Technische Universität München
Lehrstuhl für Informatik V
Dr. Tobias Neckel
M. Sc. Friedrich Menhorn

SS 2018

June 27, 2018

Algorithms for Uncertainty Quantification

Tutorial 9: Sobol' indices for global sensitivity analysis

In this worksheet, we focus on the variance-based global sensitivity analysis. In particular, we focus on the computation of the Sobol' indices for global sensitivity analysis.

Variance-based global sensitivity analysis

Variance-based sensitivity analysis provides a way to numerically quantify the contribution (take individually or in combination thereof) of uncertain inputs into output uncertainty; as the name suggests, the measure for output uncertainty is the output variance. The uncertain inputs' contribution to the total resulted variance can be measured via the so called (total) Sobol' indices for global sensitivity analysis.
As discussed in the lecture the sensitivity analysis is based on the ANOVA decomposition:

$$f(t, \boldsymbol{\omega}) = f_0(t) + \sum_{i=1}^{d} f_i(t, \omega_i) + \sum_{1 \le i < j \le d} f_{ij}(t, \omega_i, \omega_j) + \tag{1}$$

$$\sum_{1 \le i_1 < \ldots < i_s \le d} f_{i_1, \ldots, i_s}(t, \omega_{i_1}, \ldots, \omega_{i_s}) + \ldots + f_{12 \ldots d}(t, \boldsymbol{\omega}) \ , \tag{2}$$

which is often approximated by only up to second order terms

$$f(t, \boldsymbol{\omega}) = f_0(t) + \sum_{i=1}^{d} f_i(t, \omega_i) + \sum_{1 \le i < j \le d} f_{ij}(t, \omega_i, \omega_j). \tag{3}$$

This is, however, not a unique representation and further assumptions have to be imposed. Therefore, we add the following orthogonality assumption:

$$\int_{\Gamma^d} f_{i_1 \dots i_n}(t, \omega_{i_1}, \dots, \omega_{i_n}) f_{i_1 \dots i_m}(t, \omega_{i_1}, \dots, \omega_{i_m}) d\boldsymbol{\omega} = 0, \tag{4}$$

when at least one index differs in the sets $\{i_1, ..., i_n\}$ and $\{i_1, ..., i_m\}$.
Denoting

$$D_{i_1 \dots i_s}(t) = \int_{\Gamma^s} f^2_{i_1 \dots i_s}(f, \omega_{i_1} \dots \omega_{i_s}) d\omega_{i_1} \dots d\omega_{i_s} \tag{5}$$

we are now interested in the contributions of the different terms. Finally, the local and total Sobol indices are defined as follows:

- local/partial Sobol' indices: measure individual contributions OR "interactions" between inputs

$$S_{i_1 \dots i_s}(t) = \frac{D_{i_1 \dots i_s}}{\mathrm{Var}[f(t, \boldsymbol{\omega})]}$$

- total Sobol' indices: measure individual contributions AND "interactions" between inputs

$$S_i^T(t) = \sum_{i \in \{i_1, \dots, i_s\}} S_{i_1, \dots, i_s}(t)$$

Optional: Assignment 1

Given the ANOVA decomposition in eq. (1) and using the orthogonality assumption from eq. (4), derive the variance of the decomposition

$$\mathrm{Var}[f(t, \boldsymbol{\omega})]$$

using the notation in eq. (5).
For brevity we ignore the time dependency for now, i.e. $f(t, \boldsymbol{\omega}) =: f(\boldsymbol{\omega})$. Also, the indices directly tell the stochastic dimensions of ω that are involved, i.e. $f_{i_1, \dots, i_s}(\omega_{i_1}, \dots, \omega_{i_s}) =: f_{i_1, \dots, i_s}$. Just to show what that means we write out the equations in the first step and

substitute the abbreviations in the second one.

$$\mathrm{Var}[f(t,\boldsymbol{\omega})] =$$

$$\mathrm{Var}[f_0(t) + \sum_{i=1}^{d} f_i(t,\omega_i) + \sum_{1\le i_1<i_2\le d} f_{i_1 i_2}(t,\omega_i,\omega_j) \ + \sum_{1\le i_1<...<i_s\le d} f_{i_1,...,i_s}(t,\omega_{i_1},...,\omega_{i_s}) + ... + f_{12...d}(t,\boldsymbol{\omega})] =$$

$$Var[f_0 + \sum_{i=1}^{d} f_i + \sum_{1\le i_1<i_2\le d} f_{i_1 i_2} \ + \sum_{1\le i_1<...<i_s\le d} f_{i_1,...,i_s} + \ldots + f_{12...d}] =$$

