Introduction to Uncertainty Quantification
Module 4: Global Sensitivity Analysis

1 Introduction

Consider a model of a system given generally by $Y = \mathcal{M}(\boldsymbol{X})$ where $\boldsymbol{X}$ is a random vector of inputs to the model. For our purposes, we will consider $\boldsymbol{X}$ to have independent components such that its distribution is given by the product of its marginal distribution as $f_{\boldsymbol{X}}(\boldsymbol{x} = \prod_{i=1}^{d} f_i(x_i)$. As defined by Saltelli et al. [1], *sensitivity analysis* is "the study of how uncertainty in the output of a model (numerical or otherwise) can be apportioned to different sources of uncertainty in the model input."

Paragraph on why we perform GSA.

The broad field of sensitivity analysis can be divided into two distinct, but equally important tasks referred to as *local sensitivity analysis* (LSA) and *global sensitivity analysis* (GSA). It is important to make this distinction because LSA and GSA serve different purposes and innately provide different information about the model and its dependence on various inputs. In the following, we will briefly introduce both approaches.

1.1 Local Sensitivity Analysis

Local Sensitivity Analysis (LSA) is concerned with derivatives. In fact, one simple definition of a local sensitivity index is simple the partial derivative of the model with respect to each given variable, i.e.

$$S^l_{X_i} = \frac{\partial \mathcal{M}}{\partial X_i}. \tag{1}$$

However, unless the model is linear in all components, this local sensitivity index depends on the location at which the derivative is computed. That is

$$S^l_{X_i}(\boldsymbol{x}^*) = \left.\frac{\partial \mathcal{M}}{\partial X_i}\right|_{\boldsymbol{x}=\boldsymbol{x}^*} \tag{2}$$

Therefore, they provide information only in the vicinity of $\boldsymbol{x} = \boldsymbol{x}^*$. This can be very useful for certain tasks that can exploit local linearity, such as model order reduction methods. Moreover, computationally efficient implementations mean they can often be computed inexpensively. However, they provide little information about the influence of uncertainty on the model output *unless* either the model is linear (or near linear) or uncertainty is very small and therefore it has influence only over a small region where the model can be approximated as linear.

1.2 Global Sensitivity Analysis

We are instead concerned with the influence of uncertainty in each of the model parameters on the overall response of the system. That is, we want to assess how important each uncertain variable is in terms of

its impact on the output, in general, not just at a single location. This approach is referred to as Global sensitivity analysis (GSA).

The major question in GSA is how to assess the "overall" performance of the model. In fact, one can imagine many representative statistics that quantify the overall performance of the model in some meaningful way. In principle, GSA can be used in the context of any *well-defined* measure of model performance such that we aim to understand the extent of the influence (or importance) that each input has on this statistical measure. It is therefore essential that we clearly define *importance*. What does it mean for a particular input to be important? The definition of the performance measure and the means of assessing importance define the problem setting. Saltelli et al. [1] describe several common problem settings. Most of these settings employs the variance of the model response as the statistical measure of interest, but assess the importance of each factor differently.

- **Factor Prioritization**: In this setting, we aim to idenitify a factor (or factors) that, when fixed to its true (or nominal) value, leads to the greatest reduction in the variance of the output. This means that we seek the input variable (factor) that has the largest contribution to the variance of the output. This may be important, for example, in the setting of epistemic uncertainty where we aim to collect additional data that will reduce uncertainty in the output. If we can identify the factor that contributes most to the output variance, then we can prioritize this factor in efforts to reduce uncertainty – e.g. peforming experiments or measurements to reduce uncertainty in this factor.

- **Factor Fixing**: Here, we aim to identify factors (inputs) that have little or no influence on the variance of the output. That is, no matter what value these inputs take, the variance of the output does not change, or its change is negligibly small. This is important, for example, in dimension reduction. Consider that we have a set of d inputs to the model but we cannot afford to propagation uncertainty through the model because we need a huge number of simulations due to the *curse of dimensionality*. If we can identify a set of d_u factors that have little or no influence, the we can fix them and propagate uncertainty in only the $d_{\text{eff}} = d - d_u$ factors, which may be significantly more efficient computationally.

- **Variance Cutting**: In this setting, we aim to reduce the variance of the output to below a given threshold. This may be important, for example, when it is important to assign a given level of confidence to a quantified uncertainty. When assess the statistical response of the system (i.e. mean, variance, probability of failure, etc.) we may require that the statistic have a certain level of confidence. We can assign confidence in the estimator by reducing its variance below a given threshold. In this way, we can perform uncertainty quantification using the smallest possible set of factors that will ensure we achieve our required confidence.

