GSA of Model Output With Dependent Input: Part II Sensitivity Indices

Thierry A. Mara

University of Reunion

12$^{\text{th}}$ SAMO Summer School
June 24-28 2024
University of Parma

Outline

Case 1: Marginal and conditional cdfs known

Interpretation of sensitivity indices

Case 2: Marginal and conditional pdfs unknown

Rejection sampling (MCMC)

About Shapley

Let $\boldsymbol{x} \sim p_x \neq p_{x_1} p_{x_2} \dots p_{x_d}$ be the input vector of the model response $y = f(\boldsymbol{x})$. If $f(\boldsymbol{x})$ is square-integrable we can still obtain a unique decomposition of the form:

$$f(\boldsymbol{x}) = f_0 + \sum_{i_1=1}^{d} f_{i_1}(x_{i_1}) + \sum_{i_2>i_1}^{d} f_{i_1,i_2}(x_{i_1}, x_{i_2}) + \cdots + f_{1,\dots,d}(x_1, \dots, x_d)$$

But this time the functions $f_{i_1,i_2,\dots}$ are not orthogonal to each other. However, if we know $\boldsymbol{u} \sim \mathcal{U}(0,1)^d$ the Rosenblatt transform of $\boldsymbol{x}$, we can write,

$$f(\boldsymbol{u}) = f_0 + \sum_{i_1=1}^{d} f_{i_1}(u_{i_1}) + \sum_{i_2>i_1}^{d} f_{i_1,i_2}(u_{i_1}, u_{i_2}) + \cdots + f_{1,\dots,d}(u_1, \dots, u_d)$$

with the $f_{i_1,i_2,\dots}$s orthogonal to each other. But the Rosenblatt is not unique $\Leftrightarrow$ the ANOVA decomposition (in the Sobol' sense) is not unique.

The interpretation of the sensitivity indices of u_i as those of x_i is the following:

- Because u_{i_1} stems from the unconditional transformation of x_{i_1}, the sensitivity indices of u_{i_1} are interpreted as the full sensitivity indices of x_{i_1} that account for its mutual contribution to the variance of y because of its dependence on the other variables
- Because u_{i_2} stems from the transformation of $x_{i_2}|x_{i_1}$, the sensitivity indices of u_{i_2} are interpreted as those of x_{i_2} without its mutual contribution due to its dependence on x_{i_1}
- The sensitivity indices of u_{i_3} are those of x_{i_3} without its mutual contribution due to its dependence on (x_{i_1}, x_{i_2})
- $\vdots$
- The sensitivity indices of u_{i_d} are those of x_{i_d} that do not account for its mutual dependence contribution with the other variables.

The interpretation of the sensitivity indices of u_i as those of x_i is the following:

- Because u_{i_1} stems from the unconditional transformation of x_{i_1}, the sensitivity indices of u_{i_1} are interpreted as the full sensitivity indices of x_{i_1} that account for its mutual contribution to the variance of y because of its dependence on the other variables
- Because u_{i_2} stems from the transformation of $x_{i_2}|x_{i_1}$, the sensitivity indices of u_{i_2} are interpreted as those of x_{i_2} without its mutual contribution due to its dependence on x_{i_1}

Hence, one can define two kind of indices: the Independent and Full Sensitivity Indices

The first one measures the importance of a variable without the mutual contribution due to its dependence on the other variables, while the second one does.

The interpretation of the sensitivity indices of u_i as those of x_i is the following:

- Because u_{i_1} stems from the unconditional transformation of x_{i_1}, the sensitivity indices of u_{i_1} are interpreted as the full sensitivity indices of x_{i_1} that account for its mutual contribution to the variance of y because of its dependence on the other variables
- Because u_{i_2} stems from the transformation of $x_{i_2}|x_{i_1}$, the sensitivity indices of u_{i_2} are interpreted as those of x_{i_2} without its mutual contribution due to its dependence on x_{i_1}

For instance,
First-order Sobol' indices (S_i^{full}, S_i^{ind}), pdf-based importance measure $(\delta_i^{full}, \delta_i^{ind})$, etc.
Total-order Sobol' indices - (T_i^{full}, T_i^{ind}), cdf-based importance measure $(\tau_i^{full}, \tau_i^{ind})$, etc.

