Introduction to Global Sensibility Analysis

Model exploration for approximation of complex,
high-dimensional problems

MSIAM

inputs $\longrightarrow$ model $\longrightarrow$ output

$$\left.\begin{array}{c} X_1 \\ \cdot \\ \cdot \\ \cdot \\ X_d \end{array}\right\} \quad \text{---}\ \mathcal{M}\ \text{---} \quad Y = \mathcal{M}(X_1, \ldots, X_d)$$

Experimental design:
planification, sampling

$\longleftrightarrow$

Sensitivity analysis:
sensitivity indices' inference

Background :

$$\mathcal{M} : \begin{cases} \mathbb{R}^d & \rightarrow & \mathbb{R} \\ \mathbf{x} & \mapsto & y = \mathcal{M}(x_1, \ldots, x_d) \end{cases}$$

Goal : find how model outputs vary with inputs changes.

Different strategies :

- Qualitative analysis : non-linear behaviors? possible interactions?
  ex. : screening .
- Quantitative analysis : factorial hierarchisation, statistical tests H_0 "negligible input"
  ex. : sensitivity Sobol' indices

Sensitivity analysis may help identifying inappropriate models.

Background :

$$\mathcal{M} : \begin{cases} \mathbb{R}^d & \rightarrow & \mathbb{R} \\ \mathbf{x} & \mapsto & y = \mathcal{M}(x_1, \ldots, x_d) \end{cases}$$

Goal : find how model outputs vary with inputs changes.

Different strategies :

- Qualitative analysis : non-linear behaviors? possible interactions?
  ex. : screening .
- Quantitative analysis : factorial hierarchisation, statistical tests H_0 "negligible input"
  ex. : sensitivity Sobol' indices

Sensitivity analysis may help identifying inappropriate models.

Background :

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

Goal : find how model outputs vary with inputs changes.

Different strategies :

- Qualitative analysis : non-linear behaviors? possible interactions?
  ex. : screening .
- Quantitative analysis : factorial hierarchisation, statistical tests H_0 "negligible input"
  ex. : sensitivity Sobol' indices

Sensitivity analysis may help identifying inappropriate models.

Various approaches for quantitative sensitivity :

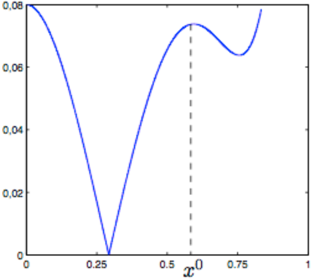
Local approaches :

$\mathcal{M}(\mathbf{x}) \approx \mathcal{M}(\mathbf{x^0}) + \sum_{i=1}^{d} \left(\frac{\partial \mathcal{M}}{\partial x_i}\right)_{\mathbf{x^0}} (x_i - x_i^0)$ (Taylor approximation).

First order sensitivity index for input i : $\left(\frac{\partial \mathcal{M}}{\partial x_i}\right)_{\mathbf{x^0}}$.

Pros : Low computational cost even for large d

Cons : local approaches, not well-suited for highly nonlinear models



Global approaches :

From expert knowledge or observations, we attribute a probability law to the inputs vector.

ex.: If independent inputs, then only margins are needed.

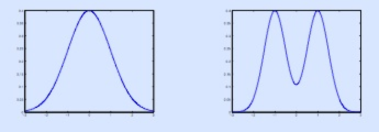


Figure: law (left) unimodal , (right) bimodal

We vary inputs w.r.t. their probability distribution.

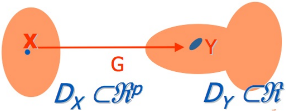


Figure: Local versus Global ($G := \mathcal{M}$)

We vary input w.r.t. their probability law

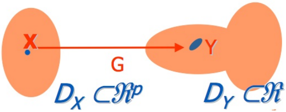
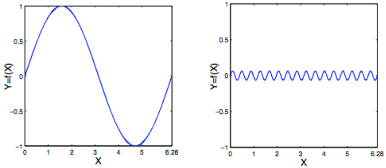


Figure: Local versus Global ($G := \mathcal{M}$), illustration.