- **Factor Mapping**: Here, we are interested in understanding which values of the input lead to outputs in a specific range. For example, in reliability analysis it is common to identify which factors and their ranges cause model outputs that exceed a specified threshold that corresponds to system failure.

These are all common settings in which GSA is applied, although many others certainly exist. One of the common attributes of most (but not all) settings is that they apply the variance of the output as the measure of model performance. For this reason, we will focus primarily on *variance-based* GSA. This is the topic of the next section.

2 Variance-based sensitivity analysis and Sobol' indices

The variance is a natural statistic for GSA because it provides, in a simple scalar value, a measure of the magnitude of variation of a given quantity – in this case the output of the model. To understand

the influence of any single factor on the variance of the model output, it is therefore helpful to define the *conditional variance*.

Consider again our model, defined $Y = \mathcal{M}(\boldsymbol{X})$ where $\boldsymbol{X} = \{X_1, X_2, \ldots, X_d\}^T$ is a d-dimensional random vector. We will denote the variance of the model output by $V(Y)$. Consider that we fix the input X_i at a particular value x_i^*. The variance of the model under this condition is given by $V_{\boldsymbol{X}_{\sim i}}(Y|X_i = x^*)$, where the subscript $\sim i$ denotes that the variance is taken over all input *except* i. This is referred to as the *conditional variance*.

It may be tempting to use $V_{\boldsymbol{X}_{\sim i}}(Y|X_i = x^*)$ directly as a measure of importance of each individual input such that inputs with high conditional variance are more important. This is problematic for two reasons (although this is sometimes done for certain *one factor at-a-time* approaches where e.g. it is common to set X_i equal to its mean value, $X_i = \mu_{X_i}$). First, the conditional variance depends on the specific value of x^* that is chosen and can vary greatly if x^* is changed. Second, for certain fixed values x_i^* in some models, it may occur that $V_{\boldsymbol{X}_{\sim i}}(Y|X_i = x^*) > V(Y)$. That is, the conditional variance may be large than the total variance of the output. This is clearly undesirable.

A more prudent approach is to take the expectation of the conditional variance over the variable X_i. In this way, we average the conditional variances over all values of X_i and are therefore not dependent on the specific selection of a point x_i^*. This expectation is expressed by $E_{X_i}[V_{\boldsymbol{X}_{\sim i}}(Y|X_i)]$. In fact, the *Law of Total Variance* given by

$$E_{X_i}[V_{\boldsymbol{X}_{\sim i}}(Y|X_i)] + V_{X_i}(E_{\boldsymbol{X}_{\sim i}}[Y|X_i]) = V(Y) \tag{3}$$

states that this expectation is strictly less than or equal to the total variance $V(Y)$. Analysis of Eq. (3) reveals that a small value of $E_{X_i}[V_{\boldsymbol{X}_{\sim i}}(Y|X_i)]$ implies that fixing X_i removes substantial variance and therefore X_i is an important factor. This term of often referred to as a *residual* because it is the variance that remains, on average, after fixing X_i. This further suggests that a large value of $V_{X_i}(E_{\boldsymbol{X}_{\sim i}}[Y|X_i])$, referred to as the *first-order effect*, or *main effect* of X_i on $V(Y)$, implies that factor X_i is important. From the first-order effect, we then define the *first-order senstivity index* as

$$S_i = \frac{V_{X_i}(E_{\boldsymbol{X}_{\sim i}}[Y|X_i])}{V(Y)}. \tag{4}$$

S_i is a scalar value that always lies in the range $[0, 1]$ such that a higher value implies more importance.

2.1 High-Dimensional Model Representation (HDMR) & Sobol' Indices

Consider that our model can be expressed as a square integrable function $f(\boldsymbol{X})$ over the unit hypercube, where $\boldsymbol{X} \in [0, 1]^d$. Equivalently, we consider that the model $Y = \mathcal{M}(\boldsymbol{X})$ operates on uniform random variables and that Y has finite variance. The Russian mathematician I.M. Sobol' considered the following expansion of this function into terms of increasing dimension

$$f(\boldsymbol{X}) = f_0 + \sum_i f_i(_i) + \sum_i \sum_{j>i} f_{ij}(X_i, X_j) + \cdots + f_{12\ldots d} \tag{5}$$

called the *high-dimensional model representation* (HDMR). Note that the HDMR is not an infinite series, like Fourier or Taylor series. It has a finite number of terms (2^d), and decomposes the function into a hierarchy of terms of increasing complexity starting with a constant term f_0, d first-order terms $f_i(X_i)$, $\binom{d}{2}$ second-order terms $f_{ij}(X_i, X_j)$, and so on.

