Global sensitivity analysis and dimension reduction

Clémentine PRIEUR
Grenoble Alpes University

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

Part III

From variance-based to more general sensitivity indices: I will present recent results on sensitivity analysis targeted to the analysis of models with outputs in general metric spaces.

See, e.g., Da Veiga et al. (2021).

We have seen in Part II that is possible to extend variance based GSA for Y valued in $\mathbb{R}^p$ or in a separable Hilbert space $\mathcal{H}$ (see (Da Veiga et al., 2021, Chapter 3, Section 3.3) and references therein).

In the following, one wants to go to more complex outputs, e.g., considering

$$\mathcal{M} : \quad \mathbb{R}^d \quad \rightarrow \quad \mathcal{Y}$$

with $\mathcal{Y}$ not necessarily a Hilbert space.

Also, one wants to investigate sensitivity beyond variance.

The topic is in full expansion. We refer to (Da Veiga et al., 2021, Chapter 6) for a recent list of references.

Let $Y = \mathcal{M}(X_1, \ldots, X_d) \in \mathbb{R}^p$ be the code output and F be its cumulative distribution function defined as

$$F(t) = \mathbb{P}\left(Y \leqslant t\right) = \mathbb{E}\left[\mathbb{1}_{\{Y \leqslant t\}}\right] = \mathbb{E}\left[Z(t)\right], \; t = (t_1, \ldots, t_p) \in \mathbb{R}^p.$$

Let $F^{\mathbf{u}}(t)$ be the conditional cumulative distribution function of Y conditionally on $X_{\mathbf{u}}$ defined as

$$F^{\mathbf{u}}(t) = \mathbb{P}\left(Y \leqslant t | X_{\mathbf{u}}\right) = \mathbb{E}\left[\mathbb{1}_{\{Y \leqslant t\}} | X_{\mathbf{u}}\right] = \mathbb{E}\left[Z(t) | X_{\mathbf{u}}\right].$$

We perform the Hoeffding decomposition of $Z(t)$:

$$\begin{aligned} Z(t) = \mathbb{1}_{\{Y \leqslant t\}} = & \underbrace{\mathbb{E}\left[Z(t)\right]}_{\text{Mean effect}} \\ & + \underbrace{\left(\mathbb{E}\left[Z(t)|X_{\mathbf{u}}\right] - \mathbb{E}\left[Z(t)\right]\right) + \left(\mathbb{E}\left[Z(t)|X_{-\mathbf{u}}\right] - \mathbb{E}\left[Z(t)\right]\right)}_{\text{First order effects}} \\ & + \underbrace{R(t, \mathbf{u}).}_{\text{Remainder term: higher order effects}} \end{aligned}$$

We then compute the variance of both sides of the previous equation:

$$\mathrm{Var}\left[Z(t)\right] = \mathbb{E}\left[\left(F^{\mathbf{u}}(t) - F(t)\right)^2\right] + \mathbb{E}\left[\left(F^{-\mathbf{u}}(t) - F(t)\right)^2\right] + \mathrm{Var}\left[R(t,\mathbf{u})\right]$$

using orthogonality in the Hoeffding decomposition.
Finally by integrating with respect to the distribution of $Z(t)$ and by normalizing we get:

$$S^{\mathbf{u}}_{2,CVM} := \frac{\int_{\mathbb{R}^p} \mathbb{E}\left[\left(F(t) - F^{\mathbf{u}}(t)\right)^2\right] dF(t)}{\int_{\mathbb{R}^p} F(t)(1 - F(t))dF(t)},$$

involving the Cramér-von Mises distance between the distribution of $Z(t)$ and the one of $Z(t)|X_{\mathbf{u}}$.

Properties of the Cramér-von Mises indices:

1. the different contributions sum to 1;
2. invariance by any translation and by any nondegenerated scaling of the components of Y.

Cramér-von Mises indices have no clear dual formulation, however they can be estimated with a Pick-Freeze scheme.

Other estimation procedures such as U-statistics or rank-based inference (only for scalar inputs and $\mathbf{u}$ a singleton) are also interesting alternatives (see Gamboa et al. (2018)).

