Metamodelling with Dependent Inputs. A Comparison of Two Approaches

Miguel Munoz Zuniga, Sergei Kucherenko, Nilay Shah

CPSE, Imperial College London, London, SW7 2AZ, UK

August 29 - September 2, 2011

In the context of uncertainty and sensitivity analysis the behavior of a realistic system can be represented by a model with a random input vector and a random output vector. The computation of the model for a given input value may require from milliseconds to several days of CPU time for a single run. The dimension of the uncertain input vector may range from a few to several hundred variables. In the extreme cases of computationally expensive models the metamodelling technique which maps inputs and outputs is a very useful and practical way of making computations tractable. Commonly used metamodelling techniques include the spline, generalized linear model, partial least squares model, neural network, support vector machines, kriging, RS(QRS)-HDMR and polynomial chaos methods. The majority of known methods deal with models with independent input variables. However, in practical applications input variables are often dependent. The objective of this work is to compare two different metamodelling techniques for models with dependent input variables. Both techniques are based on polynomial chaos expansions. In the polynomial chaos method the model function is decomposed using suitable tensored polynomials. The choice of a polynomial basis depends on the distributions of the input vector. The first technique consist of transforming the dependent input vector into a Gaussian independent random vector and then applying decomposition of the model using the tensored Hermite polynomial basis. The second approach is based on the direct decomposition of the model function, without making the random input vector independent, into a basis which is based on the marginal distributions of the input components and their joint distribution. For both methods, the use of the copula formalism and the Monte Carlo approximations are discussed. We present numerical and analytical results and discuss the efficiency of both approaches on representational benchmark examples.