Moreover, the HDMR is not unique. There are potentially many (even infinitely many) possible HDMRs for a given function. Sobol' proved that one way to construct the HDMR is through conditional expectations

of the model output, Y. That is:

$$\begin{aligned} f_0 &= E[Y] \\ f_i(X_i) &= E[Y|X_i] - f_0 \\ f_{ij}(X_i, X_j) &= E[Y|X_i, X_j] - f_i(X_i) - f_j(X_j) - f_0 \\ &\vdots \end{aligned} \tag{6}$$

This construction leads to a decomposition in which all term are orthogonal, and therefore uncorrelated. As a result, the variance of the output $V(Y)$ can be computed as the sum of the variance of the individual terms. That is

$$V(Y) = \sum_i V_i + + \sum_i \sum_{j>i} V_{ij} + \cdots + V_{12\ldots d} \tag{7}$$

where we see that

$$V_i = \mathrm{Var}(f_i(X_i)) = V(E[Y|X_i]) \tag{8}$$

is the *first-order effect* defined above. Note that the subscripts X_i and $\boldsymbol{X}_{\sim i}$ have been dropped and are implied by the conditioning on X_i. We likewise define the *second-order effects* as

$$V_{ij} = \mathrm{Var}(f_{ij}(X_i, X_j)) = V(E[Y|X_i, X_j]) - V(E[Y|X_i]) - V(E[Y|X_j]) \tag{9}$$

This term expresses the influece that the pair of inputs (X_i, X_j) have on Y with the first-order effects of X_i and X_j removed. Similar terms can be written to express *higher-order effects*.

Dividing Eq. (7), known as the ANOVA-HDMR, by the total variance $V(Y)$ yields

$$\sum_i S_i + + \sum_i \sum_{j>i} S_{ij} + \cdots + S_{12\ldots d} = 1 \tag{10}$$

where S_i are the first-order sensitivity indices given in Eq. (4) – also referred to as Sobol' indices. It follows that the *second-order sensitivity indices* are given by:

$$S_{ij} = \frac{V(E[Y|X_i, X_j]) - V(E[Y|X_i]) - V(E[Y|X_j])}{V(Y)} = \frac{V(E[Y|X_i, X_j])}{V(Y)} - S_i - S_j \tag{11}$$

Additional higher-order indices can be computed in the same manner.

Let us now consider that we now condition on all factors *except* X_i, i.e.‘ $\boldsymbol{X}_{\sim i}$. We can express the conditional variance as $V(E[Y|\boldsymbol{X}_{\sim i}])$ and the corresponding sensitivity index as

$$S_{\boldsymbol{X}_{\sim i}} = \frac{V(E[Y|\boldsymbol{X}_{\sim i}])}{V(Y)} \tag{12}$$

which expresses the contributions of all terms of any order that do not include X_i. Given the summation in Eq. (10), we can express its complement as

$$S_{Ti} = 1 - \frac{V(E[Y|\boldsymbol{X}_{\sim i}])}{V(Y)} \tag{13}$$

where S_{Ti} is referred to as the *total effect* of factor X_i and expresses the relative contribution of all terms of any order that *include* factor X_i. For example, if we consider a system with three factors X_1, X_2, X_3 then the total effect indices are given by

$$\begin{aligned} S_{T1} &= S_1 + S_{12} + S_{13} + S_{123} \\ S_{T2} &= S_2 + S_{12} + S_{23} + S_{123} \\ S_{T3} &= S_3 + S_{23} + S_{13} + S_{123} \end{aligned} \tag{14}$$

Recalling the *law of total variance* in Eq. (3) we can derive the total effect in another way. Instead of conditioning on X_i in the law of total variance, we can condition on all factors except X_i, $\boldsymbol{X}_{\sim i}$, as follows

$$E[V(Y|\boldsymbol{X}_{\sim i})] + V(E[Y|\boldsymbol{X}_{\sim i})]) = V(Y) \tag{15}$$

where $E[V(Y|\boldsymbol{X}_{\sim i})] = V(Y) - V(E[Y|\boldsymbol{X}_{\sim i})])$ is the residual variance of Y that would remain, on average, if we specify all terms except X_i. This is exactly the *total effect* of factor X_i. Dividing by the total variance yields:

$$S_{Ti} = \frac{E[V(Y|\boldsymbol{X}_{\sim i})]}{V(Y)} = 1 - \frac{V(E[Y|\boldsymbol{X}_{\sim i}])}{V(Y)} \tag{16}$$