Let us consider the more general case where $Y = \mathcal{M}(X_1, \ldots, X_d)$ valued in $\mathcal{Y}$, a general metric space. Let $m \in \mathbb{N}^*$ and $a = (a_i)_{i=1,\ldots,m} \in \mathcal{Y}^m$. We consider the family of test functions

$$\begin{cases} \mathcal{Y}^m \times \mathcal{Y} & \to & \mathbb{R} \\ (a, y) & \mapsto & T_a(y). \end{cases}$$

We assume $T_a(\cdot) \in L^2(\mathbb{P}^{\otimes m} \otimes \mathbb{P})$ with $\mathbb{P}$ the probability distribution of Y.

The general metric space sensitivity index with respect to $\mathbf{u}$, introduced in Fort et al. (2021), is defined as

$$S^{\mathbf{u}}_{2,GMS} := \frac{\int_{\mathcal{Y}^m} \mathbb{E}_{X_{\mathbf{u}}} \left[\left(\mathbb{E}_Y[T_a(Y)] - \mathbb{E}_Y[T_a(Y)|X_{\mathbf{u}}] \right)^2 \right] d\mathbb{P}^{\otimes m}(a)}{\int_{\mathcal{Y}^m} \mathrm{Var}(T_a(Y)) d\mathbb{P}^{\otimes m}(a)}$$

$$= \frac{\int_{\mathcal{Y}^m} \mathrm{Var}\left[\mathbb{E}\left(T_a(Y)|X_{\mathbf{u}} \right) \right] d\mathbb{P}^{\otimes m}(a)}{\int_{\mathcal{Y}^m} \mathrm{Var}(T_a(Y)) d\mathbb{P}^{\otimes m}(a)}.$$

Particular examples:

1. for $\mathcal{Y} = \mathbb{R}$, $m = 0$ and $T_a(y) = y$, one recovers Sobol' indices;
2. for $\mathcal{Y} = \mathbb{R}^p$, $m = 1$ and $T_a(y) = \mathbb{1}_{\{y \leqslant a\}}$, one recovers the index based on the Cramér-von-Mises distance;
3. for $\mathcal{Y} = \mathcal{M}$ a manifold, $m = 2$ and

$$T_a(y) = \mathbb{1}_{y \in \widetilde{B}(a_1, a_2)} = \mathbb{1}_{\|y-(a_1+a_2)/2\| \leqslant \|a_1 - a_2\|/2},$$

   where $\widetilde{B}(a_1, a_2)$ is the ball in $\mathcal{M}$ of diameter $\overline{a_1 a_2}$, one recovers the indices introduced in Fraiman et al. (2020).

General metric space indices can be estimated with either a pick-freeze scheme or U-statistics. For scalar inputs and first-order indices, a rank-based inference procedure is also an alternative.

Principle:

- multiple Monte-Carlo estimation procedure (one to handle the integration part, one to handle the pick-freeze part);
- cost to estimate all first-order indices: $N(m+d+1)$;
- non trivial proof of the CLT using Donsker theorem and the functional delta method (see Fort *et al.*, 2021).

Design of experiments:

- a classical pick-freeze N-sample, that is two N-samples of Y: $(y^{(k)}, y^{\mathbf{u},(k)})$, $1 \leqslant k \leqslant N$;
- m other N-samples of Y independent of $(Y^{(k)}, Y^{\mathbf{u},(k)})_{1\leqslant k\leqslant N}$, namely $w_i^{(k)}$, $1 \leqslant i \leqslant m$, $1 \leqslant k \leqslant N$.

The estimator of the numerator of $S^{\mathbf{u}}_{2,\mathrm{GMS}}$ is then given by

$$\frac{1}{N^m}\sum_{1\leqslant i_1,\dots,i_m\leqslant N}\left\{\left[\frac{1}{N}\sum_{k=1}^{N} T_{w_1^{(i_1)},\cdots,w_m^{(i_m)}}(y^{(k)})\,T_{w_1^{(i_1)},\cdots,w_m^{(i_m)}}(y^{\mathbf{u},(k)})\right]\right.$$
$$\left.-\left[\frac{1}{2N}\sum_{k=1}^{N}\left(T_{w_1^{(i_1)},\cdots,w_m^{(i_m)}}(y^{(k)})+T_{w_1^{(i_1)},\cdots,w_m^{(i_m)}}(y^{\mathbf{u},(k)})\right)\right]^2\right\}$$