Interpretation of the sensitivity indices: Mara & Tarantola (RESS, 2012)

The interpretation of the sensitivity indices of u_i as those of x_i is the following:

- Because u_{i_1} stems from the unconditional transformation of x_{i_1}, the sensitivity indices of u_{i_1} are interpreted as the full sensitivity indices of x_{i_1} that account for its mutual contribution to the variance of y because of its dependence on the other variables
- Because u_{i_2} stems from the transformation of $x_{i_2}|x_{i_1}$, the sensitivity indices of u_{i_2} are interpreted as those of x_{i_2} without its mutual contribution due to its dependence on x_{i_1}

N.B.: But because the Rosenblatt transformation is not unique, we need to consider several RTs to get the overall indices (i.e. full+independent).

Suppose that the RT is applied to $(x_1, \ldots, x_d)$, that is

$$\begin{cases} u_1 = & F_{x_1}(x_1) \\ u_2 = & F_{x_2|x_1}(x_2|x_1) \\ \vdots & \\ u_d = & F_{x_d|\boldsymbol{x}_{\sim d}}(x_d|\boldsymbol{x}_{\sim d}) \end{cases}$$

then we can compute the following sensitivity indices (among others):

Variance-based: (see Kucherenko et al. 2012, Tarantola and Mara 2017)

$$S_1^{full} = \frac{\mathrm{Var}\left[\mathbb{E}\left[y|u_1\right]\right]}{\mathrm{Var}\left[y\right]} \tag{1}$$

$$ST_1^{full} = \frac{\mathbb{E}\left[\mathrm{Var}\left[y|\boldsymbol{u}_{\sim 1}\right]\right]}{\mathrm{Var}\left[y\right]} \tag{2}$$

$$S_d^{ind} = \frac{\mathrm{Var}\left[\mathbb{E}\left[y|u_d\right]\right]}{\mathrm{Var}\left[y\right]} \tag{3}$$

$$ST_d^{ind} = \frac{\mathbb{E}\left[\mathrm{Var}\left[y|\boldsymbol{u}_{\sim d}\right]\right]}{\mathrm{Var}\left[y\right]} \tag{4}$$

Suppose that the RT is applied to $(x_1, \ldots, x_d)$, that is

$$\begin{cases} u_1 = & F_{x_1}(x_1) \\ u_2 = & F_{x_2|x_1}(x_2|x_1) \\ \vdots & \\ u_d = & F_{x_d|\boldsymbol{x}_{\sim d}}(x_d|\boldsymbol{x}_{\sim d}) \end{cases}$$

then we can compute the following sensitivity indices (among others):
pdf-based: (see Mara and Becker 2021)

$$\delta_1^{full} = \frac{1}{2}\int_{\mathbb{R}}\int_0^1 |p_y - p_{y|u_1}|\mathrm{d}y\mathrm{d}u_1 \tag{1}$$

$$\delta_d^{ind} = \frac{1}{2}\int_{\mathbb{R}}\int_0^1 |p_y - p_{y|u_d}|\mathrm{d}y\mathrm{d}u_d \tag{2}$$

Suppose that the RT is applied to $(x_1, \ldots, x_d)$, that is

$$\begin{cases} u_1 = & F_{x_1}(x_1) \\ u_2 = & F_{x_2|x_1}(x_2|x_1) \\ \quad\vdots & \\ u_d = & F_{x_d|\boldsymbol{x}_{\sim d}}(x_d|\boldsymbol{x}_{\sim d}) \end{cases}$$

then we can compute the following sensitivity indices (among others):
<u>cdf-based</u>:

$$\tau_1^{full} = \int_0^1 \sup |F_y - F_{y|u_1}| \mathrm{d}u_1 \quad (1)$$

$$\tau_d^{ind} = \int_0^1 \sup |F_y - F_{y|u_d}| \mathrm{d}u_d \quad (2)$$

Suppose that the RT is applied to $(x_2, x_3, \ldots, x_d, x_1)$, that is

$$\begin{cases} u_2 = & F_{x_1}(x_2) \\ u_3 = & F_{x_3|x_2}(x_3|x_2) \\ \quad\vdots & \\ u_d = & F_{x_d|x_2,x_3,\ldots,x_{d-1}}(x_d|x_2,x_3,\ldots,x_{d-1}) \\ u_1 = & F_{x_1|\boldsymbol{x}_{\sim 1}}(x_1|\boldsymbol{x}_{\sim 1}) \end{cases}$$

then we can compute the following sensitivity indices (among others):

