Sensitivity Analysis and FANOVA Graphs for Computer Experiments

Jana Fruth and Sonja Kuhnt

Abstract In industrial research non-stochastic simulation models such as finite element simulations often replace real experiments nowadays. Such computer experiments can be highly complex and time-consuming, therefore it has become common practice to fit easier to evaluate metamodels. The most popular metamodel is the Gaussian process model, also known as Kriging model.
Sensitivity analysis investigates how the input variables contribute to the variation of the outcome of the experiment. A popular method is the use of Sobol indices which quantify the importance of individual input variables or groups of them.
This contribution first introduces computer experiments and sensitivity analysis. We then present a new procedure described in [3] where a different kind of index, the total interaction index, is applied to provide a deeper insight into the interaction structure of the unknown function which can be displayed in a so-called FANOVA graph. The information contained in the graph can then be used to derive data-driven kernels to fit improved Kriging models. Finally we show an application to a computer experiment in sheet metal forming.

Key words: Computer Experiments, Sensitivity Analysis, Graph, Kriging

1 Kriging

In computer experiments, Kriging is a standard tool for prediction and analysis of expensive functions (see e.g. [5], [2]). We denote the outcome of the computer experiment for a vector of input variables $\mathbf{x}$ by $Y(\mathbf{x})$. Then the assumed Kriging model is

Jana Fruth
Faculty of Statistics, TU Dortmund University, e-mail: fruth@statistik.tu-dortmund.de

Sonja Kuhnt
Faculty of Statistics, TU Dortmund University, e-mail: kuhnt@statistik.tu-dortmund.de

$$Y(\mathbf{x}) = \sum_{k=1}^{p} \beta_k f_k(\mathbf{x}) + Z(\mathbf{x}) \quad (1)$$

where $\sum_{k=1}^{p} \beta_k f_k(\mathbf{x})$ models the trend at $\mathbf{x}$ and $Z(.)$ a centred Gaussian process with covariance function, or kernel K. It is assumed that $Z(.)$ is stationary, thereby K depends only on the difference between two points $\mathbf{x}^1$ and $\mathbf{x}^2$. K can therefore be written as $K(\mathbf{x}^1, \mathbf{x}^2) = k(\mathbf{x}^1 - \mathbf{x}^2)$ with $k(.) = \sigma^2 R(.;\theta)$, σ^2 the process variance, R the correlation function and θ a vector of parameters. Since the departure from the trend relies on the kernel, its specification is an essential part of model building. In computer experiments, kernels are usually obtained as tensor products of 1-dimensional kernels g_k (e.g. Gaussian or Matérn 5/2):

$$k(\mathbf{h}) = \sigma^2 \prod_{k=1}^{d} g_k(h_k; \theta_k) \, . \quad (2)$$

One reason for the success of Kriging is that it interpolates the data, which is desirable for deterministic functions like computer experiments. It also gives a measure of uncertainty at unknown points, due to its probabilistic nature.

2 Sensitivity Analysis

Following the definition of [4] sensitivity analysis is the study of how the variation in the output of a model can be apportioned to different sources of variation. It is an important tool in the understanding and interpretation of the model, and also for variable ranking and reduction. There are several different methods for sensitivity analysis including scatter plots, correlations coefficients, screening methods and regression analysis which are restricted to local sensitivities or to specific model behaviours. A very popular global and model independent method for sensitivity analysis is given by Sobol indices. This method is based on the so-called *Functional ANOVA* (FANOVA) decomposition ([1], [6]).

Let X be a random vector over a domain Δ with integration measure $\mathrm{d}\nu$ and independent components $X_1, \dots, X_d$. Consider a function f such that $f(X)$ is square integrable. Then the FANOVA decomposition of f is given by:

$$f(X) = \mu_0 + \sum_{i=1}^{d} \mu_i(X_i) + \sum_{j<k} \mu_{jk}(X_j, X_k) + \sum_{j<k<l} \mu_{jkl}(X_j, X_k, X_l) + \dots + \mu_{12\dots d}(X_1, X_2, \dots, X_d) \quad (3)$$

where each term is centred ($E(\mu_J(X_J)) = 0$) and orthogonal ($E(\mu_J(X_J)\mu_{J'}(X_{J'})) = 0, \forall J' \neq J$). For $i = 1, \dots, d$ the term $\mu_i(x_i)$ can be interpreted as main effects of X_i, for $j < k$ the term $\mu_{j,k}(x_j, x_k), j < k$ as second order interaction effects and so on.