"Globalized" local approaches : e.g. (1) $\mathbb{E}_X \left[\frac{\partial \mathcal{M}}{\partial x_i} \Big|_{\mathbf{x}} \right]$, ou (2) $\mathbb{E}_X \left[\left(\frac{\partial \mathcal{M}}{\partial x_i} \Big|_{\mathbf{x}} \right)^2 \right]$.

Avantages : particularly interesting if adjoint available

Cons :
(1) does not discriminate enc



(2) is known as Derivative-based Global Sensitivity Measures, see Sobol' & Gresham (1995), Sobol' & Kucherenko (2009).

This index is more adapted for screening than for hierarchization (e.g. Lamboni *et al.*, 2013).

This lecture targets global approaches that allow to efficiently rank input factors.

Sensitivity measures, definition, estimation

I- Measures based on linear regressions

II- Functional variance analysis

III- Sobol indices inference

- Monte Carlo techniques,
- Spectral techniques

IV- *Distributional* indices

V- Further topics

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

- Linear correlation

$$\rho_i = \rho(X_i, Y) = \frac{\mathrm{Cov}(X_i, Y)}{\sqrt{\mathrm{Var}(X_i)}\sqrt{\mathrm{Var}(Y)}}$$

- Partial correlation

$$PCC_i = PCC(X_i, Y) = \rho\left(Y - \widehat{Y}^{-(i)}, X_i - \widehat{X_i}^{-(i)}\right)$$

Remarks :

• if $Y = \sum_{i=1}^{d} \beta_i X_i$, and if inputs are independent , $\sum_{i=1}^{d} \rho^2(X_i, Y) = 1$;

• if inputs are correlated , the PCC are more suitable.

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

- Linear correlation

$$\rho_i = \rho\left(X_i, Y\right) = \frac{\mathrm{Cov}(X_i, Y)}{\sqrt{\mathrm{Var}(X_i)}\sqrt{\mathrm{Var}(Y)}}$$

- Partial correlation

$$PCC_i = PCC\left(X_i, Y\right) = \rho\left(Y - \widehat{Y}^{-(i)}, X_i - \widehat{X_i}^{-(i)}\right)$$

Remarks :

- if $Y = \sum_{i=1}^{d} \beta_i X_i$, and if inputs are independent , $\sum_{i=1}^{d} \rho^2\left(X_i, Y\right) = 1$;
- if inputs are correlated , the PCC are more suitable.

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

- Linear correlation

$$\rho_i = \rho(X_i, Y) = \frac{\mathrm{Cov}(X_i, Y)}{\sqrt{\mathrm{Var}(X_i)}\sqrt{\mathrm{Var}(Y)}}$$

- Partial correlation

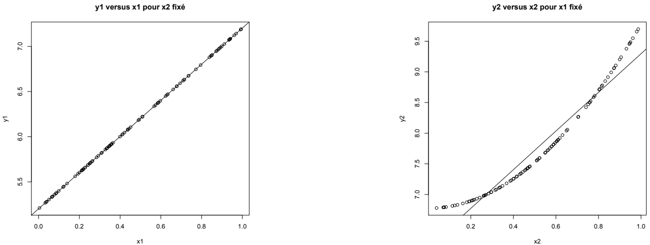
$$PCC_i = PCC(X_i, Y) = \rho\left(Y - \widehat{Y}^{-(i)}, X_i - \widehat{X_i}^{-(i)}\right)$$

Remarks :

- if $Y = \sum_{i=1}^{d} \beta_i X_i$, and if inputs are independent , $\sum_{i=1}^{d} \rho^2(X_i, Y) = 1$;
- if inputs are correlated , the PCC are more suitable.

Assessment of linear model?

