Global sensitivity analysis and dimension reduction

Clémentine PRIEUR
Grenoble Alpes University

Summer school on GSA & Poincaré inequalities
6-8 July 2022, Toulouse

Part IV

Global Sensitivity Analysis for dependent inputs: we present variance-based GSA for dependent inputs.

Variance-based sensitivity analysis with dependent inputs

Introduction

In this talk, we consider

$$\mathcal{M}\colon \left\{ \begin{array}{ccc} \mathcal{X} = \mathcal{X}_1 \times \ldots \mathcal{X}_d & \rightarrow & \mathcal{Y} \\ \boldsymbol{x} = (x_1, \ldots, x_d) & \mapsto & y = \mathcal{M}(\boldsymbol{x}) \end{array} \right.$$

with

- $\mathcal{M}$: mathematical or numerical model,
- $\boldsymbol{x}$: uncertain input parameters,
- y : output.

We model the uncertain input parameters by a probability distribution P on $\mathcal{X}$ and get

$$Y = \mathcal{M}(X_1, \ldots, X_d)$$

with the vector $\mathbf{X} = (X_1, \ldots, X_d)$ distributed as P.

Introduction

Independent framework: $P(d\boldsymbol{x}) = P_1(dx_1)\dots P_d(dx_d)$

Why is the independent framework not always the right one?

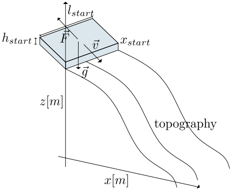
In the following, we consider an application to long-term avalanche hazard assessment. The model under consideration is:

- a snow avalanche model, joint work with INRAE (Grenoble, FRANCE).

Snow avalanche modeling

Model based on depth-averaged Saint-Venant equations (see [Heredia et al., 2020] for more details)

$$\begin{aligned} \frac{\partial h}{\partial t} + \frac{\partial hv}{\partial x} &= 0 \\ \frac{\partial hv}{\partial t} + \frac{\partial}{\partial x}\left(hv^2 + \frac{h^2}{2}\right) &= h\left(g\sin\theta - \mathrm{F}\right) \end{aligned}$$



with $v = \|\vec{\mathbf{v}}\|$ the flow velocity, h the flow depth, θ the local angle, t the time, g the gravity constant and $\mathrm{F} = \|\vec{\mathbf{F}}\|$ a frictional force. The model uses the Voellmy frictional force $\mathrm{F} = \mu g cos\theta + g/(\xi h)v^2$, where μ and ξ are friction parameters.

The equations are solved with a finite volume scheme [Naaim, 1998]. The topography is the one of a path located in Bessans, France.

Let us present one of the two scenarii presented in [Heredia et al., 2020].

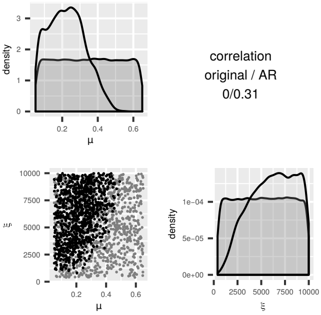
Input	Description	Distribution
μ	Static friction coefficient	$\mathcal{U}[0.05, 0.65]$
ξ	Turbulent friction [m/s^2]	$\mathcal{U}[400, 10000]$
$\mathrm{l_{start}}$	Length of the release zone [m]	$\mathcal{U}[5, 300]$
$\mathrm{h_{start}}$	Mean snow depth in the release zone [m]	$\mathcal{U}[0.05, 3]$
$\mathrm{x_{start}}$	Release abscissa [m]	$\mathcal{U}[0, 1600]$

Let's $\mathrm{vol_{start}} = l_{start} \times h_{start} \times 72.3/\cos(35^\circ)$ instead of $\mathrm{h_{start}}$ and $\mathrm{l_{start}}$.

AR rules:

- avalanche simulation is flowing in $[1600m, 2412m]$,
- $vol > 7000m^3$,
- runout distance $< 2500m$ (end of the path).