$$\mathbb{E}\Bigg[\bigg(f_0 + \sum_{i=1}^{d} f_i + \sum_{1\le i_1<i_2\le d} f_{i_1 i_2} \ + \sum_{1\le i_1<...<i_s\le d} f_{i_1,...,i_s} + \ldots + f_{12...d} -$$

$$\mathbb{E}[f_0 + \sum_{i=1}^{d} f_i + \sum_{1\le i_1<i_2\le d} f_{i_1 i_2} \ + \sum_{1\le i_1<...<i_s\le d} f_{i_1,...,i_s} + ... + f_{12...d}]\bigg)^2\Bigg] =$$

$$\mathbb{E}\Bigg[\bigg(f_0 + \sum_{i=1}^{d} f_i + \sum_{1\le i_1<i_2\le d} f_{i_1 i_2} \ + \sum_{1\le i_1<...<i_s\le d} f_{i_1,...,i_s} + \ldots + f_{12...d} -$$

$$[\int_\Omega f_0 + \sum_{i=1}^{d} f_i + \sum_{1\le i_1<i_2\le d} f_{i_1 i_2} \ + \sum_{1\le i_1<...<i_s\le d} f_{i_1,...,i_s} + ... + f_{12...d} d\mu(\omega)]\bigg)^2\Bigg] =$$

$$\mathbb{E}\Bigg[\bigg(f_0 + \sum_{i=1}^{d} f_i + \sum_{1\le i_1<...<i_s\le d} f_{i_1,...,i_s} + \ldots + f_{12...d} -$$

$$[\int_\Omega f_0 d\mu(\omega) + \int_\Omega \sum_{i=1}^{d} f_i d\mu(\omega) + \int_\Omega \sum_{1\le i_1<...<i_s\le d} f_{i_1,...,i_s} d\mu(\omega) + ... + \int_\Omega f_{12...d} d\mu(\omega)]\bigg)^2\Bigg] =$$

$$\mathbb{E}\Bigg[\bigg(f_0 + \sum_{i=1}^{d} f_i + \sum_{1\le i_1<...<i_s\le d} f_{i_1,...,i_s} + \ldots + f_{12...d} -$$

$$\int_\Omega f_0 d\mu(\omega) - \int_\Omega \sum_{i=1}^{d} f_i d\mu(\omega) - \int_\Omega \sum_{1\le i_1<...<i_s\le d} f_{i_1,...,i_s} d\mu(\omega) - ... - \int_\Omega f_{12...d} d\mu(\omega)]\bigg)^2\Bigg] =$$

$$\mathbb{E}[f_0 * f_0 + f_0 * \sum_{i=1}^{d} f_i + f_0 * \sum_{1\le i_1<\ldots<i_s\le d} f_{i_1,\ldots,i_s} + \ldots + f_0 f_{12\ldots d} -$$

$$f_0 * \int_\Omega f_0 d\mu(\omega) - f_0 * \int_\Omega \sum_{i=1}^{d} f_i d\mu(\omega) - f_0 * \int_\Omega \sum_{1\le i_1<\ldots<i_s\le d} f_{i_1,\ldots,i_s} d\mu(\omega) - \ldots - f_0 * \int_\Omega f_{12\ldots d} d\mu(\omega) +$$

$$\sum_{i=1}^{d} f_i * f_0 + \sum_{i=1}^{d} f_i * \sum_{i=1}^{d} f_i + \sum_{i=1}^{d} f_i * \sum_{1\le i_1<\ldots<i_s\le d} f_{i_1,\ldots,i_s} + \ldots + \sum_{i=1}^{d} f_i f_{12\ldots d} -$$

$$\sum_{i=1}^{d} f_i * \int_\Omega f_0 d\mu(\omega) - \sum_{i=1}^{d} f_i * \int_\Omega \sum_{i=1}^{d} f_i d\mu(\omega) - \sum_{i=1}^{d} f_i * \int_\Omega \sum_{1\le i_1<\ldots<i_s\le d} f_{i_1,\ldots,i_s} d\mu(\omega) - \ldots - \sum_{i=1}^{d} f_i * \int_\Omega f_{12\ldots d} d\mu(\omega)$$

$$\vdots$$

$$f_{12\ldots d} * f_0 + f_{12\ldots d} * \sum_{i=1}^{d} f_i + f_{12\ldots d} * \sum_{1\le i_1<\ldots<i_s\le d} f_{i_1,\ldots,i_s} + \ldots + f_{12\ldots d} f_{12\ldots d} -$$

$$f_{12\ldots d} * \int_\Omega f_0 d\mu(\omega) - f_{12\ldots d} * \int_\Omega \sum_{i=1}^{d} f_i d\mu(\omega) - f_{12\ldots d} * \int_\Omega \sum_{1\le i_1<\ldots<i_s\le d} f_{i_1,\ldots,i_s} d\mu(\omega) - \ldots - f_{12\ldots d} * \int_\Omega f_{12\ldots d} d\mu(\omega) -$$