Toy example : $Y = 2X_1 + 3X_2{}^2 + 5$, $X_i \sim \mathcal{U}([0,1])$, $i = 1, 2$, $X_1 \perp\!\!\!\perp X_2$.



We can approximate this model by a linear model :
$Y = \beta_1 X_1 + \beta_2 X_2 + \beta_0 + \varepsilon$, $\varepsilon \sim \mathcal{N}(0, \sigma^2)$.

Learning sample : $y_k = \mathcal{M}(x_{1,k}, \ldots, x_{d,k})$, $k = 1, \ldots, 100$

$\Rightarrow \quad \hat{y} = \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \hat{\beta}_0 = 2.06x1 + 3.15x2 + 4.34.$

Which measure to assess the fit of this model?

⋆ Coefficient R^2

$$R^2 = \frac{SCE}{SCT} = \frac{\sum_{k=1}^{m}\left(\widehat{y}_k - \overline{y}\right)^2}{\sum_{k=1}^{m}\left(y_k - \overline{y}\right)^2},$$

$\widehat{y}_k = \sum_{i=1}^{d} \hat{\beta}_i x_{i,k}$, $\overline{y} = \frac{1}{m}\sum_{k=1}^{m} y_k$.

⋆ Prediction error, e.g. cross-validation :

$$\frac{1}{m}\frac{\sum_{k=1}^{m}\left(\widehat{y}_k^{-(k)} - y_k\right)^2}{\frac{1}{m}\sum_{k=1}^{m}\left(y_k - \overline{y}\right)^2},$$

$\widehat{y}_k^{-(k)} = \sum_{i=1}^{d} \widehat{\beta}_i^{-(k)} x_{i,k}$, $\widehat{\beta}_i^{-(k)}$ inferred from

$$(y_j, \mathbf{x}_j),\ j = 1, \ldots, k-1, k+1, \ldots, m.$$

⋆ Coefficient R^2

$$R^2 = \frac{SCE}{SCT} = \frac{\sum_{k=1}^{m}\left(\widehat{y}_k - \overline{y}\right)^2}{\sum_{k=1}^{m}\left(y_k - \overline{y}\right)^2},$$

$\widehat{y}_k = \sum_{i=1}^{d} \hat{\beta}_i x_{i,k}$, $\overline{y} = \frac{1}{m}\sum_{k=1}^{m} y_k$.

⋆ Prediction error, e.g. cross-validation :

$$\frac{1}{m}\frac{\sum_{k=1}^{m}\left(\widehat{y}_k^{-(k)} - y_k\right)^2}{\frac{1}{m}\sum_{k=1}^{m}\left(y_k - \overline{y}\right)^2},$$

$\widehat{y}_k^{-(k)} = \sum_{i=1}^{d} \widehat{\beta}_i^{-(k)} x_{i,k}$, $\widehat{\beta}_i^{-(k)}$ inferred from

$$(y_j, \mathbf{x}_j),\ j = 1, \dots, k-1, k+1, \dots, m.$$

If the relationship input/output is no more linear but simply monotonic, we work with ranks.

y_k, $x_{i,k}$, $k = 1, \dots, m$, $i = 1, \dots, d$

$r_{i,k}$ rank of $x_{i,k}$ in $(x_{i,1}, \dots, x_{i,m})$, r_k rank of y_k in $(y_1, \dots, y_m)$

- $\rho_i^S = \frac{\sum_{k=1}^m (r_{i,k} - \bar{r}_i)(r_k - \bar{r})}{\sqrt{\sum_{k=1}^m (r_{i,k} - \bar{r}_i)^2} \sqrt{\sum_{k=1}^m (r_k - \bar{r})^2}}$

- idem for pcc_i

ANVOA basics

Y quantity of interest, X_1 (*resp.* X_2) qualitative factor with I (*resp.* J) levels.

model : $Y_{ij} = \mu + \alpha_i + \beta_j + \gamma_{i,j} + \varepsilon_{ij}$, ε_{ij} i.i.d. $\mathcal{N}(0, \sigma^2)$.

