^4
10^5
10^6
10^7
10^8

Outline

- Time dependent problems: compute $E[u(T,x)]$ for

$$\partial_t u(t,x) - \nabla\cdot(\kappa(t,x)\nabla u(t,x)) + \gamma(x)u(t,x)\xi_t = f(x), \quad x \in D$$
$$u(t,x) = g(t,x), \quad x \in \partial D$$
$$u(0,x) = h(x), \quad x \in \overline{D}$$

- Either Dirichlet or Neumann or mixed problems
- Nonlinear random PDEs: stochastic representations by forward-backward SDEs
- Non-local problems: $-Lu(x) = f(x)$

$$L = b_i \frac{\partial}{\partial x_i} + \frac{1}{2} a_{ij} \frac{\partial^2}{\partial x_i \partial x_j} \iff dX_t = b(X_t)dt + \sigma(X_t)dW_t$$
$$L = -(-\Delta)^{\alpha/2} \iff X_t \text{ is } \alpha\text{-stable process } (0 < \alpha < 2)$$

L "fractional Laplacian with random coefficients" $\iff$ X solves SDE driven by stable process.

- Time dependent problems: compute $E[u(T,x)]$ for

$$\partial_t u(t,x) - \nabla \cdot (\kappa(t,x)\nabla u(t,x)) + \gamma(x)u(t,x)\xi_t = f(x), \quad x \in D$$
$$u(t,x) = g(t,x), \quad x \in \partial D$$
$$u(0,x) = h(x), \quad x \in \overline{D}$$

- Either Dirichlet or Neumann or mixed problems
- Nonlinear random PDEs: stochastic representations by forward-backward SDEs
- Non-local problems: $-Lu(x) = f(x)$

$$L = b_i \frac{\partial}{\partial x_i} + \frac{1}{2} a_{ij} \frac{\partial^2}{\partial x_i \partial x_j} \iff dX_t = b(X_t)dt + \sigma(X_t)dW_t$$

$$L = -(-\Delta)^{\alpha/2} \iff X_t \text{ is } \alpha\text{-stable process } (0 < \alpha < 2)$$

L "fractional Laplacian with random coefficients" $\iff$ X solves SDE driven by stable process.

- Time dependent problems: compute $E[u(T,x)]$ for

$$\partial_t u(t,x) - \nabla \cdot (\kappa(t,x)\nabla u(t,x)) + \gamma(x)u(t,x)\xi_t = f(x), \quad x \in D$$
$$u(t,x) = g(t,x), \quad x \in \partial D$$
$$u(0,x) = h(x), \quad x \in \overline{D}$$

- Either Dirichlet or Neumann or mixed problems
- Nonlinear random PDEs: stochastic representations by forward-backward SDEs
- Non-local problems: $-Lu(x) = f(x)$

$$L = b_i \frac{\partial}{\partial x_i} + \frac{1}{2} a_{ij} \frac{\partial^2}{\partial x_i \partial x_j} \iff dX_t = b(X_t)dt + \sigma(X_t)dW_t$$
$$L = -(-\Delta)^{\alpha/2} \iff X_t \text{ is } \alpha\text{-stable process } (0 < \alpha < 2)$$

L “fractional Laplacian with random coefficients” $\iff$ X solves SDE driven by stable process.

- Multilevel Monte Carlo: special care is needed for determining the hitting time at the boundary.
- Adaptivity possible for time-stepping and sampling

Figure: $E[u(x)]$

Figure: Density of points

- Multilevel Monte Carlo: special care is needed for determining the hitting time at the boundary.
- Adaptivity possible for time-stepping and sampling

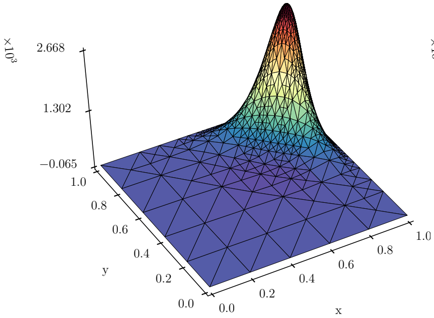


Figure: $E[u(x)]$

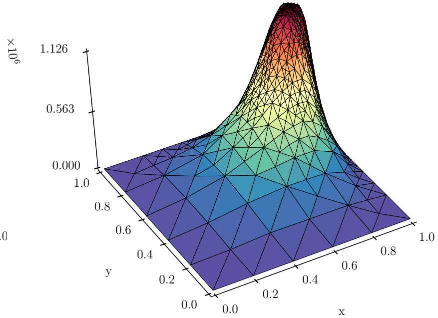


Figure: Density of points

- F. Anker, C. B., M. Eigel, M. Ladkau, J. Neumann, J. Schoenmakers: *SDE based regression for random PDEs*, preprint, 2016.
- C. B., A. Szepessy, R. Tempone: *Adaptive week approximation of stopped and reflected diffusion*, MCMA, 2010.
- G. Milstein, M. Tretyakov: *Stochastic numerics for mathematical physics*, Springer, 2004.
- G. Migliorati: *Analysis of the stability and accuracy of the discrete least-squares approximation on multivariate polynomial spaces*, presentation at UQAW 2016.