while the one of the denominator is

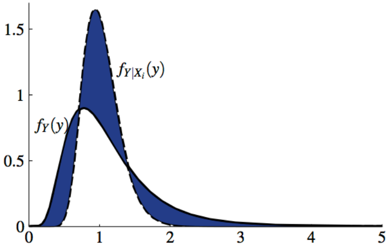
$$\frac{1}{N^m}\sum_{1\leqslant i_1,\dots,i_m\leqslant N}\left\{\frac{1}{2N}\sum_{k=1}^{N}\left[\left(T_{w_1^{(i_1)},\cdots,w_m^{(i_m)}}(y^{(k)})\right)^2+\left(T_{w_1^{(i_1)},\cdots,w_m^{(i_m)}}(y^{\mathbf{u},(k)})\right)^2\right]\right.$$
$$\left.-\left[\frac{1}{2N}\sum_{k=1}^{N}\left(T_{w_1^{(i_1)},\cdots,w_m^{(i_m)}}(y^{(k)})+T_{w_1^{(i_1)},\cdots,w_m^{(i_m)}}(y^{\mathbf{u},(k)})\right)\right]^2\right\}.$$

In Borgonovo (2007), the following index is introduced:

$$\delta_i = \frac{1}{2}\mathbb{E}_{X_i}\left(S_i(X_i)\right) \text{ with } S_i(X_i) = \int \left|p_Y(y) - p_{Y|X_i}(y)\right| dy.$$

Note that $S_i(X_i)$ is the total variation distance between $\mathbb{P}_Y$ and $\mathbb{P}_{Y|X_i}$.

The definition can be generalized as: $S_i(X_i) = \int_{\mathbb{R}} f\left(\frac{p_Y(y)}{p_{Y|X_i}(y)}\right) p_{Y|X_i}(y)dy$ for f any convex function with $f(1) = 0$. E.g., for $f(t) = -\ln(t)$ or $f(t) = t\ln(t)$ one recovers the Kullback-Leibler divergence.



We focus on sensitivity indices based on Maximum Mean Discrepancy introduced in Barr and Rabitz (2022); Da Veiga (2021).

For **P** *and* **Q** *two probability measures defined on $\mathcal{Z}$, we define an Integral Probability Metric between* **P** *and* **Q** *as:*

$$\gamma_{\mathcal{F}}(\mathbf{P}, \mathbf{Q}) = \sup_{f \in \mathcal{F}} \left| \int_{\mathcal{X}} f \, d\mathbf{P} - \int_{\mathcal{X}} f \, d\mathbf{Q} \right|$$

with $\mathcal{F}$ a class of real-valued bounded measurable functions on $\mathcal{Z}$.

Different examples for $\mathcal{F}$ and associated distance:

- bounded continuous functions $\rightarrow$ Dudley metric;
- bounded variation functions $\rightarrow$ Kolmogorov metric;
- Lipschitz bounded functions $\rightarrow$ Wasserstein distance;
- characteristic functions on Borel sets $\rightarrow$ total variation distance.

For **P** *and* **Q** *two probability measures defined on* $\mathcal{Z}$*, we define an Integral Probability Metric between* **P** *and* **Q** *as:*

$$\gamma_{\mathcal{F}}(\mathbf{P},\mathbf{Q}) = \sup_{f\in\mathcal{F}} \left| \int_{\mathcal{X}} f\, d\mathbf{P} - \int_{\mathcal{X}} f\, d\mathbf{Q} \right|$$

with $\mathcal{F}$ *a class of real-valued bounded measurable functions on* $\mathcal{Z}$*.*

Different examples for $\mathcal{F}$ and associated distance:

- bounded continuous functions $\rightarrow$ Dudley metric;
- bounded variation functions $\rightarrow$ Kolmogorov metric;
- Lipschitz bounded functions $\rightarrow$ Wasserstein distance;
- characteristic functions on Borel sets $\rightarrow$ total variation distance.