$$\vdots$$

$$\int_\Omega f_0 d\mu(\omega) * f_0 - \int_\Omega f_0 d\mu(\omega) * \sum_{i=1}^{d} f_i - \int_\Omega f_0 d\mu(\omega) * \sum_{1\le i_1<\ldots<i_s\le d} f_{i_1,\ldots,i_s} - \ldots - \int_\Omega f_0 d\mu(\omega) f_{12\ldots d} +$$

$$\int_\Omega f_0 d\mu(\omega) * \int_\Omega f_0 d\mu(\omega) + \int_\Omega f_0 d\mu(\omega) * \int_\Omega \sum_{i=1}^{d} f_i d\mu(\omega) + \int_\Omega f_0 d\mu(\omega) * \int_\Omega \sum_{1\le i_1<\ldots<i_s\le d} f_{i_1,\ldots,i_s} d\mu(\omega) + \ldots + \int_\Omega f_0 d\mu(\omega) * \int_\Omega f_{12\ldots d} d\mu(\omega) -$$

$$\vdots$$

$$\int_\Omega f_{12\ldots d} d\mu(\omega) * f_0 - \int_\Omega f_{12\ldots d} d\mu(\omega) * \sum_{i=1}^{d} f_i - \int_\Omega f_{12\ldots d} d\mu(\omega) * \sum_{1\le i_1<\ldots<i_s\le d} f_{i_1,\ldots,i_s} - \ldots - \int_\Omega f_{12\ldots d} d\mu(\omega) f_{12\ldots d} -$$

$$\int_\Omega f_{12\ldots d} d\mu(\omega) * \int_\Omega f_0 d\mu(\omega) + \int_\Omega f_{12\ldots d} d\mu(\omega) * \int_\Omega \sum_{i=1}^{d} f_i d\mu(\omega) + \int_\Omega f_{12\ldots d} d\mu(\omega) * \int_\Omega \sum_{1\le i_1<\ldots<i_s\le d} f_{i_1,\ldots,i_s} d\mu(\omega) + \ldots + \int_\Omega f_{12\ldots d} d\mu(\omega) * \int_\Omega f_{12\ldots d} d\mu(\omega) =$$

$] =$

all terms $\int f_{i_1,\ldots,i_s} \sum f_{i_1,\ldots,i_r} d\mu(\omega) = \sum \int f_{i_1,\ldots,i_s} f_{i_1,\ldots,i_r} d\mu(\omega) = 0$ due to orthogonality.

all terms (apart from 0!) $\int f_{i_1,\ldots,i_s} d\mu(\omega) = 0$ due to constraint in approximation condition.

$f_0 * f_0$ cancel each other out due to independence on ω.

$$= \mathbb{E}[f_0 * \sum_{i=1}^{d} f_i + f_0 * \sum_{1 \le i_1 < ... < i_s \le d} f_{i_1,...,i_s} + ... + f_0 f_{12...d} -$$

$$\sum_{i=1}^{d} f_i * f_0 + \sum_{i=1}^{d} f_i * \sum_{i=1}^{d} f_i + \sum_{i=1}^{d} f_i * \sum_{1 \le i_1 < ... < i_s \le d} f_{i_1,...,i_s} + ... + \sum_{i=1}^{d} f_i f_{12...d} -$$

$$\vdots$$

$$f_{12...d} * f_0 + f_{12...d} * \sum_{i=1}^{d} f_i + f_{12...d} * \sum_{1 \le i_1 < ... < i_s \le d} f_{i_1,...,i_s} + ... + f_{12...d} f_{12...d}$$

$$] =$$

$$\int_\Omega \Big[f_0 * \sum_{i=1}^{d} f_i + f_0 * \sum_{1 \le i_1 < ... < i_s \le d} f_{i_1,...,i_s} + ... + f_0 f_{12...d} -$$

$$\sum_{i=1}^{d} f_i * f_0 + \sum_{i=1}^{d} f_i * \sum_{i=1}^{d} f_i + \sum_{i=1}^{d} f_i * \sum_{1 \le i_1 < ... < i_s \le d} f_{i_1,...,i_s} + ... + \sum_{i=1}^{d} f_i f_{12...d} -$$

$$\vdots$$

$$f_{12...d} * f_0 + f_{12...d} * \sum_{i=1}^{d} f_i + f_{12...d} * \sum_{1 \le i_1 < ... < i_s \le d} f_{i_1,...,i_s} + ... + f_{12...d} f_{12...d} \Big] d\mu(\omega)$$