Identifiability constraints : $\sum_{i=1}^{I} \alpha_i = 0$, $\sum_{j=1}^{J} \beta_j = 0$, $\sum_{i=1}^{I} \gamma_{ij} = 0$, $\sum_{j=1}^{J} \gamma_{ij} = 0$.

Effect inference :

Complete and balanced design with $r > 1$ replica $\rightarrow y_{ijk}$, $k = 1, \ldots, r$

$$\hat{\mu} = \overline{y}, \quad \hat{\alpha}_i = \overline{y}_{i..} - \overline{y}, \quad \hat{\beta}_j = \overline{y}_{.j.} - \overline{y}, \quad \hat{\gamma}_{ij} = \overline{y}_{ij.} - \overline{y}_{i..} - \overline{y}_{.j.} + \overline{y},$$

with usual notation $\overline{y} = \frac{1}{IJr} \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{r} y_{ijk}$,

$$\overline{y}_{i..} = \frac{1}{Jr} \sum_{j=1}^{J} \sum_{k=1}^{r} y_{ijk}, \quad \overline{y}_{.j.} = \frac{1}{Ir} \sum_{i=1}^{I} \sum_{k=1}^{r} y_{ijk}, \quad \overline{y}_{ij.} = \frac{1}{r} \sum_{k=1}^{r} y_{ijk}.$$

In the previous model, we define :

- the forecasts $\hat{y}_{ijk} = \overline{y}_{ij\cdot}$,
- the residuals $\hat{\varepsilon}_{ijk} = y_{ijk} - \hat{y}_{ijk} = y_{ijk} - \overline{y}_{ij\cdot}$.

Variance decomposition :

$$\begin{array}{lclcl} SCT & = & SCM & + & SCR \\ \text{total variance} & = & \text{variance explained by the model} & + & \text{residual variance} \end{array}$$

$$SCM = SCX_1 + SCX_2 + SCX_1X_2 \qquad \text{with}$$

$$SCX_1 = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{r}\left(\overline{y}_{i\cdot\cdot} - \overline{y}\right)^2, \qquad SCX_2 = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{r}\left(\overline{y}_{\cdot j\cdot} - \overline{y}\right)^2,$$

$$SCX_1X_2 = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{r}\left(\overline{y}_{ij\cdot} - \overline{y}_{i\cdot\cdot} - \overline{y}_{\cdot j\cdot} + \overline{y}\right)^2,$$

$$SCR = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{r}\left(\hat{y}_{ijk} - y_{ijk}\right)^2 = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{r}\left(\overline{y}_{ij\cdot} - y_{ijk}\right)^2.$$

In the previous model, we define :

- the forecasts $\hat{y}_{ijk} = \overline{y}_{ij\cdot}$,
- the residuals $\hat{\varepsilon}_{ijk} = y_{ijk} - \hat{y}_{ijk} = y_{ijk} - \overline{y}_{ij\cdot}$.

Variance decomposition :

SCT	=	*SCM*	+	*SCR*
total variance	=	variance explained by the model	+	residual variance

$$SCM = SCX_1 + SCX_2 + SCX_1X_2 \qquad \text{with}$$

$$SCX_1 = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{r}(\overline{y}_{i\cdot\cdot} - \overline{y})^2, \qquad SCX_2 = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{r}(\overline{y}_{\cdot j\cdot} - \overline{y})^2,$$

$$SCX_1X_2 = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{r}(\overline{y}_{ij\cdot} - \overline{y}_{i\cdot\cdot} - \overline{y}_{\cdot j\cdot} + \overline{y})^2,$$

$$SCR = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{r}(\hat{y}_{ijk} - y_{ijk})^2 = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{r}(\overline{y}_{ij\cdot} - y_{ijk})^2.$$

Hypothesis testing: many possible tests

ex. : H_0 : additive model versus H_1 : complete model

Test statistic :

$$T = \frac{SCX_1X_2/(IJ-I-J+1)}{SCR/(IJr-IJ)} \underset{H_0}{\sim} F(IJ-I-J+1, IJr-IJ).$$

