Local and Global Sensitivity Analysis

Omar Knio[1,2]

[1]Duke University
Department of Mechanical Engineering & Materials Science
omar.knio@duke.edu

[2]KAUST
Division of Computer, Electrical, Mathematical Science & Engineering
omar.knio@kaust.edu.sa

In this section, we will be mainly interested with the following, fairly straightforward but as it turns out delicate and important, question : Given $f(x, y)$ does f vary more with x or with y ?

Why is this important ?

- Planning
- Decision analysis and support
- Designing experiments or field campaigns, etc...

For the discussion that follows, assume we have a minimum of information, specifying :

- shape of f (equation, table, spreadsheet, simulation code, etc...)
- range of x and y
- optionally a design point, say x_0, y_0

L_2 functions over unit-hypercubes

Let $L_2(\mathcal{U}^d)$ be the space of real-valued squared-integrable functions over the d-dimensional hypercube $\mathcal{U}$:

$$f : \boldsymbol{x} \in \mathcal{U}^d \mapsto f(\boldsymbol{x}) \in \mathbb{R}, \quad f \in L_2(\mathcal{U}^d) \Leftrightarrow \int_{\mathcal{U}^d} f(\boldsymbol{x})^2 d\boldsymbol{x} < \infty.$$

$L_2(\mathcal{U}^d)$ is equipped with the inner product $\langle \cdot, \cdot \rangle$,

$$\forall f, g \in L_2(\mathcal{U}^d), \quad \langle u, v \rangle := \int_{\mathcal{U}^d} f(\boldsymbol{x}) g(\boldsymbol{x}) d\boldsymbol{x},$$

and norm $\| \cdot \|_2$,

$$\forall f \in L_2(\mathcal{U}^d), \quad \|f\|_2 := \langle f, f \rangle^{1/2}.$$

L_2 functions over unit-hypercubes

Let $L_2(\mathcal{U}^d)$ be the space of real-valued squared-integrable functions over the d-dimensional hypercube $\mathcal{U}$:

$$f : \boldsymbol{x} \in \mathcal{U}^d \mapsto f(\boldsymbol{x}) \in \mathbb{R}, \quad f \in L_2(\mathcal{U}^d) \Leftrightarrow \int_{\mathcal{U}^d} f(\boldsymbol{x})^2 d\boldsymbol{x} < \infty.$$

NB : all subsequent developments immediately extend to product-type situations, where

$$\boldsymbol{x} \in A = A_1 \times \cdots \times A_d \subseteq \mathbb{R}^d,$$

and weighted spaces $L_2(A, \rho)$,

$$\rho : \boldsymbol{x} \in A \mapsto \rho(\boldsymbol{x}) \geq 0, \quad \rho(\boldsymbol{x}) = \rho_1(x_1) \times \cdots \times \rho_d(x_d).$$

(*e.g.* : ρ is a pdf of a random vector $\boldsymbol{x}$ with mutually independent components.)

Ensemble notations

Let $\mathfrak{D} = \{1, 2, \ldots, d\}$.
Given $\mathfrak{i} \subseteq \mathfrak{D}$, we denote $\mathfrak{i}_\sim := \mathfrak{D} \setminus \mathfrak{i}$ its complement set in $\mathfrak{D}$, such that

$$\mathfrak{i} \cup \mathfrak{i}_\sim = \mathfrak{D}, \quad \mathfrak{i} \cap \mathfrak{i}_\sim = \emptyset.$$

For instance

- $\mathfrak{i} = \{1, 2\}$ and $\mathfrak{i}_\sim = \{3, \ldots, d\}$,
- $\mathfrak{i} = \mathfrak{D}$ and $\mathfrak{i}_\sim = \emptyset$.

Ensemble notations

Let $\mathfrak{D} = \{1, 2, \ldots, d\}$.
Given $\mathfrak{i} \subseteq \mathfrak{D}$, we denote $\mathfrak{i}_\sim := \mathfrak{D} \setminus \mathfrak{i}$ its complement set in $\mathfrak{D}$, such that

$$\mathfrak{i} \cup \mathfrak{i}_\sim = \mathfrak{D}, \quad \mathfrak{i} \cap \mathfrak{i}_\sim = \emptyset.$$