2.2 Estimating Sobol' indices by Monte Carlo simulation

A naive approach to estimating a given Sobol' index, S_i, is the following

1. Fix X_i
2. Conduct a Monte Carlo simulation using N samples
3. Estimate the conditional expectation $E[Y|X_i]$
4. Repeat steps 1–3 N_{X_i} times
5. Estimate the variance of the expectations

Of course, this procedure is computationlly cumbersome becaause it requires $N \times N_{X_i}$ simulations for *each* Sobol' index.

Sobol' [2] originally proposed a more efficient method for the estimation of first-order and total order sensitivity indices. This method was then extended by Homma and Saltelii [3] and then further by Saltelli [4]. The method of Saltelli [4] is presented here.

Consider the d-dimensional input random vector $\boldsymbol{X} = \{X_1, X_2, \ldots, X_d\}$. First, we generate two matrices of N samples, denoted A and B and given by

$$A = \begin{bmatrix} \boldsymbol{x}_a^{(1)} \\ \boldsymbol{x}_a^{(2)} \\ \vdots \\ \boldsymbol{x}_a^{(N)} \end{bmatrix} \qquad B = \begin{bmatrix} \boldsymbol{x}_b^{(1)} \\ \boldsymbol{x}_b^{(2)} \\ \vdots \\ \boldsymbol{x}_b^{(N)} \end{bmatrix} \tag{17}$$

where again each sample corresponds to the vector $\boldsymbol{x}^{(i)} = \{x_1^{(i)}, x_2^{(i)}, \ldots, x_d^{(i)}\}$. Next, we create a third matrix of samples, C_i, by replacing the i^{th} column of matrix B with the i^{th} column of A. That is, replacing the i^{th} component of every sample in B with the corresponding component sampel from A as

$$C_i = \begin{bmatrix} x_{b,1}^{(1)} & x_{b,2}^{(1)} & \ldots & x_{a,i}^{(1)} & \ldots & x_{b,d}^{(1)} \\ x_{b,1}^{(2)} & x_{b,2}^{(2)} & \ldots & x_{a,i}^{(2)} & \ldots & x_{b,d}^{(2)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{b,1}^{(N)} & x_{b,2}^{(N)} & \ldots & x_{a,i}^{(N)} & \ldots & x_{b,d}^{(N)} \end{bmatrix} \tag{18}$$

We then run the model on all samples in the matrices A, B, and $C_i, i = 1, \ldots, d$ to obtain all first-order and total-order indices as follows. The first-order indices are estimated as

$$S_i = \frac{V(E[Y|X_i])}{V(Y)} = \frac{y_A \cdot y_{C_i} - f_0^2}{y_A \cdot y_A - f_0^2} = \frac{\frac{1}{N}\sum_{j=1}^{N} y_A^{(j)} y_{C_i}^{(j)} - f_0^2}{\frac{1}{N}\sum_{j=1}^{N} (y_A^{(j)})^2 - f_0^2} \tag{19}$$

where $y_A = f(A)$, $y_B = f(B)$, and $y_{C_i} = f(C_i)$ are the vectors of model responses for sample sets A, B, and C_i, respectively, and

$$f_0 = \left(\frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} y_A^{(j)} \right)^2 \tag{20}$$

is the Monte Carlo estimator of the mean value squared using samples A.

Similarly, the total-effect sensitivities can be estimated by

$$S_{Ti} = 1 - \frac{V(E[Y|\boldsymbol{X}_{\sim i}])}{V(Y)} = 1 - \frac{y_B \cdot y_{C_i} - f_0^2}{y_A \cdot y_A - f_0^2} = \frac{\frac{1}{N}\sum_{j=1}^{N} y_B^{(j)} y_{C_i}^{(j)} - f_0^2}{\frac{1}{N}\sum_{j=1}^{N} (y_A^{(j)})^2 - f_0^2} \tag{21}$$

In terms of computational expense, we see here that computing each main effect requires $N + N$ simulations for a total cost of $N(d+1)$ and the total effect sensitivities require an additional N simulations. Therefore, the total cost for all main effect and total effect sensitivities is $N(d+2)$

2.3 Estimating Sobol' indices with surrogate models

Another common practice for estimating Sobol' indices is to construct a surrogate model from a relatively small number of simulations (See Module 3.4) and exploit the surrogate model for global sensitivity analysis. This is commonly done using both the PCE surrogates and GP surrogates surrogate discussed in Module 3.4. Here we desribe the process for estimating Sobol indices in both cases.