For $\mathbf{P}$ *and* $\mathbf{Q}$ *two probability measures defined on* $\mathcal{Z}$*, we define an Integral Probability Metric between* $\mathbf{P}$ *and* $\mathbf{Q}$ *as:*

$$\gamma_{\mathcal{F}}(\mathbf{P}, \mathbf{Q}) = \sup_{f \in \mathcal{F}} \left| \int_{\mathcal{X}} f \, d\mathbf{P} - \int_{\mathcal{X}} f \, d\mathbf{Q} \right|$$

with $\mathcal{F}$ *a class of real-valued bounded measurable functions on* $\mathcal{Z}$*.*

Different examples for $\mathcal{F}$ and associated distance:

- bounded continuous functions $\rightarrow$ Dudley metric;
- bounded variation functions $\rightarrow$ Kolmogorov metric;
- Lipschitz bounded functions $\rightarrow$ Wasserstein distance;
- characteristic functions on Borel sets $\rightarrow$ total variation distance.

For $\mathbf{P}$ *and* $\mathbf{Q}$ *two probability measures defined on* $\mathcal{Z}$*, we define an Integral Probability Metric between* $\mathbf{P}$ *and* $\mathbf{Q}$ *as:*

$$\gamma_{\mathcal{F}}(\mathbf{P},\mathbf{Q}) = \sup_{f \in \mathcal{F}} \left| \int_{\mathcal{X}} f \, d\mathbf{P} - \int_{\mathcal{X}} f \, d\mathbf{Q} \right|$$

with $\mathcal{F}$ *a class of real-valued bounded measurable functions on* $\mathcal{Z}$*.*

Different examples for $\mathcal{F}$ and associated distance:

- bounded continuous functions $\to$ Dudley metric;
- bounded variation functions $\to$ Kolmogorov metric;
- Lipschitz bounded functions $\to$ Wasserstein distance;
- characteristic functions on Borel sets $\to$ total variation distance.

For $\mathbf{P}$ *and* $\mathbf{Q}$ *two probability measures defined on* $\mathcal{Z}$*, we define an Integral Probability Metric between* $\mathbf{P}$ *and* $\mathbf{Q}$ *as:*

$$\gamma_{\mathcal{F}}(\mathbf{P},\mathbf{Q}) = \sup_{f\in\mathcal{F}} \left| \int_{\mathcal{X}} f\, d\mathbf{P} - \int_{\mathcal{X}} f\, d\mathbf{Q} \right|$$

with $\mathcal{F}$ *a class of real-valued bounded measurable functions on* $\mathcal{Z}$*.*

Different examples for $\mathcal{F}$ and associated distance:

- bounded continuous functions $\rightarrow$ Dudley metric;
- bounded variation functions $\rightarrow$ Kolmogorov metric;
- Lipschitz bounded functions $\rightarrow$ Wasserstein distance;
- characteristic functions on Borel sets $\rightarrow$ total variation distance.

Let $\mathcal{Z}$ be an arbitrary set and $\mathcal{H}$ a Hilbert space of real-valued functions $f : \mathcal{Z} \to \mathbb{R}$ on $\mathcal{Z}$ with inner product $< \cdot, \cdot >_{\mathcal{H}}$. For every $z \in \mathcal{Z}$, we define the evaluation functional $L_z : \mathcal{H} \to \mathbb{R}$ as $f \mapsto L_z(f) = f(z)$.

A Hilbert space $\mathcal{H}$ is a reproducing kernel Hilbert space (RKHS) if the evaluation functionals are continuous.

A RKHS $\mathcal{H}$ is associated to a function $k : \mathcal{Z} \times \mathcal{Z} \to \mathbb{R}$ such that

- for all $z \in \mathcal{Z}$, $k(z, \cdot) \in \mathcal{H}$;
- for all $f \in \mathcal{H}$ and for all $z \in \mathcal{Z}$, $< f, k(z, \cdot) >_{\mathcal{H}} = f(z)$.

