Conference Paper

Covariance-based sensitivity indices based on polynomial chaos functional decomposition

Author(s):
Caniou, Y.; Sudret, Bruno

Publication Date:
2013

Permanent Link:
https://doi.org/10.3929/ethz-a-010062260 →

Covariance-based sensitivity indices based on polynomial chaos functional decomposition

Yann, Caniou
PHIMECA Engineering S.A., Cournon d'Auvergne, FRANCE

Bruno, Sudret
ETH, Zürich, Institute of Structural Mechanics, Chair of Risk, Safety & Uncertainty Quantification

The art of robust engineering requires to take the random nature of design parameters into account in order to predict the dispersion of the performance of a structure. When dealing with reducing this dispersion, one has to identify the parameters the variability of the performance is the most sensitive to. Global sensitivity analysis (GSA) is a statistical field that aims at identifying and prioritizing the design parameters that contribute the most to the dispersion of the response of a model. This quantity is in most cases described by the statistical variance of the model response. The so-called ANOVA (ANalysis Of VAriance) technique ranks the parameter according the share of the model response variance they are responsible for.

Let us consider a performance Y described by a physical model $\mathcal{M}(\boldsymbol{X})$ where $\boldsymbol{X}$ is $n-$dimensional random vector with independent components. Such an apportionment of the total variance can be processed thanks to a functional decomposition of the model $\mathcal{M}$ [1] reading:

$$\begin{aligned}\mathcal{M}(\boldsymbol{X}) &= \mathcal{M}_0 + \sum_{i=1}^{n} \mathcal{M}_i(X_i) + \sum_{1 \leqslant i < j \leqslant n} \mathcal{M}_{ij}(X_i, X_j) + \cdots + \mathcal{M}_{1\ldots n}(X_1, \ldots, X_n) \\ &= \mathcal{M}_0 + \sum_{\boldsymbol{u} \subseteq \{1,\ldots,n\}} \mathcal{M}_{\boldsymbol{u}}(\boldsymbol{X}_{\boldsymbol{u}})\end{aligned} \tag{1}$$

where $\mathcal{M}_0$ is a constant and where the components have zero mean and are mutually orthogonal. This decomposition also holds when dealing with the variance of Y:

$$\mathbb{V}[Y] = \sum_{i=1}^{n} \mathbb{V}[\mathcal{M}_i(X_i)] + \sum_{1 \leqslant i < j \leqslant n} \mathbb{V}[\mathcal{M}_{ij}(X_i, X_j)] + \cdots + \mathbb{V}[\mathcal{M}_{1\ldots n}(X_1, \ldots, X_n)] \tag{2}$$

The so-called Sobol' index [2] of a variable X_i is defined by the ratio between the variance of the component that only depends on X_i and the total variance of Y, namely:

$$S_i = \frac{\mathbb{V}[\mathcal{M}_i(X_i)]}{\mathbb{V}[Y]} \tag{3}$$

The index S_i represents the share of the variance of Y that is due to both the physical role of X_i in $\mathcal{M}$ and its random nature. An index S_i close to 1 indicates a strong contribution of X_i to the dispersion of Y whereas an index close to 0 denotes a weak incidence.

Computing the ANOVA sensitivity indices requires to identify the different component of the functional decomposition. This task can be achieved by a projection method but the corresponding computing cost is substantial. On top of that, if the response of the model $\mathcal{M}$ is expensive to evaluate (if $\mathcal{M}$ is a FEM code for instance), then performing a sensitivity analysis is almost inconceivable. In order to circumvent this limitation, one may substitute the physical model by a surrogate model $\widehat{\mathcal{M}}$, namely a analytical representation built from a reasonable-sized design of experiment $\mathcal{D} = \{\boldsymbol{\mathcal{X}}, \mathcal{Y} = \mathcal{M}(\boldsymbol{\mathcal{X}})\}$ that is much cheaper to evaluate than $\mathcal{M}$. One adequate method is referred to as polynomial chaos expansion [3]. The principle is to expand the model response on a suitable polynomial basis, namely:

$$Y \approx \widehat{\mathcal{M}}(\boldsymbol{X}) = \sum_{j=0}^{P-1} a_j \Psi_j(\boldsymbol{X}) \tag{4}$$

In practice, the basis $\mathcal{B} = \{\Psi_j,\, j = 0, \ldots, P-1\}$ is truncated to a finite number P of terms, *e.g.* according to the maximal total degree of the retained polynomials or according to a sparser scheme. Then, defining a substitution model for $\mathcal{M}$ consists in evaluating the coefficients a_j of the development, using a regression method for instance.