Given $\boldsymbol{x} = (x_1, \ldots, x_d)$, we denote $\boldsymbol{x}_\mathfrak{i}$ the vector having for components the $x_{i \in \mathfrak{i}}$, that is

$$\mathfrak{D} \supseteq \mathfrak{i} = \{i_1, \ldots, i_{|\mathfrak{i}|}\} \Rightarrow \boldsymbol{x}_\mathfrak{i} = (x_{i_1}, \ldots, x_{i_{|\mathfrak{i}|}}),$$

where $|\mathfrak{i}| := \mathrm{Card}(\mathfrak{i})$. For instance

$$\int_{\mathcal{U}^{|\mathfrak{i}|}} f(\boldsymbol{x}) d\boldsymbol{x}_\mathfrak{i} = \int_{\mathcal{U}^{|\mathfrak{i}|}} f(x_1, \ldots, x_d) \prod_{i \in \mathfrak{i}} dx_i,$$

and

$$\int_{\mathcal{U}^{d-|\mathfrak{i}|}} f(\boldsymbol{x}) d\boldsymbol{x}_{\mathfrak{i}_\sim} = \int_{\mathcal{U}^{d-|\mathfrak{i}|}} f(x_1, \ldots, x_d) \prod_{i \in \mathfrak{D}}^{i \notin \mathfrak{i}} dx_i,$$

Sobol-Hoeffding decomposition

Any $f \in L_2(\mathcal{U}^d)$ has a **unique hierarchical orthogonal decomposition** of the form

$$f(\boldsymbol{x}) = f(x_1, \ldots, x_d) = f_0 + \sum_{i=1}^{d} f_i(x_i) + \sum_{i=1}^{d} \sum_{j=i+1}^{d} f_{i,j}(x_i, x_j) + \sum_{i=1}^{d} \sum_{j=i+1}^{d} \sum_{k=j+1}^{d} f_{i,j,k}(x_i, x_j, x_k) + \cdots + f_{1,\ldots,d}(x_1, \ldots, x_d).$$

Hierarchical : 1st order functionals (f_i) $\rightarrow$ 2nd order functionals ($f_{i,j}$) $\rightarrow$ 3rd order functionals ($f_{i,j,l}$) $\rightarrow \cdots \rightarrow$ d-th order functional ($f_{1,\ldots,d}$).

Decomposition in a sum of 2^k functionals

Using ensemble notations :

$$f(\boldsymbol{x}) = \sum_{\mathfrak{i} \subseteq \mathfrak{D}} f_{\mathfrak{i}}(\boldsymbol{x}_{\mathfrak{i}}).$$

Sobol-Hoeffding decomposition

Any $f \in L_2(\mathcal{U}^d)$ has a **unique hierarchical orthogonal decomposition** of the form

$$f(\boldsymbol{x}) = \sum_{\mathfrak{i} \subseteq \mathfrak{D}} f_{\mathfrak{i}}(\boldsymbol{x}_{\mathfrak{i}}).$$

Orthogonal : the functionals if the S-H decomposition verify the following orthogonality relations :

$$\int_{\mathcal{U}} f_{\mathfrak{i}}(\boldsymbol{x}_{\mathfrak{i}}) dx_j = 0, \qquad \forall \mathfrak{i} \subseteq \mathfrak{D},\ j \in \mathfrak{i},$$

$$\int_{\mathcal{U}^d} f_{\mathfrak{i}}(\boldsymbol{x}_{\mathfrak{i}}) f_{\mathfrak{j}}(\boldsymbol{x}_{\mathfrak{j}}) d\boldsymbol{x} = \langle f_{\mathfrak{i}}, f_{\mathfrak{j}} \rangle = 0, \qquad \forall \mathfrak{i}, \mathfrak{j} \subseteq \mathfrak{D},\ \mathfrak{i} \neq \mathfrak{j}.$$

It follows the hierarchical construction

$$f_{\emptyset} = \int_{\mathcal{U}^d} f(\boldsymbol{x}) d\boldsymbol{x} = \langle f \rangle_{\emptyset_\sim = \mathfrak{D}}$$

$$f_{\{i\}} = \int_{\mathcal{U}^{d-1}} f(\boldsymbol{x}) d\boldsymbol{x}_{\{i\}_\sim} - f_{\emptyset} = \langle f \rangle_{\mathfrak{D} \setminus \{i\}} - f_{\emptyset} \qquad i \in \mathfrak{D}$$

$$f_{\mathfrak{i}} = \int_{\mathcal{U}^{|\mathfrak{i}_\sim|}} f(\boldsymbol{x}) d\boldsymbol{x}_{\mathfrak{i}_\sim} - \sum_{\mathfrak{j} \subsetneq \mathfrak{i}} f_{\mathfrak{j}} = \langle f \rangle_{\mathfrak{i}_\sim} - \sum_{\mathfrak{j} \subsetneq \mathfrak{i}} f_{\mathfrak{j}} \qquad \mathfrak{i} \in \mathfrak{D},\ |\mathfrak{i}| \geq 2.$$