The kernel mean embedding $\mu_{\mathbf{Q}} \in \mathcal{H}$ of a probability distribution $\mathbf{Q}$ on $\mathcal{Z}$ is given by

$$\mu_{\mathbf{Q}} = \mathbb{E}_{\zeta \sim \mathbf{Q}} k_{\mathcal{Z}}(\zeta, \cdot) = \int_{\mathcal{Z}} k_{\mathcal{Z}}(\zeta, \cdot) d\mathbf{Q}(\zeta).$$

Let $\mathcal{M}_1^+(\mathcal{Z})$ the set of probability measures on $\mathcal{Z}$.

A kernel is characteristic if the kernel embedding $\mu : \mathcal{M}_1^+(\mathcal{Z}) \to \mathcal{H}$ *is injective.*

Choosing $\mathcal{F}$ (in the definition of $\gamma_{\mathcal{F}}$) as the space of functions in the unit ball of a characteristic RKHS leads to a MMD distance.

Due to the definition of kernel embedding we get

$$\text{MMD}^2(P, Q; \mathcal{H}) = \|\mu_{\mathbf{P}} - \mu_{\mathbf{Q}}\|_{\mathcal{H}}^2.$$

How can we use MMD for GSA?

Let $\mathcal{H}$ be a RKHS on $\mathcal{Y}$.
We define

$$S_j^{\mathcal{H},\text{unnorm}} = \mathbb{E}_{X_j}\text{MMD}^2(\mathbf{P}_Y, \mathbf{P}_{Y|X_j}; \mathcal{H}).$$

Let ζ_1, ζ_1' iid $\sim P_Y$, independent of ζ_2, ζ_2' iid $\sim P_{Y|X_j}$.

Moreover, due to the reproducing property and kernel embedding it is possible to prove that:

$$S_j^{\mathcal{H},\text{unnorm}} = \mathbb{E}_{X_j}\mathbb{E}_{\zeta_2,\zeta_2'} k_{\mathcal{Y}}(\zeta_2,\zeta_2') - \mathbb{E}_{\zeta_1,\zeta_1'} k_{\mathcal{Y}}(\zeta_1,\zeta_1').$$

Examples :

For $\mathcal{Y} \subset \mathbb{R}$ and $k_{\mathcal{Y}}(y,y') = yy'$ (not a characteristic kernel), one recovers first-order Sobol' index S_j.

For $\mathcal{Y}$ a compact set, one has from Mercer's theorem that $k_{\mathcal{Y}}(y,y') = \sum_{r=1}^{+\infty} \Phi_r(y)\Phi_r(y')$ with $\{\Phi_r\,,\ r \geq 1\}$ orthogonal functions in $\mathbb{L}^2(\mathcal{Y})$. Then

$$S_j^{\mathcal{H},\text{unnorm}} = \sum_{r=1}^{+\infty} \mathrm{Var}[\mathbb{E}(\Phi_r(Y)|X_j)].$$

Coming back to the formulation

$$S_j^{\mathcal{H},\text{unnorm}} = \mathbb{E}_{X_j}\mathbb{E}_{\zeta_2,\zeta_2'} k_{\mathcal{Y}}(\zeta_2, \zeta_2') - \mathbb{E}_{\zeta_1,\zeta_1'} k_{\mathcal{Y}}(\zeta_1, \zeta_1'),$$

one can propose a pick-freeze scheme to estimate $S_j^{\mathcal{H},\text{unnorm}}$:

$$\widehat{S}_j^{\mathcal{H},\text{unnorm}} = \frac{1}{N}\sum_{i=1}^{N}\Big(k_{\mathcal{Y}}(y^i, y^{j,i}) - k_{\mathcal{Y}}(y^i, y'^i)\Big)$$

with $y^i = \mathcal{M}(\mathbf{x}^{1,i})$, $y^{j,i} = \mathcal{M}(x_1^{2,i}, \ldots, x_{j-1}^{2,i}, x_j^{1,i}, x_{j+1}^{2,i}, \ldots, x_d^{2,i})$ as previously and $y'^i = \mathcal{M}(x^{2,i})$.