GSA techniques for models with *independent* input parameters are well-established and computationally efficient when coupled with surrogate models. When the input parameters are no longer independent, the functional decomposition in (1) does not hold since the components of the decomposition are no longer orthogonal. A generalization of the ANOVA for models with correlated input has been introduced in [4]. The principle of the ANCOVA (ANalysis of COVAriance) is to express the variance of Y as its covariance with the functional decomposition of $\mathcal{M}$, namely:

$$\begin{aligned} \mathbb{V}[Y] &= \mathbb{C}[Y, \mathcal{M}(\boldsymbol{X})] \\ &= \mathbb{C}\left[Y, \sum_{\boldsymbol{u} \subseteq \{1,\ldots,n\}} \mathcal{M}_{\boldsymbol{u}}(\boldsymbol{X}_{\boldsymbol{u}})\right] \\ &= \sum_{\boldsymbol{u} \subseteq \{1,\ldots,n\}} \Big[\mathbb{V}[\mathcal{M}_{\boldsymbol{u}}(\boldsymbol{X}_{\boldsymbol{u}})] + \mathbb{C}[\mathcal{M}_{\boldsymbol{u}}(\boldsymbol{X}_{\boldsymbol{u}}), Y - \mathcal{M}_{\boldsymbol{u}}(\boldsymbol{X}_{\boldsymbol{u}})]\Big] \end{aligned} \tag{5}$$

The following triplet of indices $(S_{\boldsymbol{u}}, S_{\boldsymbol{u}}^U, S_{\boldsymbol{u}}^C)$ can be derived from (5):

$$S_{\boldsymbol{u}} = \frac{\mathbb{C}[\mathcal{M}_{\boldsymbol{u}}(\boldsymbol{X}_{\boldsymbol{u}}), Y]}{\mathbb{V}[Y]}, \quad S_{\boldsymbol{u}}^U = \frac{\mathbb{V}[\mathcal{M}_{\boldsymbol{u}}(\boldsymbol{X}_{\boldsymbol{u}})]}{\mathbb{V}[Y]}, \quad S_{\boldsymbol{u}}^C = S_{\boldsymbol{u}} - S_{\boldsymbol{u}}^U \tag{6}$$

The index $S_{\boldsymbol{u}}^U$ represents the uncorrelated contribution of $\boldsymbol{X}_{\boldsymbol{u}}$ to the variance of Y, that is the contribution that would be left if the variables where independent. On the contrary, the index $S_{\boldsymbol{u}}^C$ represents the contribution of the correlation of $\boldsymbol{X}_{\boldsymbol{u}}$ with the other parameters. The global contribution index $S_{\boldsymbol{u}} = S_{\boldsymbol{u}}^U + S_{\boldsymbol{u}}^C$ is the sum of the two contributions.

The issue of the functional decomposition is solved here by using the one provided by the polynomial chaos expansion in (4). Since the expansion of the correlated parameters is not expressed in the physical space because of the isoprobabilistic transformation, the approach proposed in [5] is to build the expansion with the joint distribution of the input random vector $\boldsymbol{X}$ featuring an independent copula to preserve the link between the physical and standard variables and to evaluate the variances and covariances by simulating realizations of $\boldsymbol{X}$ with its true dependence structure.

The ANCOVA technique coupled with polynomial chaos expansion is first applied on analytical test functions found in the literature in order to exhibit how the uncorrelated an correlated parts behave when the correlation between the input parameters varies. It is then carried out on a simple mechanical application.

References:

[1] W. Hoeffding, "A class of statistics with asymptotically normal distributions", *Annals of Statistics*, **19**:293–325, 1948.

[2] I.M. Sobol', "Sensitivity estimates for nonlinear Mathematical models", *Matem. Modelirovanie*, **2**(1):112–118, 1990 (in Russian), *MMCE*, **1**(4):407–414, 1993 (in English).

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

[4] G. Li and H. Rabitz, "Global Sensitivity Analysis for Systems with Independent and/or Correlated Inputs", *J. Phys. Chem.*, **114**:6022–6032, 2010.

[5] Y. Caniou, "Global sensitivity analysis for nested and multiscale modelling", Université Blaise Pascal – Clermont II, 2012.

[Yann Caniou; PHIMECA Engineering S.A., 34 rue de Sarlive, F-63800 Cournon d'Auvergne]
[`caniou@phimeca.com` – `www.phimeca.com`]