Polynomial Chaos Expansions

Consider that we have constructed a PCE surrogate model as described in Module 3.4 and given by:

$$Y(\boldsymbol{x}) \approx Y^{\mathrm{PC}}(\boldsymbol{x}) = \sum_{\boldsymbol{\alpha} \in \mathcal{A}} \beta_{\boldsymbol{\alpha}} \Psi_{\boldsymbol{\alpha}}(\boldsymbol{x}) \tag{22}$$

Sudret [5], and shortly thereafter Crestaux et al. [6], showed that the Sobol' indices can be derived directly from the PCE as follows. The first step is to construct the Sobol' HDMR (Eqs. (5) and (6)) from the PCE expansion. We first define the set $\mathcal{I}_{i_1,\ldots,i_s}$ as the set that indexes the polynomials depending only on x_i. Formally this can be defined by

$$\mathcal{I}_{i_1,\ldots,i_s} = \left\{ \boldsymbol{\alpha} : \begin{array}{llll} \alpha_k > 0 & \forall k = 1,\ldots,d, & k \in (i_1,\ldots,i_s) \\ \alpha_k = 0 & \forall k = 1,\ldots,d, & k \notin (i_1,\ldots,i_s) \end{array} \right\}, \tag{23}$$

which is the set of $\boldsymbol{\alpha}$ tuples such that only the indices $(i_1,\ldots,i_s)$ are non-zero. Using this index set, we can now gather the PCE terms according to the Sobol' HDMR as follows

$$\begin{aligned} Y^{\mathrm{PC}}(\boldsymbol{x}) = \beta_0 + \sum_{i=1}^{d} \sum_{\boldsymbol{\alpha} \in \mathcal{I}_i} \beta_{\boldsymbol{\alpha}} \Psi_{\boldsymbol{\alpha}}(x_i) + \sum_{1 \le i_1 \le i_2 \le d} \sum_{\boldsymbol{\alpha} \in \mathcal{I}_{i_1 i_2}} \beta_{\boldsymbol{\alpha}} \Psi_{\boldsymbol{\alpha}}(x_{i_1}, x_{i_2}) + \ldots \\ + \sum_{1 \le i_1 \le \cdots \le i_s \le d} \sum_{\boldsymbol{\alpha} \in \mathcal{I}_{i_1,\ldots,i_s}} \beta_{\boldsymbol{\alpha}} \Psi_{\boldsymbol{\alpha}}(x_{i_1},\ldots,x_{i_s}) + \cdots + \sum_{\boldsymbol{\alpha} \in \mathcal{I}_{1,2,\ldots,d}} \beta_{\boldsymbol{\alpha}} \Psi_{\boldsymbol{\alpha}}(x_1,\ldots,x_d) \end{aligned} \tag{24}$$

Gaussian Process Regression

3 Other methods for GSA

3.1 Fourier Amplitude Sensitivity Test (FAST)

3.2 Morris Method

3.3 Chaterjee sensitivity indices

3.4 Cramér-von Mises sensitivity indices

3.5 Distribution-based GSA

References

[1] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, S. Tarantola, Global sensitivity analysis: the primer, John Wiley & Sons, 2008.

[2] I. Soboĺ, Sensitivity estimates for nonlinear mathematical models, Math. Model. Comput. Exp. 1 (1993) 407.

[3] T. Homma, A. Saltelli, Importance measures in global sensitivity analysis of nonlinear models, Reliability Engineering & System Safety 52 (1) (1996) 1–17.

[4] A. Saltelli, Making best use of model evaluations to compute sensitivity indices, Computer physics communications 145 (2) (2002) 280–297.

[5] B. Sudret, Global sensitivity analysis using polynomial chaos expansions, Reliability engineering & system safety 93 (7) (2008) 964–979.

[6] T. Crestaux, O. Le Maıtre, J.-M. Martinez, Polynomial chaos expansion for sensitivity analysis, Reliability Engineering & System Safety 94 (7) (2009) 1161–1172.