ANOVA assumptions :

• the factors only impact the mean of the quantitative variable Y, but not its variance;

• All other variations are Gaussian and independent

Functional framework : (Antoniadis, 1984)

$$Y(s,t) = \mathcal{M}(s,t) + \varepsilon(s,t)\ , \quad (s,t) \in S \times T$$

with

- $\varepsilon(s,t)$ zero-mean Gaussian process with covariance $K(s,t)$,
- S and T two metric compacts spaces

More general setup : (Hoeffding, 1948; Sobol', 1993)

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

In the following, we assume :

i) the X_i are independent ;

ii) $\forall\, i = 1, \ldots, d$, $X_i \sim \mathcal{U}([0,1])$.

Functional framework : (Antoniadis, 1984)

$$Y(s,t) = \mathcal{M}(s,t) + \varepsilon(s,t)\ , \quad (s,t) \in S \times T$$

with

- $\varepsilon(s,t)$ zero-mean Gaussian process with covariance $K(s,t)$,
- S and T two metric compacts spaces

More general setup : (Hoeffding, 1948; Sobol', 1993)

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

In the following, we assume :

i) the X_i are independent ;

ii) $\forall i = 1, \ldots, d$, $X_i \sim \mathcal{U}([0,1])$.

Functional framework : (Antoniadis, 1984)

$$Y(s,t) = \mathcal{M}(s,t) + \varepsilon(s,t)\ , \quad (s,t) \in S \times T$$

with

- $\varepsilon(s,t)$ zero-mean Gaussian process with covariance $K(s,t)$,
- S and T two metric compacts spaces

More general setup : (Hoeffding, 1948; Sobol', 1993)

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

In the following, we assume :

i) the X_i are independent ;

ii) $\forall\, i = 1, \ldots, d$, $X_i \sim \mathcal{U}\left([0,1]\right)$.

In the following, we assume :

i) the X_i are independent ;

ii) $\forall\, i = 1, \ldots, d$, $X_i \sim \mathcal{U}([0,1])$.

Assumption ii) is not restrictif : with the inverse technique, $Y = \mathcal{M}(X_1, \ldots, X_d)$ can be written as

$$Y = \mathcal{M}(F_{X_1}^{-1}(U_1), \ldots, F_{X_d}^{-1}(U_d)) = \widetilde{\mathcal{M}}(U_1, \ldots, U_d)$$

with $U_i, i = 1, \ldots d$ independent and for all i, $U_i \sim \mathcal{U}([0,1])$, $F_{X_i}^{-1}$ inverse of the cumulative distribution function of X_i.

The complex case of correlated inputs will be mentioned at the end of this lecture.

In the following, we assume :

i) the X_i are independent ;

ii) $\forall\, i = 1, \dots, d$, $X_i \sim \mathcal{U}([0,1])$.

Assumption ii) is not restrictif : with the inverse technique, $Y = \mathcal{M}(X_1, \dots, X_d)$ can be written as

$$Y = \mathcal{M}(F_{X_1}^{-1}(U_1), \dots, F_{X_d}^{-1}(U_d)) = \widetilde{\mathcal{M}}(U_1, \dots, U_d)$$

with $U_i, i = 1, \dots d$ independent and for all i, $U_i \sim \mathcal{U}([0,1])$, $F_{X_i}^{-1}$ inverse of the cumulative distribution function of X_i.

The complex case of correlated inputs will be mentioned at the end of this lecture.

Towards Sobol sensitivity indices

Is the output Y more or less variable when input are fixed?

$\mathrm{Var}\,(Y|X_i = x_i)$, how to choose x_i? $\Rightarrow E\,[V\,(Y|X_i)]$

the smaller this quantity, (i.e. fixing X_i), the smaller is the variance of Y when fixing the ith input: variable X_i has a strong impact.