Consider a 3D example :

$$Y = f(X_1, X_2, X_3)$$

Then SH yields :

$$f_\varnothing = \mathbb{E}[Y]$$

$$f_{12}(X_1, X_2) = \mathbb{E}[Y|X_1, X_2] - f_1 - f_2 - f_\varnothing$$

$$f_1(X_1) = \mathbb{E}[Y|X_1] - f_\varnothing \qquad f_{13}(X_1, X_3) = \mathbb{E}[Y|X_1, X_3] - f_1 - f_3 - f_\varnothing$$

$$f_2(X_2) = \mathbb{E}[Y|X_2] - f_\varnothing \qquad f_{23}(X_2, X_3) = \mathbb{E}[Y|X_2, X_3] - f_2 - f_3 - f_\varnothing$$

$$f_3(X_3) = \mathbb{E}[Y|X_3] - f_\varnothing$$

$$\begin{aligned} f_{123}(X_1, X_2, X_3) &= \mathbb{E}[Y|X_1, X_2, X_3] - f_1 - f_2 - f_3 \\ &- f_{12} - f_{13} - f_{23} - f_\varnothing \\ \\ &= f(X_1, X_2, X_3) - f_1 - f_2 - f_3 \\ &- f_{12} - f_{13} - f_{23} - f_\varnothing \end{aligned}$$

Parametric sensitivity analysis

Consider $\boldsymbol{x}$ as a set of d independent random parameters uniformly distributed on $\mathcal{U}^d$, and $f(\boldsymbol{x})$ a model-output depending on these random parameters. It is assumes that f is a 2nd order random variable : $f \in L_2(\mathcal{U}^d)$. Thus, f has a unique S-H decomposition

$$f(\boldsymbol{x}) = \sum_{\mathfrak{i} \subseteq \mathfrak{D}} f_{\mathfrak{i}}(\boldsymbol{x}_{\mathfrak{i}}).$$

Further, the integrals of f with respect to $\mathfrak{i}_\sim$ are in this context conditional expectations,

$$\mathbb{E}[f|\boldsymbol{x}_{\mathfrak{i}}] = \int_{\mathcal{U}^{|\mathfrak{i}_\sim|}} f(\boldsymbol{x}) d\boldsymbol{x}_{\mathfrak{i}_\sim} = g(\boldsymbol{x}_{\mathfrak{i}}) \quad \forall \mathfrak{i} \subseteq \mathfrak{D},$$

so the S-H decomposition follows the hierarchical structure

$$\begin{aligned} f_\emptyset &= \mathbb{E}[f] \\ f_{\{i\}} &= \mathbb{E}\left[f|\boldsymbol{x}_{\{i\}}\right] - \mathbb{E}[f] && i \in \mathfrak{D} \\ f_{\mathfrak{i}} &= \mathbb{E}[f|\boldsymbol{x}_{\mathfrak{i}}] - \sum_{\mathfrak{j} \subsetneq \mathfrak{i}} f_{\mathfrak{j}} && \mathfrak{i} \subseteq \mathfrak{D},\ |\mathfrak{i}| \geq 2. \end{aligned}$$

Variance decomposition

Because of the orthogonality of the S-H decomposition the variance $\mathbb{V}[f]$ of the model-output can be decomposed as

$$\mathbb{V}[f] = \sum_{\mathfrak{i} \subseteq \mathfrak{D}}^{\mathfrak{i} \neq \emptyset} \mathbb{V}[f_{\mathfrak{i}}], \quad \mathbb{V}[f_{\mathfrak{i}}] = \langle f_{\mathfrak{i}}, f_{\mathfrak{i}} \rangle .$$

$\mathbb{V}[f_{\mathfrak{i}}]$ is interpreted as the contribution to the total variance $\mathbb{V}[f]$ of the interaction between parameters $x_{i \in \mathfrak{i}}$.

The S-H decomposition thus provide a rich mean of analyzing the respective contributions of individual or sets of parameters to model-output variability. However, as there are $2^d - 1$ contributions, so one needs more "abstract" characterizations.

