STOCHASTIC PDE'S: A MULTISCALE MODELLING CONCEPT

WOLFGANG DAHMEN

Abstract. The following are lecture notes for Prof. Dahmen's Spring 2008 High Dimensional Approximation Seminar. This lecture serves as an elaboration on some topics introduced by Prof. Yotov and also as a bridge to Dr. Wan's talk.

1. Introduction

Data in real life application are usually corrupt by noise and thud subject to uncertainties. This refers to the description of computational domains as well as to parameters appearing in partial differential equation modelling a physical or technological process. Even further, stochastic PDE's, that is, PDE's involving stochastic parameters, can be used to model processes involving to large a range of scales trying to capture the effect of unresolved scales on a macroscale.

Examples of such situations include flows in porous media, or heterogeneous materials like concrete. In these settings, the fine scales (pores and grains) are important to the larger scales (flows and crack propagation) but cannot be resolved by discretizations.

In this lecture we discuss a typical situation that briefly appeared already in an earlier lecture. We shall focus on a typical method of transforming a stochastic PDE into a deterministic PDE, which, however, may then depend on a large number of additional variables leading to an operator equation in high dimensions. If will be indicated that, under strong assumptions on the stochastic input, one can derive a priori estimates for the expectation of the output.

This can be seen as a special instances of learning an implicitly given function or manifold learning.

The primary purpose, however, is the lecture next week, which will focus on a posteriori estimates for such problems.

2. Setting

Here is the setting to be considered [2, 3]. Let $D \subset \mathbb{R}^d$ be a [bounded] Lipschitz domain given $f \in L_2(D)$ $(H^{-1}(D))$. We wish to find $u = u(\cdot, \omega)$ such that

$$-\nabla \cdot (a(\cdot,\omega)\nabla u)(x) = f(x) \text{ in } D \tag{2.1}$$

$$u|_{\partial D} = 0 \tag{2.2}$$

where

$$a(\cdot,\omega) : \omega \to L_\infty(D) =: X,\ \omega \in \Omega$$

Date: March 2008.

and (Ω, Σ, P) is a probability space on the data space X. Thus, the input $a(\cdot,\omega)$ of the PDE is a random field.

For simplicity assume that

$$\forall\, \alpha > 0 \text{ such that } \alpha \leq \inf_{x\in D} a(x,\omega) \leq \|a(\cdot,\omega)\|_{L_\infty} \leq \frac{1}{\alpha}. \tag{2.3}$$

Proposition 2.1. *(2.1) is well-posed, that is, if (2.3) holds then there exists a unique $u(\cdot,\omega) : H \to H_0^1(D)$ which is P-measurable.*

It follows from (2) that

$$a \in L_2(\Omega, X), \text{ that is } \int_\Omega \|a(\cdot,\omega)\|_X^2 dP < \infty. \tag{2.4}$$

3. Goal

Now that the output $u(\cdot,\omega)$ is a random variable, a typical objective is to compute moments of $u(\cdot,\omega)$. Let us focus here on the lowest order moment: The expectation,

$$\bar{u}(x) = \mathbb{E}_\Omega(u(x,\cdot)).$$

Note that this can *not* be achieved by just solving (2.1) with $a(\cdot,\omega)$ replaced by its mean,

$$\mathbb{E}(a(x,\cdot)) = \langle a\rangle(x).$$

In order to capture the effect of the fluctuation of the random field $a(\cdot,\omega)$ around its mean on the expectation of $u(\cdot,\omega)$ one considers expansions of the form

$$a(x,\omega) = \langle a\rangle(x) + \tag{3.1}$$

where

$$\tilde{a}(x,\omega) := \sum_{m=1}^{\infty} a_m(s) y_m(\omega)$$

is the fluctuation with a zero-mean. One possible approach is based on the Karhunen-Lowe expansion (cf. SVD, Hilbert-Schmidt[1]) [3] [2] [1], which will be described shortly.

[1]Hilbert-Schmidt Theory and SVD yield convergence in $L_2(D)$, but uniformly only for continuous C_a as per Mercer's Theorem.

$$C_a(x,x') = \sum_{m=1}^{\infty} \lambda + m \cdot \phi_m(x)\phi_m(x')$$

for

$$\mathcal{C}_a\phi_m = \lambda_m\phi_m = \int_D C_a(\cdot,x')\phi_m(x')dx'.$$

Hence, by KLE:

$$\tilde{a}(x,\omega) = \sum_{m=1}^{\infty} \langle \tilde{a}(\cdot,\omega), \phi_m\rangle_D \phi_m(x)$$

converges in $L_2(d \times \Omega)$

$$\langle \tilde{a}(\cdot,\omega), \phi_m\rangle_D = \int_D a(x,\omega)\phi_m(x)dx.$$