We define the normalizing constant $\mathrm{MMD}^2_{\mathrm{tot};\mathcal{H}}$ as

$$\mathrm{MMD}^2_{\mathrm{tot};\mathcal{H}} = \sum_{A \subseteq \{1,\ldots,d\}} \mathrm{MMD}^2_{A;\mathcal{H}}$$

with $\mathrm{MMD}_{A;\mathcal{H}}$ defined as

$$\mathrm{MMD}^2_{A;\mathcal{H}} = \sum_{B \subseteq A} (-1)^{|A|-|B|} \mathbb{E}_{X_B} \mathrm{MMD}^2(\mathrm{P}_Y, \mathrm{P}_{Y|B}; \mathcal{H}).$$

This normalizing constant will lead to an ANOVA-like decomposition. Then we define

$$S_j^{\mathcal{H}} = \frac{S_j^{\mathcal{H},\mathrm{unnorm}}}{\mathrm{MMD}^2_{\mathrm{tot};\mathcal{H}}}.$$

It is possible to define MMD indices of any order and total MMD indices.

We assume that for any $\mathbf{x} = (x_1, \ldots, x_d) \in \mathcal{X} = \mathcal{X}_1 \times \ldots \mathcal{X}_d$, $\mathcal{M}(\mathbf{x})$ is a probability distribution on $\mathbb{R}$, denoted by $\mu_{\mathbf{x}}$.

Let's assume that for any $\mathbf{x}$, the probability measure $\mu_{\mathbf{x}}$ belongs to $\mathcal{Y} = \mathcal{W}_2(\mathbb{R})$ the space of all probability distributions on $\mathbb{R}$ with finite second-order moment w.r.t. the 2-Wasserstein distance W_2. We consider the r.v. $\mu_{\mathbf{X}}$ with values in $\mathcal{Y}$. We denote by $\mathbb{P}$ its probability distribution.

Let $\tilde{\mu}$ and $\tilde{\tilde{\mu}}$ be two elements in $\mathcal{W}_2(\mathbb{R})$. The general metric space indices in this framework $S^{\mathbf{u}}_{2,W_2}$ can be defined as in (Fort *et al.*, 2021):

$$\frac{\int_{\mathcal{W}_2(\mathbb{R})\times\mathcal{W}_2(\mathbb{R})} \mathrm{Var}\left[\mathbb{E}\left(\mathbb{1}_{W_2(\tilde{\mu},\mu_{\mathbf{X}})\le W_2(\tilde{\mu},\tilde{\tilde{\mu}})} | X_{\mathbf{u}}\right)\right] d\mathbb{P}^{\otimes 2}(\tilde{\mu},\tilde{\tilde{\mu}})}{\int_{\mathcal{W}_2(\mathbb{R})\times\mathcal{W}_2(\mathbb{R})} \mathrm{Var}(\mathbb{1}_{W_2(\tilde{\mu},\mu_{\mathbf{X}})\le W_2(\tilde{\mu},\tilde{\tilde{\mu}})}) d\mathbb{P}^{\otimes 2}(\tilde{\mu},\tilde{\tilde{\mu}})}.$$

In practice one can only obtain an empirical approximation of the measure $\mu_{\mathbf{x}}$ computed from n evaluations $\mathcal{M}(\mathbf{x}, d^{(j)})$, $j = 1, \ldots, n$. Note that in general, the $d^{(j)}$ are not observed.

Finally, the general design of experiments is the following:

$$\mathbf{x}^{(1)}, d^{(1,1)}, \ldots, d^{(1,n)} \longrightarrow \mathcal{M}(\mathbf{x}^{(1)}, d^{(1,1)}), \ldots, \mathcal{M}(\mathbf{x}^{(1)}, d^{(1,n)})$$

$$\ldots$$

$$\mathbf{x}^{(N)}, d^{(N,1)}, \ldots, d^{(N,n)} \longrightarrow \mathcal{M}(\mathbf{x}^{(N)}, d^{(N,1)}), \ldots, \mathcal{M}(\mathbf{x}^{(N)}, d^{(N,n)})$$

For any $k = 1, \ldots, N$, we define the approximations of $\mu_{\mathbf{x}^{(j)}}$ as:

$$\widehat{\mu}_{\mathbf{x}^{(k)}} = \frac{1}{n} \sum_{j=1}^{n} \delta_{\mathcal{M}(\mathbf{x}^{(k)}, d^{(k,j)})} .$$

Then the indices $S_{2,W_2}^{\mathbf{u}}$ can be estimated either with a pick-freeze scheme, either with U-statistics or with a rank-based approach (for $\mathbf{u}$ a singleton and for scalar inputs).

