Sparse regression-based Poincaré expansions for global sensitivity analysis

N. Lüthen
ETH Zürich

Supervisor(s): Prof. B. Sudret (ETH Zürich)

Ph.D. expected duration: Apr. 2018 - Mar. 2022

Address: Chair of Risk, Safety & Uncertainty Quantification, Stefano-Franscini-Platz 5, 8093 Zürich

Email: nluethen@ethz.ch

Joint work with: F. Gamboa (Institut de Mathématiques de Toulouse), B. Iooss (EDF R&D), S. Marelli (ETH Zürich) and O. Roustant (INSA Toulouse).

Abstract:

Surrogate modeling is a widely used technique in the field of uncertainty quantification, which reduces the computational cost of uncertainty analyses by introducing a few additional assumptions on the regularity of the model of interest. More precisely, a few evaluations of the computational model are used to construct an accurate proxy, the so-called *surrogate model*, which typically has a simpler form than the original computational model and can be evaluated at much smaller cost. Subsequent analysis such as uncertainty propagation or sensitivity analysis can then be performed using the inexpensive surrogate model. In this work, we assume that in addition to model evaluations, partial derivatives of the model are available. These might be obtained e.g. from automatic differentiation techniques, or when using adjoint methods to solve differential equations.

A popular surrogate modeling method is *polynomial chaos expansions (PCE)*, due to their rigorous mathematical foundation and good performance in applied science and engineering. A PCE is nothing but the expansion of the model of interest onto an orthogonal polynomial basis of the Hilbert space $\mathcal{L}^2_{f_{\boldsymbol{X}}}$ induced by the distribution $f_{\boldsymbol{X}}$ of the input random vector $\boldsymbol{X}$. For independent input random variables $X_1, \dots, X_d$, the corresponding basis simplifies to the tensor product of univariate polynomial bases of $\mathcal{L}^2_{f_{X_i}}$, $i = 1, \dots, d$, each induced by its respective marginal distribution f_{X_i}. A further advantage of PCE is that many quantities of interest, such as moments and Sobol' sensitivity indices, can be obtained by postprocessing the expansion coefficients [3].

Of course, there are many other orthogonal bases of $\mathcal{L}^2_{f_{\boldsymbol{X}}}$ that could be used for an expansion. One recently suggested alternative is the so-called *Poincaré expansion* [1, 2]. Assuming that the input random variables are independent, the multivariate basis functions are tensor products of univariate orthogonal functions just as in the case of PCE. The univariate, in general non-polynomial, orthogonal systems are the eigenfunctions of the associated so-called *Poincaré differential operator*, which is defined as follows. For input variable X_i, define V by $f_{X_i}(x_i) = \exp(-V(x_i))$. Assume that the support of f_{X_i} is in a bounded interval $[a, b]$, and that V is continuous and piecewise C^1 on this interval. Then the Poincaré differential operator is defined by $Lh := h'' - V'h'$ and the associated eigenproblem is to find $h \in \mathcal{H}^2_{f_{X_i}}$ that fulfills

$$h'' - V'h' = \lambda h, \tag{1}$$
$$h'(a) = h'(b) = 0. \tag{2}$$

Solutions of this equation constitute a basis for $\mathcal{L}^2_{f_{X_i}}$, with the first one being the constant function associated to $\lambda = 0$.

A consequence of this definition of the Poincaré eigenfunctions is the following equality of inner products:

$$\langle h', e_n' \rangle_{f_{X_i}} = \lambda_n \langle h, e_n \rangle_{f_{X_i}} \tag{3}$$

for any $h \in \mathcal{H}^1_{f_{X_i}}$ and any Poincaré eigenfunction e_n, which implies the remarkable property that partial derivatives of a multivariate Poincaré basis form again an orthogonal system in $\mathcal{L}^2_{f_{\boldsymbol{X}}}$.

Just as for PCE, Sobol' indices can be computed analytically from the expansion coefficients of a Poincaré expansion, using that the expansion is an approximation of the Sobol'-Hoeffding decomposition [3]. Partial variances for a subset of variables are computed as the sum of squared coefficients associated to terms involving these variables. In the case of Poincaré expansions, [2] suggested a second possibility for computing partial variances, namely to use the partial derivatives of Poincaré basis functions to approximate the partial derivatives of the model, and to convert the obtained coefficients to the original coefficients using Eq. (3). Using the projection method for the computation of the coefficients, [2] demonstrated that Sobol' indices can be computed accurately from Poincaré expansions, especially when relying on the derivative approximation.

In this work, we use sparse regression techniques to compute sparse Poincaré (derivative) expansions, in analogy to sparse PCE which has shown to outperform projection-based and non-sparse regression-based PCE. We investigate the performance of sparse regression for Poincaré (derivative) expansions for computing Sobol' indices on several example problems, analyze their convergence and compare the results to results obtained with PCE. Preliminary results show that sparse regression estimates outperform projection estimates, and that estimates of Sobol' indices computed from Poincaré derivative expansions are especially accurate for low-importance variables, suggesting that this method might be a promising new tool for low-cost screening.

References

[1] O. Roustant, F. Barthe, and B. Iooss. Poincaré inequalities on intervals – application to sensitivity analysis. *Electronic Journal of Statistics*, 11(2):3081–3119, 2017.

[2] O. Roustant, F. Gamboa, and B. Iooss. Parseval inequalities and lower bounds for variance-based sensitivity indices. *Electronic Journal of Statistics*, 14(1):386–412, 2020.

[3] B. Sudret. Global sensitivity analysis using polynomial chaos expansions. *Reliab. Eng. Sys. Safety*, 93(7):964–979, 2008.

Short biography – Nora Lüthen has studied mathematics with an emphasis on numerical mathematics at the University of Bonn, Germany. Since 2018, she is a PhD student in Prof. Sudret's Chair of Risk, Safety and Uncertainty Quantification at ETH Zürich in Switzerland. Her PhD is part of the project "Surrogate modelling for stochastic simulators" funded by the Swiss National Science Foundation.