Note

$$\int_\Omega \langle \tilde{a}(\cdot,\omega), \phi_m\rangle_D dP(\omega) = 0$$

Given $a(\cdot,\omega) \in L_2(\Omega, X)$ define

$$\begin{aligned} c_a(x,x') &= \int_\Omega (a(x,\omega) - (\mathbb{E}a)(x))(a(x',\omega) - (\mathbb{E}a)(x'))dP(\omega) \\ &= c_a(x',x) \end{aligned} \tag{3.2}$$

and consider the compact integral operator

$$(\mathcal{C}_a v)(x) = \int_D c_a(x,x')v(x')dx' \tag{3.3}$$

to understand the meaning of $\mathcal{C}_a$ note that

$$\begin{aligned} \langle \mathcal{C}_a v, v\rangle_D &= \int_D v(x)\int_D v(x')\int_\Omega (a(x,\omega) - (\mathbb{E}a)(x))(a(x',\omega) - (\mathbb{E}a)(x'))dP(\omega)dxdx' \\ &= \int_\Omega (\int_D (a(x',\omega) - (\mathbb{E}a)(x'))dx')^2 dP(\omega) \\ &= \mathrm{Var}(\int_D a(x,\omega)v(x)dx) \\ &= \|\langle \tilde{a}(\cdot,\cdot), v\rangle_D\|^2_{L_2\Omega} \\ &\geq 0 \end{aligned}$$

since

$$\begin{aligned} \mathbb{E}(\int_D a(x,\omega)v(x)dx) &= \int_\Omega\int_D a(x,\omega)v(x)dxdP(\omega) \\ &= \int_D (\mathbb{E}a)(x)v(x)dx. \end{aligned}$$

Hence, there exists a sequence of eigenvalues $\{\lambda_m\}m = 1^\infty \subset \mathbb{R}_+$

$$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_m \to 0,\ m \to \infty$$

and a corresponding family of eigenfunctions $\{\phi_m\}_{m=1}^\infty \subset L_2(D)$ such that $\mathcal{C}_a\phi_m = \lambda_m\phi_m$ which form an orthonormal basis for $L_2(D)$.

Now, define

$$y_m(\omega) := \frac{1}{\sqrt{lambda_m}} \underbrace{\int_D \tilde{a}(x,\omega)\phi_m(x)dx}_{\langle \tilde{a}(\cdot,\omega),\phi_m(x)\rangle_D \text{a.k.a. the ``Fourier coefficients''}},\ m = 1, 2, \ldots. \tag{3.4}$$

Clearly,

$$\int_\Omega y_m(\omega)y_\ell(\omega)d\omega = 0 \text{ and } \|y_m\|L_2(\Omega) = 1.$$

---

and

$$\begin{aligned} \int_\Omega \langle \tilde{a}(\cdot,\omega),\phi_m\rangle_D^2 dP(\omega) &= Var(\int_D a(x,\omega)\phi_m(x)dx) \\ &= \langle \mathcal{C}_a\phi_m,\phi_m\rangle_D \\ &= \lambda_m\|\phi_m\|^2_{L_2(D)}. \end{aligned}$$

Hence, defining

$$y_m(\omega) = \lambda_m^{-1/2}\int_D a(x,\omega)\phi_m(x)dx$$

we have $\|y_m\|_{L_2} = 1$. The y_m are uncorrelated, but generally not stochastically independent.

Remark 3.1. The given i.i.d. realizations $a(x,\omega_j)$, $j = 1,2,\dots$ of $a(x,\omega)$ give rise to realizations $y_m(\omega_j)$ for each m.

In this case, (3.1) may take the form

$$a(x,\omega) = \langle a\rangle(x) + \sum_{m=1}^{\infty} \langle \tilde{a}(\cdot,\omega), \phi_m\rangle_D \phi_m(x)$$

and so we have

$$a(x,\omega) = \langle a\rangle(x) + \sum_{m=1}^{\infty} \sqrt{\lambda_m}\phi_m(x) y_m(\omega). \tag{3.5}$$

Since this expression cannot be computed in full on has to truncate

$$a_M(x,\omega) = \langle a\rangle(x) + \sum_{m=1}^{M} \sqrt{\lambda_m}\phi_m(x) y_m(\omega). \tag{3.6}$$

4. Decay Estimates

Further reading on the topics in this section may be found in [3].

Proposition 4.1. (1) *If $c_a x, x' \in C^k(D\times D)$ for $k > 1$ then for all $s > 0$ there exist $c_s > 0$ such that for all m we have*

$$\|\phi_m\|_{L_\infty} < c_s \lambda_m^{-s},\ \lambda_m \le cM^{-k/d}.$$

(2) *If $c_a(x,x') \in \mathcal{A}(D\times D)$ (analytic) then $\lambda_m \le ce^{-bm^{1/d}}$ and*

$$\|a - a_M\|_{L_\infty(D)} \le c \begin{cases} M^{\frac{k}{2D}+\varepsilon} \\ e^{-bM^{1/d}} \end{cases} \quad P\text{-a.s.}. \tag{4.1}$$

Note that (2) holds even if $c_a(x,x')$ is piecewise analytic.