From $n_0 = 100\,000$ initial runs, we keep $n_1 = 6152$ constrained ones.



Variance based SA in the general framework

We still consider $\mathcal{M}: \left\{ \begin{array}{ccc} \mathbb{R}^d & \to & \mathbb{R} \\ \boldsymbol{x} = (x_1, \ldots, x_d) & \mapsto & y = \mathcal{M}(\boldsymbol{x}) \end{array} \right.$

Uncertain parameters are no longer assumed independent, thus $P(d\boldsymbol{x})$ is not necessarily equal to $P_1(dx_1) \ldots P_d(dx_d)$. We have $F_{\mathbf{X}}(\boldsymbol{x}) = C\left(F_{X_1}(x_1), \ldots, F_{X_d}(x_d)\right)$ (Sklar's Theorem) with $F_{X_i}(\cdot)$ and $F_{\mathbf{X}}(\cdot)$ the cdf of X_i, $\boldsymbol{X}$. If the F_{X_i} are continuous, then the copula C is unique.

We still define, for any $i \in \{1, \ldots, d\}$: $S_i = \dfrac{V\left[\mathbb{E}\left[Y|X_i\right]\right]}{V\left[Y\right]}$ and $S_i^{\text{tot}} = \dfrac{\mathbb{E}\left[V\left[Y|\boldsymbol{X}_{-i}\right]\right]}{V\left[Y\right]}$.

However, nice properties due to orthogonality are lost.

An alternative, the Shapley effects

Let $\mathcal{D} = \{1, \ldots, d\}$. Let team $u \subseteq \mathcal{D}$ create value $\mathbf{val}(u)$. Total value is $\mathbf{val}(\mathcal{D})$. We attribute ϕ_i of this to $i \in \mathcal{D}$.

Shapley axioms [Shapley, 1953]

- Efficiency $\sum_{i=1}^{d} \phi_i = \mathbf{val}(\mathcal{D})$
- Dummy If $\mathbf{val}(u \cup \{i\}) = \mathbf{val}(u)$ for all $u \subseteq \mathcal{D}$, then $\phi_i = 0$
- Symmetry If $\mathbf{val}(u \cup \{i\}) = \mathbf{val}(u \cup \{j\})$ for all $u \cap \{i, j\} = \emptyset$, then $\phi_i = \phi_j$
- Additivity If games **val**, **val'** have values ϕ, ϕ', then **val** + **val'** has value $\phi + \phi'$

Unique solution

$$\phi_i = \frac{1}{d} \sum_{u \subseteq -\{i\}} \binom{d-1}{|u|}^{-1} \left(\mathbf{val}(u+i) - \mathbf{val}(u)\right)$$

Let $X_1, \ldots, X_d$ be the team members trying to explain the variability of $\mathcal{M}$. The value of any $u \in \mathcal{D}$ is how much can be explained by $\boldsymbol{X}_u$.

We choose $\mathbf{val}(u) = \dfrac{V\left[\mathbb{E}\left[Y|\boldsymbol{X}_u\right]\right]}{V\left[Y\right]}$ which leads to the definition of Shapley effects [Owen, 2014]:

$$\phi_i = \frac{1}{d} \sum_{u \subseteq -\{i\}} \binom{d-1}{|u|}^{-1} \left(\frac{V\left[\mathbb{E}\left[Y|\boldsymbol{X}_u, X_i\right]\right]}{V\left[Y\right]} - \frac{V\left[\mathbb{E}\left[Y|\boldsymbol{X}_u\right]\right]}{V\left[Y\right]} \right)$$

It is equivalent to consider to choose $\widetilde{\mathbf{val}}(u) = \dfrac{\mathbb{E}\left[V\left[Y|\boldsymbol{X}_{-u}\right]\right]}{V\left[Y\right]}$ [Song et al., 2016].