Taking the overall variance of the function the FANOVA decomposition leads to an ANOVA like variance decomposition where the variance of each term μ_J gives a sensitivity index for its effect: $D_J = \text{Var}(\mu_J(X_J))$. Sobol indices are obtained through standardization by the overall variance:

$$S_J = \frac{D_J}{\text{Var}\,(f(X))} \,. \tag{4}$$

They are an attractive tool for investigating a function f as they measure the importance of main effects and interactions of any order and do not require limiting assumptions. Their calculation can be done numerically [4]. A related index is the total effect S_k^T, which is the sum of all indices S_J with $k \subseteq J$. It is frequently used as a measure for input variable importance since it includes the single as well as interaction effect of the variable.

3 Kriging Kernels from FANOVA Graphs

In statistics mathematical graphs, defined as tuple $G = (V, E)$ of a finite set of vertices V and a set of edges E combining the vertices in V, are used in different contexts, e.g. for variable selection and for modelling dependence structure of random vectors. Here the intention is to deal with the *curse of dimension* in sensitivity analysis: in general there are $2^d - 1$ terms in the functional decomposition which is even for medium values of d, e.g. $d = 5$, a huge amount. Therefore often only main effects and total effects are considered. Here a methodology is suggested [3], which reduces the number of effects to be calculated but still gives good insight in the interaction structure of the function f by using graphs.

A graph, called *FANOVA graph*, is set up such that each vertex in V represents one input variable. The basic idea is that two vertices/input variables X_j, X_k are connected in the graph if there is any term index set J which includes j,k with $\mu_J(x_J) \not\equiv 0$. An index to identify the edges is the so-called *total interaction index*, which only takes a positive value if the vertices are connected. It is defined by:

$$\mathfrak{D}_{ij} := \text{Var}\left(\sum_{I \supseteq \{i,j\}} \mu_I(X_I)\right) = \sum_{I \supseteq \{i,j\}} D_I \,. \tag{5}$$

Assuming that a FANOVA graph is given, maximal complete subgraphs (cliques) $C_1, \ldots, C_L$ can be identified resulting in an additive decomposition of f. Now in our framework, we can specify Kriging kernels to take advantage of the additional knowledge given by the FANOVA decomposition, since it implies an additive decomposition for the kernel of Z

$$k(\mathbf{h}) = \sum_{l=1}^{L} k_{C_l}(\mathbf{h}_{C_l}) \tag{6}$$

where each k_{C_l} is a kernel defined on the input variables given by the clique C_l. This data-driven adaptation can improve the Kriging fit over the tensor-product kernel (2) since it adapts the special interaction structure of f.

4 Application

The procedure is applied to data from a case study for the springback analysis during sheet metal forming in [3], where the output reflects the amount of springback after the forming process. The experiment has been simulated by the engineering software Autoform. There are 3 independent input variables and a 3^3 full factorial design is available as learning data set as well as another 101 runs for validation purposes. Applying the total interaction indices for estimating the graph results in a graph with just on edge $(1,3)$ illustrated in Fig. 1, which can be interpreted such that the second influence parameter just has additive influence. Adapting the Kriging model according to the graph and applying it to the validation data set results in an increase in precision of about 15% compared to the standard Kriging model (RMSE standard Kriging model: 0.0867, RMSE adapted Kriging model: 0.0758).

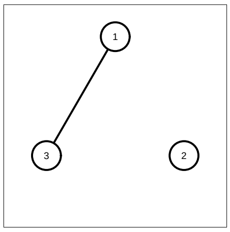

Fig. 1 Resulting FANOVA graph for the Autoform data set. Only edge (1,3) is active.

Acknowledgements This contribution stems from joint work with T. Muehlenstaedt and O. Roustant. It is based on investigations of the collaborative research centre SFB 708, project C3. The authors would like to thank the Deutsche Forschungsgemeinschaft (DFG) for funding.

References

1. Efron, B., Stein, C.: The jackknife estimate of variance. Ann. Stat. **9**(3), 586–596 (1981)
2. Fang, K., Li, R., Sudjianto, A.: Design and modeling for computer experiments. Computer science and data analysis series. Chapman & Hall/CRC, Boca Raton and London (2006)
3. Muehlenstaedt, T., Roustant, O., Carraro, L., Kuhnt, S.: Data-driven Kriging models based on FANOVA-decomposition. Stat. Comput. **22**, 723–738 (2012)
4. Saltelli, A., Chan, K., Scott, E.: Sensitivity analysis. Wiley series in probability and statistics. Wiley, Chichester (2000)
5. Santner, T., Williams, B., Notz, W.: The Design and analysis of computer experiments. Springer series in statistics. Springer, New York (2003)
6. Sobol, I.: Sensitivity estimates for non linear mathematical models. Math. Modelling Comput. Exp. **1**, 407–414 (1993)