Now, replace a by a_M in (2.1), then by (2.3) one still has

$$\frac{\alpha}{2} \le a_M(x,\omega) \le \frac{2}{\alpha}$$

so that

$$\begin{aligned} -\nabla\cdot(a_M(x,\omega)\nabla u_M) &= f \text{ in } \Omega \\ u_M|_{\partial\Omega} &= 0 \end{aligned}$$

is still well posed and

$$u_M = u_m(x; y_1(\omega),\dots,y_M(\omega))$$

satisfies

$$\|u - u_M\|_{L_\infty(\Omega, H_0^1(D))} \le C \begin{cases} M^{-k/d+\varepsilon} \\ e^{-bM^{1/d}} \end{cases}. \tag{4.2}$$

Next, one changes the measure to switch to a deterministic problem in many variables.

For simplicity, suppose that the $y_m(\omega)$ are uniformly distributed in $\left[-\frac{1}{2},\frac{1}{2}\right]$, that is,

$$(y_1(\omega),\dots,y_M(\omega)) \in \left[-\frac{1}{2},\frac{1}{2}\right]^M,\ dP_M = dy_1\dots dy_M = d\vec{y}.$$

Thus, one arrives at

$$-\nabla \cdot (a_M(x,\vec{y}))\nabla \hat{u}_M = f\mathrm{D} \times \left[-\frac{1}{2},\frac{1}{2}\right]^M \tag{4.3}$$
$$\hat{u}_M|_{\partial\Omega} = 0$$

(4.4)

$$\hat{u}_M(x,\vec{y}) \in L^\infty(\left[-\frac{1}{2},\frac{1}{2}\right]^M, H_0^1(D))$$
$$a_M(x,\vec{y}) = \langle a\rangle(x) + \sum_{m=1}^{M} \lambda_m^{1/2}\phi_m(x) y_m.$$

In case of rapid decay, we have

$$\hat{u}_M(x,\vec{y}) \in \mathcal{A}(\left[-\frac{1}{2},\frac{1}{2}\right]^M, H_0^1(D)) \text{ and}$$
$$\hat{u}_M(x, y_1(\omega), \ldots, y_M(\omega)) = u_M(x, y_1(\omega), \ldots, y_M(\omega)).$$

Finally, we must discretize (4.3) over some finite dimensional trial space.

5. Comments and Suggestions

An alternative strategy could be formulated based on realizations of $a(\cdot,\omega_i)$, computing $u(cdot,\omega_i)$, and using Monte Carlo methods.

One advantage of this alternative strategy is that no information on the underlying probability distribution is needed.

However, there are several disadvantages:

- The convergence is slow; n realizations give $o(n^{-1/2})$ accuracy. In other words, the PDE has to be solve frequently with very little chance of exploiting local phenomena in solutions.
- Sometimes one few realizations of $a(\cdot,\omega_i)$ are available due to experimental constraints.

Given these drawbacks to the Monte Carlo-based method, we still have the question as to whether the approach described in this lecture can beat it in terms of work-accuracy balance. In order to answer this question fairly, we need to consider some points:

- One need strong "stochastic information," for example, that the $c_a(x,x')$ are assumed to be *known*. If not, one would have to learn them data, again requiring many realizations of $a(\cdot,\omega)$. How would the resulting errors effect the outcome?
- The ϕ_m need to be computed. This results in an eigenvalue problem.
- Different errors need to be balanced. Particularly those involved in the truncation from a to a_M, in the calculation of the ϕ_m, and in the Galerkin discretization are of concern. *A priori* estimates are simply not quantitative enough.
- The effective dimension of the problem is unknown. How can one tell which stochastic components in the expansion of a are relevant? In this aspect, could nonlinear recovery techniques such as L_1-recovery or sparce recovery be of help?

One idea to address these points is the consideration of a posteriori estimates. These will be covered in Wan's talk next week.

References

1. M Loeve, *Probability theory*, Springer-Verlag, Berlin, 1977.
2. H.G. Matthies and A. Keese, *Galerkin methods for linear and nonlinear elliptic stochastic partial differential equations*, Tech. report, Institut fur Wissenschaftlishes Rechmen, TV Braunshweig, July 2004.
3. R.A. Todor and C. Schwab, *Convergence rates for sparse chaos approximations of elliptic problems with stochastic coefficients*, Tech. Report 2006-05, ETH Zurich, 2006.

University of South Carolina
E-mail address: dahmen@...