Sensitivity indices

To facilitate the hierarchization of the respective influence of each parameter x_i, the partial variances $\mathbb{V}[f_{\mathfrak{i}}]$ are normalized by $\mathbb{V}[f]$ to obtain the sensitivity indices :

$$S_{\mathfrak{i}}(f) = \frac{\mathbb{V}[f_{\mathfrak{i}}]}{\mathbb{V}[f]} \leq 1, \quad \sum_{\mathfrak{i} \subseteq \mathfrak{D}}^{\mathfrak{i} \neq \emptyset} S_{\mathfrak{i}}(f) = 1.$$

The order of the sensitivity indices $S_{\mathfrak{i}}$ is equal to $|\mathfrak{i}| = \mathrm{Card}(\mathfrak{i})$.

1st order sensitivity indices. The d first order indices $S_{\{i\} \in \mathfrak{D}}$ characterize the fraction of the variance due the parameter x_i **only**, *i.e.* without any interaction with others. Therefore,

$$1 - \sum_{i=1}^{d} S_{\{i\}}(f) \geq 0,$$

measures globally the effect on the variability of all interactions between parameters. If $\sum_{i=1}^{d} S_{\{i\}} = 1$, the model is said **additive**, because its S-H decomposition is

$$f(x_i, \ldots, x_d) = f_0 + \sum_{i=1}^{d} f_i(x_i),$$

and the impact of the parameters can be studied separately.

Sensitivity indices

To facilitate the hierarchization of the respective influence of each parameter x_i, the partial variances $\mathbb{V}[f_{\mathfrak{i}}]$ are normalized by $\mathbb{V}[f]$ to obtain the sensitivity indices :

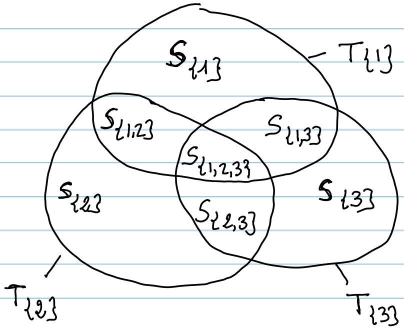
$$S_{\mathfrak{i}}(f) = \frac{\mathbb{V}[f_{\mathfrak{i}}]}{\mathbb{V}[f]} \leq 1, \quad \sum_{\mathfrak{i} \subseteq \mathfrak{D}}^{\mathfrak{i} \neq \emptyset} S_{\mathfrak{i}}(f) = 1.$$

The order of the sensitivity indices $S_{\mathfrak{i}}$ is equal to $|\mathfrak{i}| = \mathrm{Card}(\mathfrak{i})$.

Total sensitivity indices. The first order SI $S_{\{i\}}$ measures the variability due to parameter x_i alone. The total SI $T_{\{i\}}$ measures the variability due to the parameter x_i, **including** all its interactions with other parameters :

$$T_{\{i\}} := \sum_{\mathfrak{i} \ni i} S_{\mathfrak{i}} \geq S_{\{i\}}.$$

Important point : for x_i to be deemed non-important or non-influent on the model-output, $S_{\{i\}}$ **and** $T_{\{i\}}$ have to be negligible.
Observe that $\sum_{i \in \mathfrak{D}} T_{\{i\}} \geq 1$, the excess from 1 characterizes the presence of interactions in the model-output.



Case of $d=3$.

$$\sum S_i = 1$$

$$\sum S_{\{i\}} \leq 1$$

$$\sum T_{\{i\}} \geq 1$$

Sensitivity indices

In many uncertainty problem, the set of uncertain parameters can be naturally grouped into subsets depending on the process each parameter accounts for. For instance, boundary conditions BC, material property φ, external forcing F, and $\mathfrak{D}$ is the union of these distinct subsets :

$$\mathfrak{D} = \mathfrak{D}_{BC} \cup \mathfrak{D}_{\varphi} \cup \mathfrak{D}_{F}.$$

The notion of first order and total sensitivity indices can be extended to characterize the influence of the subsets of parameters. For instance,

$$S_{\mathfrak{D}_{\varphi}} = \sum_{\mathfrak{i} \subseteq \mathfrak{D}_{\varphi}} S_{\mathfrak{i}},$$

measures the fraction of variance induced by the material uncertainty alone, while

$$T_{\mathfrak{D}_F} = \sum_{\mathfrak{i} \cap \mathfrak{D}_F \neq \emptyset} S_{\mathfrak{i}}.$$

measures the fraction of variance due to the external forcing uncertainty and all its interactions.

