Dimensionality reduction and surrogate modelling for high-dimensional UQ problems

C. Lataniotis
ETH Zürich

Supervisor(s): Prof. Dr. B. Sudret (ETH Zürich) and Dr. S. Marelli (ETH Zürich)

Ph.D. expected duration: 2015-2019

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

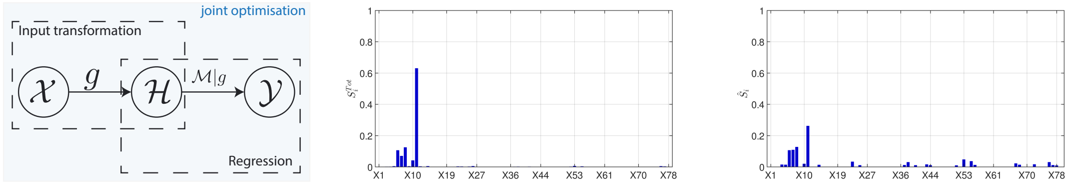
Email: lataniotis@ibk.baug.ethz.ch

Abstract:

Uncertainty Quantification (UQ) problems can be challenging in high dimensions. Powerful techniques like surrogate modelling can become infeasible when the number of input parameters is large. Moreover, it can be the case that accurate probabilistic models of the uncertain input parameters are not readily available, but only samples of their realisations. In order to deal with this issue some reduction of the input space dimension needs to be performed, either in terms of input variable selection (*i.e.* selecting the "most important" inputs) or mapping the data to a lower dimensional space using an appropriate transformation. Various techniques have been suggested or extended during the past decade in machine learning communities for reducing the input space in a data-driven fashion (see *e.g.* [4]). Nevertheless, there is still comparably little literature available on combining "the best of two worlds", that is closely coupling surrogate modelling and dimensionality reduction approaches. To that end, the goal of this contribution is to propose a new approach of optimally combining dimensionality reduction and surrogate modelling in data-driven contexts, in view of real-world engineering applications involving systems with a large number of input parameters. The concept of the proposed approach is based on recent advances in neural network-based strategies (*a.k.a.* deep learning) [2] where the dimensionality reduction step is coupled with training of slightly modified networks for regression or classification tasks. Extensions of this concept exist, *e.g.* the so-called manifold Gaussian processes [1] couple neural-networks for dimensionality reduction (*a.k.a.* autoencoders) with Gaussian processes.

To formally state the problem, a set of observations $\mathcal{X} = \{\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(N)}\,;\, \mathbf{x}^{(i)} \in \mathbb{R}^M\}$ and system responses $\mathbf{y} = \{y_1 = \mathcal{M}(\mathbf{x}^{(1)}), \dots, y_N = \mathcal{M}(\mathbf{x}^{(N)})\,;\, y_i \in \mathbb{R}\}$ is considered. Dimensionality reduction stands for a transformation g that maps the observations to a lower dimensional space, *i.e.* to a set $\mathcal{Z} = g(\mathcal{X}; \mathbf{w})$ with $\mathcal{Z} = \{\mathbf{z}^{(1)}, \dots, \mathbf{z}^{(N)}\,;\, \mathbf{z}^{(i)} \in \mathbb{R}^m\}$ with $\mathbf{w}$ a set of parameters associated with the particular method that is considered and $m < M$. Linear dimensionality reduction techniques have been extensively used in various fields to tackle this problem. The most common technique is arguably principal component analysis (PCA). However, linear transformations such as PCA can provide limited performance if the inter-dependence of the input variables is non-linear, *i.e.* $\mathcal{X}$ lies onto a non-linear manifold that is embedded in the high-dimensional input space. Consequently over the past decade there have been significant advances towards non-linear techniques. A family of non-linear dimensionality reduction methods that has gathered significant attention is kernel PCA [3], that consists in computing PCA in some high-dimensional space (possibly infinite dimensional) via the so-called *kernel trick* [5].

Kernel PCA requires one to select an appropriate kernel function and tune its parameters. An optimal set of parameters $\mathbf{w}$ is in practice obtained by minimizing some objective function related to the compressive performance of the transformation, typically the reconstruction error. However a preliminary analysis of several different non-linear compression schemes revealed that compression



(a) (b) (c)

Figure 1: (a) The proposed scheme for performing dimensionality reduction and surrogate modelling, (b) polynomial chaos expansions-based total Sobol' sensitivity indices of a real world data set with 78 input variables and (c) approximate sensitivity indices extracted from the optimal kernel PCA parameters using the proposed methodology.

efficiency is not correlated with the performance of the surrogate models in the reduced space. To this end, the goal of this contribution is to propose a new approach of optimally combining dimensionality reduction and surrogate modelling in data-driven contexts.

The proposed method treats each of the aforementioned steps in a non-intrusive (*i.e.* black-box) way, thus providing a common algorithmic foundation to deploy various combinations of dimensionality reduction and surrogate modelling methodologies (Figure 1a). In the current contribution the case of kernel PCA-based dimensionality reduction coupled with Gaussian process modelling (*a.k.a.* Kriging) as the surrogate is considered. A trade-off between the complexity of the surrogate and the computational cost of the optimization is observed which leads to the choice of simplified Kriging surrogates (using isotropic correlation families) coupled with anisotropic kernels for kernel PCA such as the Gaussian kernel (also known as *automatic relevance determination kernel*). A by-product of this particular combination is that useful insights in terms of the importance of each of the input variables with respect to the model output can be gained by a simple post-processing step. The effectiveness of such configuration is evaluated on benchmark and real-world applications. An example is shown in Figures 1b and 1c where total Sobol' sensitivity indices are compared against approximate sensitivity indices obtained by post-processing the optimal kernel-PCA parameters on a pre-existent subsurface flow dataset.

References

[1] R. Calandra, J. Peters, C. E. Rasmussen, and M. P. Deisenroth. Manifold gaussian processes for regression. In *International Joint Conference on Neural Networks (IJCNN)*, 2016.

[2] G. E Hinton. Learning multiple layers of representation. *Trends in cognitive sciences*, 11(10):428–434, 2007.

[3] B. Schölkopf, A. Smola, and K. Müller. Nonlinear component analysis as a kernel eigenvalue problem. *Neural Comput.*, 10(5):1299–1319, 1998.

[4] L.J.P. van der Maaten, E. O. Postma, and H. J. van den Herik. Dimensionality reduction: A comparative review. *J Mach Learn Res*, 10:66–71, 2008.

[5] V. N. Vapnik. *Statistical Learning Theory.* Wiley-Interscience.

Short biography – Christos Lataniotis holds a diploma in mechanical engineering (National Technical University of Athens), MSc in Robotics (ETH Zürich) and is currently a doctoral candidate at the Chair of Risk, Safety and Uncertainty Quantification (ETH Zürich). His PhD project is entitled "Machine learning-based uncertainty quantification algorithms for risk analysis of engineering systems". Prior to the start of his PhD C. Lataniotis contributed to the development of the UQLab software developed by the Chair of Risk, Safety and Uncertainty Quantification.