$=$ seperate integrals and use orthogonality condition again, introduce index j to distinguish

$$\int_\Omega \sum_{i=1}^{d} \sum_{j=1}^{d} f_i f_j d\mu(\omega)$$

$$+ \int_\Omega \sum_{1 \le i_1 < ... < i_s \le d} \sum_{1 \le j_1 < ... < j_s \le d} f_{i_1,...,i_s} f_{j_1,...,j_s} d\mu(\omega)$$

$$\vdots$$

$$f_{12...d} f_{12...d} d\mu(\omega) =$$

$=$ again, use orthogonality for different indices i and j

$$= \int_\Omega \sum_{i=1}^{d} f_i f_i d\mu(\omega) + \int_\Omega \sum_{1 \le i_1 < ... < i_s \le d} f_{i_1,...,i_s} f_{i_1,...,i_s} d\mu(\omega) + ... + f_{12...d} f_{12...d} d\mu(\omega) =$$

$=$ use definition of D

$$\sum_{i=1}^{d} D_{ij} + \sum_{1 \le i_1 < ... < i_s \le d} D_{1 \le i_1 < ... < i_s \le d} + ... + D_{12...d}$$

Remark: If we just use $\mathbb{E}[f] = f_0$ in the beginning this would shorten everything by quite a bit.

Variance-based global sensitivity analysis in chaospy

In `chaospy` it is very easy to compute the first order, second order, or total Sobol' indices for global sensitivity analysis. In the following code snippet, we assume that *model_gpc_approx* denotes the polynomial chaos approximation of the underlying model (obtained via `chaospy`'s infrastructure, i.e. $model\_gpc\_approx = cp.fit\_quadrature(...)$).

```
import chaospy as cp

...

# create the polynomial chaos approximation of the model
model_gpc_approx = cp.fit_quadrature( ... )

# assume that the underlying input distribution is distr
# first order Sobol' indices for global sensitivity analysis
first_order_Sobol_ind = cp.Sens_m(model_gpc_approx, distr)
# second order Sobol' indices for global sensitivity analysis
second_order_Sobol_ind = cp.Sens_m2(model_gpc_approx, distr)
# total Sobol' indices for global sensitivity analysis
total_Sobol_ind = cp.Sens_t(model_gpc_approx, distr)
```

Assignment 2

Consider the model problem, the linear damped oscillator

$$\begin{cases} \frac{d^2y}{dt^2}(t) + c\frac{dy}{dt}(t) + ky(t) = f\cos(\omega_O t) \\ y(0) = y_0 \\ \frac{dy}{dt}(0) = y_1. \end{cases} \tag{6}$$

Let $t \in [0, 20]$, $\Delta t = 0.01$. The output of interest is $y(10)$. Assume that $c \sim \mathcal{U}(0.08, 0.12)$, $k \sim \mathcal{U}(0.03, 0.04)$, $f \sim \mathcal{U}(0.08, 0.12)$, $y_0 \sim \mathcal{U}(0.45, 0.55)$, $y_1 \sim \mathcal{U}(-0.05, 0.05)$ and $\omega_O = 1.0$.

a)

Given $\mathbf{i} = \{c, k, f, y_0, y_1\}$ derive the first order sensitivity formulation for c, i.e. D_c, using eq. (5) and the ANOVA first order term

$$f_i(t, \omega_i) = \int_{\Gamma^{d-1}} f(t, \boldsymbol{\omega}) d\boldsymbol{\omega}_{\sim i} - f_0(t) \tag{7}$$

where $f_0(t) = \mathbb{E}[f(t, \omega)]$.

b)

Write a `python` + `chaospy` program to propagate the uncertainty in (c, k, f, y_0, y_1) through the model in Eq. (6) using the generalized polynomial chaos expansion. Asses the expansion coefficients using the pseudo-spectral approach. Consider both non-sparse and sparse 5D pseudo-spectral computation of the coefficients, constructed on Gaussian nodes. To generate multi-variate quadrature nodes and weights consider $K = 5$ (for the 1D quadrature rule); to construct multi-variate orthogonal polynomials, consider $N = 3$ (for the 1D orthogonal polynomials). Compute the first order and total Sobol' indices for global sensitivity analysis in both cases.

Optional 1: Repeat the above experiment assuming that $c = 0.1$, $k \sim \mathcal{U}(0.03, 0.04)$, $f \sim \mathcal{U}(0.08, 0.12)$, $\omega_O \sim \mathcal{U}(0.80, 1.20)$, $y_0 \sim \mathcal{U}(0.45, 0.55)$, $y_1 \sim \mathcal{U}(-0.05, 0.05)$.

Optional 2: Pick one of the setups. Plot the sobol indices over time.