Sensitivity index S_u, total sensitivity index T_u

- it is easy to verify that

$$S_u + T_{u\sim} = S_{u\sim} + T_u = 1$$

- it immediately follows that $\sum_i S_{\{i\}} \leq 1 \quad \Rightarrow \quad \sum_i T_{\{i\}} \geq 1$
- S_u high means that X_u has substantial contribution to $V(Y)$
- T_u low means that X_u has weak contribution to $V(Y)$ (and so it is a candidate to be dropped)

S-H decomposition from PC expansions.

Consider the model-output $f : \boldsymbol{\xi} \in \Xi \subset \mathbb{R}^d \mapsto \mathbb{R}$, where $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_d)$ are independent real-valued r.v. with joint-probability density function

$$p_{\boldsymbol{\xi}}(x_1, \ldots, x_d) = \prod_{i=1}^{d} p_i(x_i).$$

Let $\{ \Psi_{\boldsymbol{\alpha}} \}$ be the set of d-variate orthogonal polynomials,

$$\Psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \prod_{i=1}^{d} \psi^{(i)}_{\alpha_i}(\xi_i),$$

with $\psi^{(i)}_{l \geq 0} \in \pi_l$ the uni-variate polynomials mutually orthogonal with respect to the density p_i.

If $f \in L_2(\Xi, p_{\boldsymbol{\xi}})$, it has a convergent PC expansion

$$f(\boldsymbol{\xi}) = \lim_{\mathrm{No} \to \infty} \sum_{|\boldsymbol{\alpha}| \leq \mathrm{No}} \Psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi}) f_{\boldsymbol{\alpha}}, \quad |\boldsymbol{\alpha}| = \sum_{i=1}^{d} |\alpha_i|.$$

S-H decomposition from PC expansions.

Given the truncated PC expansion of f,

$$\hat{f}(\boldsymbol{\xi}) = \sum_{|\boldsymbol{\alpha}| \le \mathcal{A}} \Psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi}) f_{\boldsymbol{\alpha}},$$

one can readily obtain the PC approximation of the S-H functionals through

$$\hat{f}_{\mathfrak{i}}(\boldsymbol{\xi}_{\mathfrak{i}}) = \sum_{|\boldsymbol{\alpha}| \le \mathcal{A}(\mathfrak{i})} \Psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi}_{\mathfrak{i}}) f_{\boldsymbol{\alpha}},$$

where the multi-index set $\mathcal{A}(\mathfrak{i})$ is given by

$$\mathcal{A}(\mathfrak{i}) = \{\boldsymbol{\alpha} \in \mathcal{A}; \alpha_i > 0 \text{ for } i \in \mathfrak{i}, \alpha_i = 0 \text{ for } i \notin \mathfrak{i}\} \subsetneq \mathcal{A}.$$

For the sensitivity indices it comes

$$S_{\mathfrak{i}}(\hat{f}) = \frac{\sum_{\boldsymbol{\alpha} \in \mathcal{A}(\mathfrak{i})} f_{\boldsymbol{\alpha}}^2 \langle \Psi_{\boldsymbol{\alpha}}, \Psi_{\boldsymbol{\alpha}} \rangle}{\sum_{\boldsymbol{\alpha} \in \mathcal{A}} f_{\boldsymbol{\alpha}}^2 \langle \Psi_{\boldsymbol{\alpha}}, \Psi_{\boldsymbol{\alpha}} \rangle}, \quad T_{\{i\}}(\hat{f}) = \frac{\sum_{\boldsymbol{\alpha} \in \mathcal{T}(\mathfrak{i})} f_{\boldsymbol{\alpha}}^2 \langle \Psi_{\boldsymbol{\alpha}}, \Psi_{\boldsymbol{\alpha}} \rangle}{\sum_{\boldsymbol{\alpha} \in \mathcal{A}} f_{\boldsymbol{\alpha}}^2 \langle \Psi_{\boldsymbol{\alpha}}, \Psi_{\boldsymbol{\alpha}} \rangle},$$

where

$$\mathcal{T}(\mathfrak{i}) = \{\boldsymbol{\alpha} \in \mathcal{A}; \alpha_i > 0 \text{ for } i \in \mathfrak{i}\}$$