Main properties

Independent framework: $\forall\, i = 1, \ldots, d$, $\phi_i = \sum_{\mathbf{u}: i \in \mathbf{u}} \frac{1}{|\mathbf{u}|} S_{\mathbf{u}}$

We also have: $\forall\, i = 1, \ldots, d$, $0 \leq S_i \leq \phi_i \leq S_i^{\text{tot}} \leq 1$ and $\sum_{i=1}^{d} \phi_i = 1$.

Main properties

Independent framework: $\forall\, i = 1, \ldots, d$, $\phi_i = \sum_{\mathbf{u}: i \in \mathbf{u}} \frac{1}{|\mathbf{u}|} S_{\mathbf{u}}$

We also have: $\forall\, i = 1, \ldots, d$, $0 \leq S_i \leq \phi_i \leq S_i^{\text{tot}} \leq 1$ and $\sum_{i=1}^{d} \phi_i = 1$.

Dependent framework:

In this framework, we still have $0 \leq \phi_i \leq 1$ and $\sum_{i=1}^{d} \phi_i = 1$

We do not necessarily have $S_i \leq \phi_i \leq S_i^{\text{tot}}$

The Shapley allocation rule is based on an equitable principle, which ensures that $\phi_i \approx 0 \Rightarrow X_i$ has no significant contribution to $\text{Var}[Y]$, neither by its interactions nor by its dependencies with other inputs.

If output is multivariate or the discretization of a functional output $\mathbf{Y} = (Y_1, \ldots, Y_p)$, we define aggregated Shapley effects as:

$$\forall\, 1 \leq j \leq p\,,\ \forall\, 1 \leq i \leq d\,,\ \phi_i^{\text{agg}} = \frac{\sum_{j=1}^{p} V[Y_j]\phi_i^j}{\sum_{j=1}^{p} V[Y_j]}$$

with ϕ_i^j defined as the Shapley effect of Y_j associated to input X_i [Heredia et al., 2020] (see also [Lamboni et al., 2011]).

Proposition [Heredia et al., 2020, Prop. 2.1]

The set of aggregated Shapley effects $\left(\phi_i^{\text{agg}}\,,\ i \in \{1, \ldots, d\}\right)$ *correspond to the set of Shapley values with characteristic function:*

$$u \subseteq \{1, \ldots, d\} \mapsto \boldsymbol{val}(u) = \frac{\sum_{j=1}^{p} V[Y_j]\boldsymbol{val}_j(u)}{\sum_{j=1}^{p} V[Y_j]}$$

with $\boldsymbol{val}_j(u) = \dfrac{V\left[\mathbb{E}\left[Y_j|\mathbf{X}_u\right]\right]}{V\left[Y_j\right]}$ *or* $\boldsymbol{val}_j(u) = \dfrac{\mathbb{E}\left[V\left[Y_j|\mathbf{X}_{-u}\right]\right]}{V\left[Y_j\right]}$.

What about algorithms?

Algorithms to compute Shapley effects [Castro et al., 2009] are based on the value function $u \mapsto \dfrac{\mathbb{E}[V[Y|\mathbf{X}_{-u}]]}{V[Y]}$. Note that

$$\phi_i = \frac{1}{d\,!} \sum_{\pi \in \Pi(\{1,\ldots,d\})} \left(\widetilde{\mathbf{val}}(P_i(\pi) \cup \{i\})) - \widetilde{\mathbf{val}}(P_i(\pi)) \right)$$

with $\Pi(\{1,\ldots,d\})$ the set of all possible permutations of the inputs and for a permutation $\pi \in \Pi(\{1,\ldots,d\})$, the set $P_i(\pi)$ is defined as the inputs that precede input i in π.

Exact permutation algo. (moderate d) all possible permutations are covered.

Random permutation algo. ($d >> 1$) it randomly samples permutations of the inputs.

In [Song et al., 2016], $\widetilde{\mathbf{val}}(u) \to \widehat{\widetilde{\mathbf{val}}}(u)$.

For each iteration of the loop on the inputs' permutations, the expectation of a conditional variance must be computed.

The cost C of these algorithms is the following:

$$C = N_v + m(d-1)N_0 N_i$$

with N_v the sample size for the variance computation, N_0 the outer loop size for the expectation, N_i the inner loop size for the conditional variance and m the number of permutations according to the selected method.