<u>Variance-based</u>:

$$\begin{aligned} S_2^{full} &= \frac{\mathrm{Var}\left[\mathbb{E}\left[y|u_2\right]\right]}{\mathrm{Var}\left[y\right]} \\ ST_2^{full} &= \frac{\mathbb{E}\left[\mathrm{Var}\left[y|\boldsymbol{u}_{\sim 2}\right]\right]}{\mathrm{Var}\left[y\right]} \\ S_1^{ind} &= \frac{\mathrm{Var}\left[\mathbb{E}\left[y|u_1\right]\right]}{\mathrm{Var}\left[y\right]} \\ ST_1^{ind} &= \frac{\mathbb{E}\left[\mathrm{Var}\left[y|\boldsymbol{u}_{\sim 1}\right]\right]}{\mathrm{Var}\left[y\right]} \end{aligned}$$

Suppose that the RT is applied to $(x_2, x_3, \ldots, x_d, x_1)$, that is

$$\begin{cases} u_2 = & F_{x_1}(x_2) \\ u_3 = & F_{x_3|x_2}(x_3|x_2) \\ \quad\vdots & \\ u_d = & F_{x_d|x_2,x_3,\ldots,x_{d-1}}(x_d|x_2, x_3, \ldots, x_{d-1}) \\ u_1 = & F_{x_1|\boldsymbol{x}_{\sim 1}}(x_1|\boldsymbol{x}_{\sim 1}) \end{cases}$$

then we can compute the following sensitivity indices (among others):

<u>pdf-based</u>:

$$\begin{aligned} \delta_2^{full} &= \frac{1}{2}\int_{\mathbb{R}}\int_0^1 |p_y - p_{y|u_2}| \mathrm{d}y \mathrm{d}u_2 \\ \delta_1^{ind} &= \frac{1}{2}\int_{\mathbb{R}}\int_0^1 |p_y - p_{y|u_1}| \mathrm{d}y \mathrm{d}u_1 \end{aligned}$$

Suppose that the RT is applied to $(x_2, x_3, \ldots, x_d, x_1)$, that is

$$\begin{cases} u_2 = & F_{x_1}(x_2) \\ u_3 = & F_{x_3|x_2}(x_3|x_2) \\ \quad\vdots & \\ u_d = & F_{x_d|x_2,x_3,\ldots,x_{d-1}}(x_d|x_2, x_3, \ldots, x_{d-1}) \\ u_1 = & F_{x_1|\boldsymbol{x}_{\sim 1}}(x_1|\boldsymbol{x}_{\sim 1}) \end{cases}$$

then we can compute the following sensitivity indices (among others):
<u>cdf-based</u>:

$$\tau_2^{full} = \int_0^1 \sup |F_y - F_{y|u_2}| \mathrm{d}u_2 \quad (3)$$

$$\tau_1^{ind} = \int_0^1 \sup |F_y - F_{y|u_1}| \mathrm{d}u_1 \quad (4)$$

By proceeding as such with all possible circular permutations (i.e., with $(x_3, x_4, \ldots, x_d, x_1, x_2)$, $(x_4, x_5, \ldots, x_d, x_1, x_2, x_3)$, etc.) one can compute the overall full and independent sensitivity indices of interest.

Computational cost: for given data methods like the BSPCE only N model runs are necessary, for the variance-based IA-estimator $(2 + 4d)N$ model runs are required.