Pick-freeze estimation procedure

1. Generate two samples $\mathbf{x}^{(k)}, d^{(k,j)}$ and $\mathbf{x}'^{(k)}, d'^{(k,j)}$, $k = 1, \dots, N$, $j = 1, \dots, n$.
2. Generate a pick-freeze sample of size N: $\left(\mathbf{x}^{(k)}, \mathbf{x}^{\mathbf{u},(k)}\right) = \left(\mathbf{x}^{(k)}, \mathbf{x}_{\mathbf{u}}^{(k)} : \mathbf{x}'^{(k)}_{-\mathbf{u}}\right)$, $k = 1, \dots, N$.
3. For each input, compute the corresponding output n times: $\mathcal{M}(\mathbf{x}^{(k)}, d^{(k,j)})\,,\ \mathcal{M}(\mathbf{x}^{\mathbf{u},(k)}, d'^{(k,j)})\,,\ k = 1, \dots, N\,,\ j = 1, \dots, n\,.$
4. Approximate the measures by empirical measures:
$$\mu^{(k)} \approx \widehat{\mu}^{(k)} = \frac{1}{n} \sum_{j=1}^{n} \delta_{\mathcal{M}(\mathbf{x}^{(k)}, d^{(k,j)})},$$
$$\mu^{\mathbf{u},(k)} \approx \widehat{\mu}^{\mathbf{u},(k)} = \frac{1}{n} \sum_{j=1}^{n} \delta_{\mathcal{M}(\mathbf{x}^{\mathbf{u},(k)}, d'^{(k,j)})}.$$
5. We also need two additional samples of the output, indendent from the pick-freeze scheme: $\mathcal{M}(\tilde{\mathbf{x}}^{(k)}, \tilde{d}^{(k,j)})\,,\ \mathcal{M}(\tilde{\tilde{\mathbf{x}}}^{(k)}, \tilde{\tilde{d}}^{(k,j)})\,,\ k = 1, \dots, N\,,\ j = 1, \dots, n$ leading to $\widehat{\tilde{\mu}}^{(k)}, \widehat{\tilde{\tilde{\mu}}}^{(k)}$, $k = 1, \dots, N$.

Pick-freeze estimation procedure

The cost in terms of number of evaluations of $\mathcal{M}$ is $4Nn$.
In order to compute explicitly our estimator, it remains to compute terms of the form:

$$W_2(\widehat{\mu}^{(\ell)}, \widehat{\mu}^{(k)}).$$

The quantity $W_2(\nu_1, \nu_2)$ is easy to compute if ν_1 and ν_2 are two discrete measures on $\mathbb{R}$ supported on a same number of points. Namely, for

$$\nu_1 = \frac{1}{n}\sum_{k=1}^{n} \delta_{a_k},\ \nu_2 = \frac{1}{n}\sum_{k=1}^{n} \delta_{b_k},$$

the Wasserstein distance between ν_1 and ν_2 simply writes

$$W_2^2(\nu_1, \nu_2) = \frac{1}{n}\sum_{k=1}^{n}(a_{(k)} - b_{(k)})^2,$$

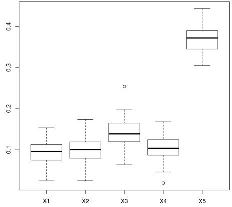
where $z_{(k)}$ is the k-th order statistics of z.

Illustration on a toy model Let us define the stochastic simulator (see Da Veiga (2021); Moutoussamy et al. (2015)) as

$$Y = (X_1 + 2X_2 + U_1)\sin(3X_3 - 4X_4 + G) + U_2 + 5X_5\,B + \sum_{i=1}^{5} i\,X_i$$

where the intrisic noise is modeled by $U_1 \sim \mathcal{U}([0,1])$, $U_2 \sim \mathcal{U}([1,2])$, $G \sim \mathcal{N}(0,1)$ and $B \sim$ Bernoulli$(1/2)$, and the uncertain parameters X_i are uniformly distributed on $[0,1]$.