Bootstrap confidence intervals can be computed. A costly model can be replaced by a metamodel. [Iooss and Prieur, 2019, Benoumechiara and Elie-Dit-Cosaque, 2019]

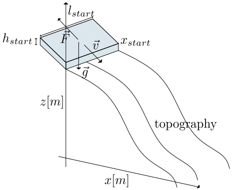
Those algorithms require the ability to sample from the distribution of $\mathbf{X}_u \mid \mathbf{X}_{-u}$, $\forall u \subsetneq \{1, \ldots, d\}$. In [Broto et al., 2020], a given data procedure based on nearest neighbors is introduced.

It is possible to plug algorithms presented in [Castro et al., 2009, Song et al., 2016, Broto et al., 2020] in the estimation of aggregated Shapley effects [Heredia et al., 2020].

Application: snow avalanche modeling

Model based on depth-averaged Saint-Venant equations (see [Heredia et al., 2020] for more details)

$$\begin{aligned} \frac{\partial h}{\partial t} + \frac{\partial hv}{\partial x} &= 0 \\ \frac{\partial hv}{\partial t} + \frac{\partial}{\partial x}\left(hv^2 + \frac{h^2}{2}\right) &= h\left(g \sin\theta - \mathrm{F}\right) \end{aligned}$$



with $v = \|\vec{\mathbf{v}}\|$ the flow velocity, h the flow depth, θ the local angle, t the time, g the gravity constant and $\mathrm{F} = \|\vec{\mathbf{F}}\|$ a frictional force. The model uses the Voellmy frictional force $\mathrm{F} = \mu g cos\theta + g/(\xi h)v^2$, where μ and ξ are friction parameters.

The equations are solved with a finite volumes scheme [Naaim, 1998]. The topography is the one of a path located in Bessans, France.

Objective: better understanding the numerical model.

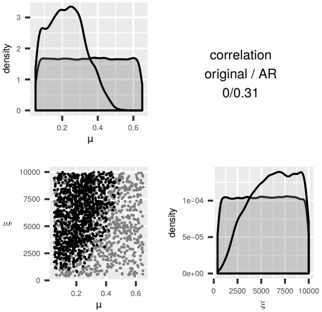
Input	Description	Distribution
μ	Static friction coefficient	$\mathcal{U}[0.05, 0.65]$
ξ	Turbulent friction [m/s^2]	$\mathcal{U}[400, 10000]$
l_{start}	Length of the release zone [m]	$\mathcal{U}[5, 300]$
h_{start}	Mean snow depth in the release zone [m]	$\mathcal{U}[0.05, 3]$
x_{start}	Release abscissa [m]	$\mathcal{U}[0, 1600]$

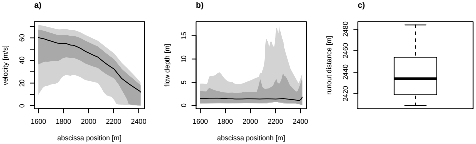
Let's $vol_{start} = l_{start} \times h_{start} \times 72.3/\cos(35°)$ instead of h_{start} and l_{start}.

AR rules:

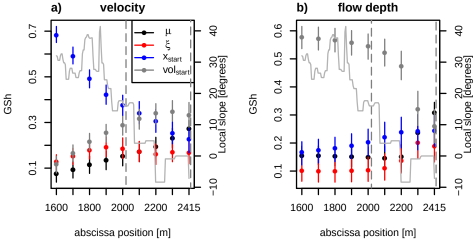
- avalanche simulation is flowing in $[1600m, 2412m]$,
- $vol > 7000m^3$,
- runout distance $< 2500m$ (end of the path).

From $n_0 = 100\,000$ initial runs, we keep $n_1 = 6152$ constrained ones.