If the u-values are unknown, the only sensitivity indices that you can compute are the **full sensitivity indices**, namely,

$$S_i^{full} = \frac{\mathrm{Var}\left[\mathbb{E}\left[y|x_i\right]\right]}{\mathrm{Var}\left[y\right]}$$

$$\delta_i^{full} = \frac{1}{2}\int_{\mathbb{R}}\int_{\mathbb{R}} |p_y - p_{y|x_i}|\mathrm{d}y\mathrm{d}x_i$$

$$\tau_i^{full} = \int_{\mathbb{R}} \sup |F_y - F_{y|x_i}|\mathrm{d}x_i$$

For the Sobol' indices we need to find an estimate $\mathbb{E}\left[y|x_i\right]$ by regression. There is also the possibility to estimate the total-independent Sobol' indices

$$T_i^{ind} = \frac{\mathbb{E}\left[\mathrm{Var}\left[y|\boldsymbol{x}_{\sim i}\right]\right]}{\mathrm{Var}\left[y\right]} = 1 - \frac{\mathrm{Var}\left[\mathbb{E}\left[y|\boldsymbol{x}_{\sim i}\right]\right]}{\mathrm{Var}\left[y\right]}$$

for this we must fit a d-1 dimensional function to estimate $\mathbb{E}\left[y|\boldsymbol{x}_{\sim i}\right]$.

Shapley effect

The Shapley effect (Owen 2014) is a concept stemming from game theory (Shapley 1953). It is defined as follows

$$Sh_i = \sum_{v \subseteq \boldsymbol{x}_{\sim i}} \frac{|\boldsymbol{v}|!\,(d-1-|\boldsymbol{v}|)!}{d!} \left(S^{clo}_{\boldsymbol{v} \cup x_i} - S^{clo}_{\boldsymbol{v}} \right)$$

where, $S^{clo}_{v} = \frac{\mathrm{Var}[\mathbb{E}[y|\boldsymbol{v}]]}{\mathrm{Var}[y]}$ is the first-order effect of the group of inputs $\boldsymbol{v} \subseteq \boldsymbol{x}_{\sim i}$.

Shapley effect

The Shapley effect (Owen 2014) is a concept stemming from game theory (Shapley 1953). It is defined as follows

$$Sh_i = \sum_{v \subseteq \boldsymbol{x}_{\sim i}} \frac{|\boldsymbol{v}|!\,(d-1-|\boldsymbol{v}|)!}{d!} \left(S^{clo}_{\boldsymbol{v} \cup x_i} - S^{clo}_{\boldsymbol{v}} \right)$$

where, $S^{clo}_{v} = \frac{\mathrm{Var}[\mathbb{E}[y|\boldsymbol{v}]]}{\mathrm{Var}[y]}$ is the first-order effect of the group of inputs $\boldsymbol{v} \subseteq \boldsymbol{x}_{\sim i}$.
Therefore, we just need to compute $S^{full,clo}_{v} = \frac{\mathrm{Var}[\mathbb{E}[y|\boldsymbol{v}]]}{\mathrm{Var}[y]}\ \forall \boldsymbol{v} \subseteq \boldsymbol{x}$

Shapley effect

The Shapley effect (Owen 2014) is a concept stemming from game theory (Shapley 1953). It is defined as follows

$$Sh_i = \sum_{v \subseteq \boldsymbol{x}_{\sim i}} \frac{|\boldsymbol{v}|!\,(d-1-|\boldsymbol{v}|)!}{d!} \left(S^{clo}_{\boldsymbol{v} \cup x_i} - S^{clo}_{\boldsymbol{v}} \right)$$

where, $S^{clo}_v = \frac{\mathrm{Var}[\mathbb{E}[y|\boldsymbol{v}]]}{\mathrm{Var}[y]}$ is the first-order effect of the group of inputs $\boldsymbol{v} \subseteq \boldsymbol{x}_{\sim i}$.

Therefore, we just need to compute $S^{full,clo}_v = \frac{\mathrm{Var}[\mathbb{E}[y|\boldsymbol{v}]]}{\mathrm{Var}[y]}$ $\forall \boldsymbol{v} \subseteq \boldsymbol{x}$

Main obstacle: We need to compute the overall $2^d - 1$ Sobol' indices.

Some References

- Kucherenko S, Tarantola S, Annoni P. (2012). Comput Phys Comm, 937–46.
- Mara TA, Tarantola S.(2012). Reliab Eng Syst Saf 115–21
- Mara TA, Becker WE. (2021). Reliab Eng Syst Saf, 107795
- Rosenblatt, M (1952). Annals Math. Stat., 470–472