Stochastic Inversion Under Functional Uncertainties

M.R. El Amri
University of Grenoble Alpes, IFPEN

Supervisor(s): C. Prieur (UGA), C. Helbert (ICJ, ECL), D. Sinoquet (IFPEN), M. Munoz Zuniga (IFPEN), O. Lepreux (IFPEN)

Ph.D. expected duration: 2016 - 2019

Adress: Bâtiment IMAG, 700 Avenue Centrale, 38401 Saint-Martin-d'Hères

Email: mohamed-reda.el-amri@ifpen.fr

Abstract:

In this present work we propose a new efficient method for solving an inversion problem under functional uncertainties. Without loss of generality, the considered numerical simulator takes inputs that can be divided into two sets, the deterministic control variables and the functional random variables. Let $f : \mathbb{X} \times \mathcal{V} \to \mathbb{R}$ denote the output of the simulator where $\mathbb{X} \subset \mathbb{R}^p$ is the search space of the control variables and the randomness of the simulator is derived from a random input defined in a functional space $\mathcal{V}$. We assume that the probability distribution of the functional input is only known through a finite set of realizations $\Xi = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ and each evaluation of f involves a time consuming computation. Let $\mathcal{X}_n = \{\mathbf{x}^1, \dots, \mathbf{x}^n\}$ and $\mathcal{V}_n = \{\mathbf{v}^1, \dots, \mathbf{v}^n\}$ denote the initial design points. The simulator responses at these design points are denoted by $\mathbf{Z} = \{f(\mathbf{x}^1, \mathbf{v}^1), \dots, f(\mathbf{x}^n, \mathbf{v}^n)\}$.

The objective of this study is to estimate the set

$$\Gamma^* = \{\mathbf{x} \in \mathbb{X} \ , \ g(\mathbf{x}) = \mathbb{E}_{\mathbf{V}}[f(\mathbf{x}, \mathbf{V})] \leq c\}, \tag{1}$$

where $c \in \mathbb{R}$ and $\mathbb{E}_{\mathbf{V}}$ is the expectation with respect to $\mathbf{V}$. The estimation of Γ^* based on computing function g at each grid point of discretized version of $\mathbb{X}$ requires far too many expensive simulations of f. Therefore, statistical methods based on a reduced number of evaluation points are widely used to overcome this latter difficulty by focusing the evaluations on the promising subregion of the control space.

Among statistical models, Gaussian Process (GP) model has received increasing interest in recent years, due to many of its good properties, such as the existence of explicit formula of statistical moments and the easy computation of the uncertainties on predictions. However, in the literature the input variables involved in Gaussian Process models are often univariate or multivariate. The purpose of this talk is first to extend the use of Gaussian Process model to cases where the inputs contain infinite dimensional variables or functional data which are collected as curves. Then we define an infill sampling criterion based on the GP model in order to solve the stochastic inversion problem (1).

After a dimension reduction of the functional space ($\mathbf{V} \in \mathcal{V} \xrightarrow[\text{Karhunen-Loève}]{} \mathbf{U} \in \mathbb{R}^{m_{\text{KL}}}$), a GP model $Z_{(\mathbf{x}, \mathbf{u})}$ is built in the joint space of control and uncertain variables $(\mathbf{x}, \mathbf{u})$. Then an averaged GP over the random variables is derived [2], $Y_{(\mathbf{x})} = \mathbb{E}_{\mathbf{U}}[Z_{(\mathbf{x}, \mathbf{u})}]$. Therefore the 'induced' Gaussian process approximates the expected response involved in inversion problem (1). Driven by the Gaussian process conditioned on the n observations $Y_{(\mathbf{x})}^n$, the proposed infill strategy consists of two steps: first we choose $\mathbf{x}^{n+1}$ by minimizing the Vorob'ev deviation [1]. Secondly we choose the uncertain point $\mathbf{u}^{n+1}$ that minimizes the variance of the process $Y_{(\mathbf{x})}^n$ evaluated at the point $\mathbf{x}^{n+1}$. By iterating this procedure, we expect a rapid estimation of the set Γ^* with a limited number of calls to the simulator f.

To illustrate the performance of the proposed method we compare it with two approaches. The first one is the nested MC approach, where the expectation with respect to $\mathbf{V}$ is estimated by Monte Carlo simulations using the simulator f, and the inversion is based on a GP model in the control space $\mathbb{X}$. The accuracy level of the MC estimation of the expectation is controlled by the number of calls k to the underlying simulator $f(\mathbf{x}, \mathbf{v})$.

The second method used for comparison is similar to the first one except that the expectation is estimated by a *Greedy Functional Quantizer* of size k. Briefly, the *Functional Quantization* of a functional random variable $\mathbf{V}$ consists of its approximation by a discrete random variable $\hat{\mathbf{V}}_k$ valued in $\Theta_k = \{\hat{\mathbf{v}}_1, \ldots, \hat{\mathbf{v}}_k\}$ [3]. The "grid" (or codebook) Θ_k minimizes over $(\mathcal{V})^k$ the quantization error induced by replacing $\mathbf{V}$ by $\hat{\mathbf{V}}_k$.

After presenting comparative results on a toy problem, we focus on the results obtained for an automotive industrial test case where the objective is to identify the set of control parameters leading to meet the pollutant emissions standards of a vehicle.

References

[1] Clément Chevalier. *Fast uncertainty reduction strategies relying on Gaussian process models.* PhD thesis, Citeseer, 2013.

[2] Janis Janusevskis and Rodolphe Le Riche. Simultaneous kriging-based estimation and optimization of mean response. *Journal of Global Optimization*, 55(2):313–336, 2013.

[3] Harald Luschgy, Gilles Pagès, and Benedikt Wilbertz. Asymptotically optimal quantization schemes for gaussian processes on hilbert spaces. *ESAIM: Probability and Statistics*, 14:93–116, 2010.

Short biography – I obtained a MSc in Applied mathematics : Computer and Stochastic Methods for Decision in 2015 at Université de Pau et des Pays de l'Adour. I started my PhD with IFPEN and University Grenoble Alpes in April 2016 in the OQUAIDO project (Optimisation et QUAntification d'Incertitudes pour les Données Onéreuses).