Aggregated Shapley effects of velocity and flow depth curves calculated over space intervals $[x, 2412m]$ where $x \in \{1600m, 1700m, \ldots, 2412m\}$



We have $n = 6152$, $N_{\text{tot}} = 2002$, $B = 500$. Effects are estimated using the first (2, *resp.* 4) fPCs [Yao et al., 2005, Ramsay and Silverman, 2005] explaining more than 95% of the variance. Local slope is drawn with a gray line. A gray dotted rectangle is drawn at $[2017m, 2412m]$ where avalanche return periods vary from 10 to 10 000 years.

In summary,

- it is fundamental to have a good approximation of the released volume and abscissa for velocity forecasting, while for flow depth forecasting, a good approximation of released volume is desirable;
- nevertheless, none of the other inputs are negligible.

To outperform the estimation accuracy at the end of the path generating a larger initial sample of avalanches is possible, but the computational burden is prohibitive.

Conclusion, perspectives

Conclusion: Shapley effects present an alternative to allocate parts of variance in the correlated framework. It is possible to define aggregated Shapley indices. There exist algorithms to estimate these indices.

Open questions

- What about goal-oriented Shapley effects? (see recent work in [Da Veiga, 2021])
- Nearest neighbor algorithm depends on many parameters to tune (number of neighbors, total cost...)? Is it possible to propose an adaptive choice of these parameters?
- How can Shapley effects be related to gradient-based measures of sensitivity?
- ...

Some references I

Benoumechiara, N. and Elie-Dit-Cosaque, K. (2019).
Shapley effects for sensitivity analysis with dependent inputs: bootstrap and kriging-based algorithms.
ESAIM: Proceedings and Surveys, 65:266–293.

Broto, B., Bachoc, F., and Depecker, M. (2020).
Variance Reduction for Estimation of Shapley Effects and Adaptation to Unknown Input Distribution.
SIAM/ASA Journal on Uncertainty Quantification, 8(2):693–716.

Castro, J., Gómez, D., and Tejada, J. (2009).
Polynomial calculation of the shapley value based on sampling.
Computers & Operations Research, 36(5):1726–1730.

Da Veiga, S. (2021).
Kernel-based anova decomposition and shapley effects–application to global sensitivity analysis.
arXiv preprint arXiv:2101.05487.

Heredia, M. B., Prieur, C., and Eckert, N. (2020).
Aggregated shapley effects: nearest-neighbor estimation procedure and confidence intervals. application to snow avalanche modeling.
https://hal.inria.fr/hal-02908480.

Some references II

Iooss, B. and Prieur, C. (2019).
Shapley effects for sensitivity analysis with correlated inputs: comparisons with sobol'indices, numerical estimation and applications.
International Journal for Uncertainty Quantification, 9(5).

Lamboni, M., Monod, H., and Makowski, D. (2011).
Multivariate sensitivity analysis to measure global contribution of input factors in dynamic models.
Reliability Engineering and System Safety, 96(4):450–459.

Naaim, M. (1998).
Dense avalanche numerical modeling: interaction between avalanche and structures.
In *25 years of snow avalanche research, Voss, NOR, 12-16 May 1998*, pages 187–191, Norway.

Owen, A. B. (2014).
Sobol'indices and shapley value.
SIAM/ASA Journal on Uncertainty Quantification, 2(1):245–251.

Ramsay, J. O. and Silverman, B. W. (2005).
Functional Data Analysis.
Springer Series in Statistics. Springer, 2nd edition.

Some references III

Shapley, L. S. (1953).
A value for n-person games.
In Kuhn, H. W. and Tucker, A. W., editors, *Contribution to the Theory of Games II (Annals of Mathematics Studies 28)*, pages 307–317. Princeton University Press, Princeton, NJ.

Song, E., Nelson, B., and Staum, J. (2016).
Shapley effects for global sensitivity analysis: Theory and computation.
SIAM/ASA Journal of Uncertainty Quantification, 4:1060–1083.

Yao, F., Müller, H.-G., and Wang, J.-L. (2005).
Functional data analysis for sparse longitudinal data.
Journal of the American Statistical Association, 100(470):577–590.

Thanks for your attention!