PC-Kriging: Combining Polynomial Chaos Expansions and Universal Kriging

R. Schöbi
ETH Zürich

Affiliation: *Chair of Risk, Safety and Uncertainty Quantification*, ETH Zürich, Stefano-Franscini-Platz 5, CH-8093 Zürich, Switzerland

Email: schoebi@ibk.baug.ethz.ch – **URL**: http://www.ibk.ethz.ch/su/people/schoebir/index_EN

Master: ETH Zürich

Ph.D. (2013-2016): ETH Zürich

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

Abstract:

There is an ongoing trend for replacing real life experiments by computer simulations. The physical behaviour is substituted by a computational model which approximates the system response of the physical system. Advances in research lead to more complex and more accurate computational models which are at the same time more costly to evaluate, *i.e.* time-consuming. There is a conflicting situation between accuracy and speed. Applications such as reliability analysis or optimization algorithms require a large number of model evaluations, *e.g.* the computation of a system's failure probability or the optimal value of a parameter. These operations are reasonable when the computational model is easy-to-evaluate, *i.e.* when a model evaluation is inexpensive and the system response of a large number of input samples is processed fast.

Metamodelling provides a framework for replacing an expensive-to-evaluate computational model $Y = \mathcal{M}(\boldsymbol{X})$ by a simple although approximative surrogate model. The surrogate model (also called metamodel) allows one to predict the system response of a large number of input samples at low cost. The metamodel is built from a *small* number N of support points called the experimental design $\boldsymbol{\mathcal{X}} = \{\boldsymbol{\chi}^{(i)},\, i = 1, \ldots, N\}$ for which the original model is evaluated. The input and output values/vectors are used to determine an appropriate metamodel with a certain metamodelling technique. Two of the more popular non-intrusive metamodelling techniques are *Polynomial Chaos Expansions* (PCE) and *Kriging* (also called Gaussian process modelling).

PCE surrogates the computational model $\mathcal{M}$ by a finite set of orthonormal polynomials in the input variables (Ghanem and Spanos, 2003). In the context of uncertainty quantification, the latter are defined in coherency with the probability distribution functions of those input variables. The coefficients of a PC expansion may be computed using *e.g.* least-square minimization algorithms. PCE assumes that the computational model is a black-box model, *i.e.* only information about the input values and model response are available (the inner structure and features of the model (nonlinearity, interaction between parameters, etc.) are assumed unknown).

Kriging is called a stochastic metamodelling technique which assumes that the computational model is a realization of a Gaussian random field whose properties are inferred from the experimental design and the associated model output (Santner et al., 2003). The experimental design points provide the information to compute the optimal correlation parameters by *e.g.* maximum likelihood method. The prediction of the surrogate at a new point results in a Gaussian variable represented by its mean value and variance value called Kriging mean prediction and prediction variance.

Although these two techniques have become popular for solving uncertainties propagation, optimization or sensitivity problems, their combination has not been considered yet. In this paper, the new metamodelling technique *Polynomial-Chaos-Kriging* (PC-Kriging) is proposed. This metamodel is based on the classical universal Kriging approach where the trend (regression part) is a sum of functions in the

general case. In PC-Kriging a sparse set of orthonormal polynomials serves as the trend of the universal Kriging model. The general formulation of the metamodel is then:

$$\mathcal{M}(\boldsymbol{X}) \approx \mathcal{M}^{(\mathsf{PCK})}(\boldsymbol{x}) = \sum_{k=1}^{P} \beta_k \, f_k(\boldsymbol{x}) + \sigma^2 \, Z(\boldsymbol{x}, \omega) \tag{1}$$

where $\sum_{k=1}^{P} \beta_k \, f_k(\boldsymbol{x})$ is the mean value of the Gaussian process (so-called trend) and $Z(\boldsymbol{x}, \omega)$ is a zero mean, unit variance Gaussian process described by a set of hyper-parameters $\omega = \{\boldsymbol{\theta}, R\}$. The autocorrelation function $R(\boldsymbol{x}, \boldsymbol{x}'; \boldsymbol{\theta})$ describes the correlation between two samples given its parameters $\boldsymbol{\theta}$. $\boldsymbol{f}(\boldsymbol{x}) = \{f_k(\boldsymbol{x}), k = 1, \ldots, P\}$ are the multidimensional orthonormal polynomials, its coefficients are β_k, $\boldsymbol{x}$ are realizations of the input variables $\boldsymbol{X}$, σ^2 is the Kriging variance. The sparse set of P polynomials is determined by using hyperbolic index sets and least angle regression based of the experimental design $\{\boldsymbol{\chi}^{(i)}\}$ (Blatman and Sudret, 2011). In an iterative manner and one-by-one, a polynomial out of the determined sparse set is added to the Kriging model. The iteration starts with the polynomial which is the most correlated to the system response. Then the classical equations to fit the Kriging model are used to determine the correlation parameters, the Kriging variance and the trend coefficients (Bachoc, 2013). The P PC-Kriging models are then compared by means of the leave-one-out error and the optimal PC-Kriging metamodel (with minimal leave-one-out error) is chosen.

The performance of PC-Kriging is compared to ordinary Kriging (constant trend β_0) and pure PCE on six easy-to-evaluate benchmark problems in the field of optimization and metamodelling. The results show that PC-Kriging performs better or at least as good as PCE and/or Kriging. Especially, for small experimental designs PC-Kriging is preferable to the two distinct approaches. From the numerical experiments is appears that some problems are better suited for PCE whereas some problems are better handled by Kriging. PC-Kriging converges to the best of the two simple approaches and leads to a smaller squared residual error. For large sample sizes, PC-Kriging performs similar to PCE, so that the added value of PC-Kriging is questionable in that case.

Heuristically, the behaviour of PC-Kriging can be explained as follows: the set of polynomials approximates the global behaviour whereas the correlation part models the local variabilities between the support points. The combination of these two effects leads to a higher accuracy and thus to a better metamodel. The validation of PC-Kriging is shown on numerous analytical benchmark functions which are easy-to-evaluate.

References

Bachoc, F. (2013). Cross validation and maximum likelihood estimations of hyper-parameters of Gaussian processes with model misspecifications. *Comp. Stat. Data An. 66*, 55–69.

Blatman, G. and B. Sudret (2011). Adaptive sparse polynomial chaos expansion based on Least Angle Regression. *J. Comput. Phys 230*, 2345–2367.

Ghanem, R. and P. Spanos (2003). *Stochastic Finite Elements : A Spectral Approach.* Courier Dover Publications.

Santner, T., B. Williams, and W. Notz (2003). *The Design and Analysis of Computer Experiments.* Springer, New York.

Short biography – R. Schöbi graduated in 2012 from ETH Zürich with a Master's degree in civil engineering where his Master's Thesis was entitled "Subset Simulation in Engineering Problems." Since May 2013, he is a Ph.D. student at the Chair of Risk, Safety and Uncertainty Quantification of ETH Zürich under the supervision of Prof. B. Sudret. His research topics include the quantification and propagation of epistemic uncertainty using mixed probabilistic/nonprobabilistic approaches.