With Sébastien's code we compute, for each input X_i, 50 independent realizations of the pick-freeze estimator of S^i_{2,W_2} with $N = 200$ and $n = 100$.



We are interested in models driven by parametrized stochastic differential equations (parametrized SDE) defined by:

$$dZ_t = b(Z_t, x)dt + \sigma(Z_t, x)dW_t, \quad Z_0 = z \tag{1}$$

on $\mathbb{R}^s$. We assume that the uncertain parameter is the realization of a random vector $X \in \mathbb{R}^d$, independent from the Brownian motion driving the SDE.
Assume there exists a unique stationary solution to (1) absolutely continuous with respect to Lebesgue. The density $p_s(\cdot, x)$ of the stationary solution is obtained by solving Fokker-Planck equation

$$\mathcal{L}^* p_s(x, \xi) = 0 \; \forall \, x \in \mathbb{R}^d$$

$$\int_{\mathbb{R}^d} p_s(x, \xi) dx = 1$$

$$p_s(x, \xi) \geq 0 \; \forall \, x \in \mathbb{R}^d \, , \quad \lim_{|x| \to +\infty} p_s(x, \xi) = 0$$

with $\mathcal{L}^*$ the adjoint of the infinitesimal generator associated to (1).

We propose to compute first-order and total MMD indices, with $\mathcal{Y} = \mathcal{M}_1^+(\mathbb{R}^s)$.

One chooses a kernel $k_{\mathcal{M}_1^+(\mathbb{R}^s)}$ defined by

$$k_{\mathcal{M}_1^+(\mathbb{R}^s)}(y, y') = \sigma^2 \exp\left(-\lambda \, \mathrm{MMD}^2(y, y'; \mathcal{G})\right), \quad y, y' \in \mathcal{M}_1^+(\mathbb{R}^s),$$

with $\sigma^2, \lambda > 0$ and RKHS $\mathcal{G}$ of functions $\mathbb{R}^s \to \mathbb{R}$ associated to kernel $k_{\mathbb{R}^s}$ defined by

$$k_{\mathbb{R}^s}(z, z') = \exp\left(\frac{-1}{2\,\tau^2}|z - z'|^2\right), \quad z, z' \in \mathbb{R}^s, \tag{2}$$

with $\tau^2 > 0$.

There remain challenges such as:

- the kernel choice;
- to speed up the computation with apropriate metamodels (e.g., stochastic Galerkin scheme for the last example, ongoing work);
- …

Research on this topic is very active, and in relation with recent developments in machine learning.

Some references

Barr, J. and Rabitz, H. (2022). A generalized kernel method for global sensitivity analysis. *SIAM/ASA Journal on Uncertainty Quantification*, 10(1):27–54.

Borgonovo, E. (2007). A new uncertainty importance measure. *Reliability Engineering and System Safety*, 92(6):771–784.

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

Da Veiga, S., Gamboa, F., Iooss, B., Prieur, C., for Industrial, S., and Mathematics, A. (2021). *Basics and Trends in Sensitivity Analysis: Theory and Practice in R*. Computational science and engineering. Society for Industrial and Applied Mathematics.

Fort, J.-C., Klein, T., and Lagnoux, A. (2021). Global sensitivity analysis and wasserstein spaces. *SIAM/ASA Journal on Uncertainty Quantification*, 9(2):880–921.

Fraiman, R., Gamboa, F., and Moreno, L. (2020). Sensitivity indices for output on a riemannian manifold. *International Journal for Uncertainty Quantification*, 10(4).

Gamboa, F., Klein, T., and Lagnoux, A. (2018). Sensitivity analysis based on cramér–von mises distance. *SIAM/ASA Journal on Uncertainty Quantification*, 6(2):522–548.

Moutoussamy, V., Nanty, S., and Pauwels, B. (2015). Emulators for stochastic simulation codes. *ESAIM: Proceedings and Surveys*, 48:116–155.

Thanks for your attention!

And a little bit of advertisement