Theorem (Total variance)

$\mathrm{Var}(Y) = \mathrm{Var}\,[\mathbb{E}\,(Y|X_i)] + \mathbb{E}\,[\mathrm{Var}\,(Y|X_i)]$.

Definition (First order Sobol' Index)

$i = 1, \ldots, d$

$$0 \leq S_i = \frac{V\,[E\,(Y|X_i)]}{\mathrm{Var}(Y)} \leq 1$$

ex. : linear output $Y = \sum_{i=1}^{d} \beta_i X_i$, we get $S_i = \frac{\beta_i^2 \mathrm{Var}(X_i)}{\mathrm{Var}(Y)} = \rho_i^2$.

Towards Sobol sensitivity indices

Is the output Y more or less variable when input are fixed?

$\mathrm{Var}\left(Y|X_i = x_i\right)$, how to choose x_i? $\Rightarrow E\left[V\left(Y|X_i\right)\right]$

the smaller this quantity, (i.e. fixing X_i), the smaller is the variance of Y when fixing the ith input: variable X_i has a strong impact.

Theorem (Total variance)

$\mathrm{Var}(Y) = \mathrm{Var}\left[\mathbb{E}\left(Y|X_i\right)\right] + \mathbb{E}\left[\mathrm{Var}\left(Y|X_i\right)\right].$

Definition (First order Sobol' Index)

$i = 1, \ldots, d$

$$0 \leq S_i = \frac{V\left[E\left(Y|X_i\right)\right]}{\mathrm{Var}(Y)} \leq 1$$

ex. : linear output $Y = \sum_{i=1}^{d} \beta_i X_i$, we get $S_i = \frac{\beta_i^2 \mathrm{Var}(X_i)}{\mathrm{Var}(Y)} = \rho_i^2$.

Towards Sobol sensitivity indices

Is the output Y more or less variable when input are fixed?

$\mathrm{Var}\left(Y|X_i = x_i\right)$, how to choose x_i? $\Rightarrow E\left[V\left(Y|X_i\right)\right]$

the smaller this quantity, (i.e. fixing X_i), the smaller is the variance of Y when fixing the ith input: variable X_i has a strong impact.

Theorem (Total variance)

$\mathrm{Var}(Y) = \mathrm{Var}\left[\mathbb{E}\left(Y|X_i\right)\right] + \mathbb{E}\left[\mathrm{Var}\left(Y|X_i\right)\right].$

Definition (First order Sobol' Index)

$i = 1, \ldots, d$

$$0 \leq S_i = \frac{V\left[E\left(Y|X_i\right)\right]}{\mathrm{Var}(Y)} \leq 1$$

ex. : linear output $Y = \sum_{i=1}^{d} \beta_i X_i$, we get $S_i = \frac{\beta_i^2 \mathrm{Var}(X_i)}{\mathrm{Var}(Y)} = \rho_i^2$.

Towards Sobol sensitivity indices

Is the output Y more or less variable when input are fixed?

$\mathrm{Var}\,(Y|X_i = x_i)$, how to choose x_i? $\Rightarrow E\,[V\,(Y|X_i)]$

the smaller this quantity, (i.e. fixing X_i), the smaller is the variance of Y when fixing the ith input: variable X_i has a strong impact.

Theorem (Total variance)

$\mathrm{Var}(Y) = \mathrm{Var}\,[\mathbb{E}\,(Y|X_i)] + \mathbb{E}\,[\mathrm{Var}\,(Y|X_i)].$

Definition (First order Sobol' Index)

$i = 1, \ldots, d$

$$0 \leq S_i = \frac{V\,[E\,(Y|X_i)]}{\mathrm{Var}(Y)} \leq 1$$

ex. : linear output $Y = \sum_{i=1}^{d} \beta_i X_i$, we get $S_i = \frac{\beta_i^2 \mathrm{Var}(X_i)}{\mathrm{Var}(Y)} = \rho_i^2$.

Towards Sobol sensitivity indices

Is the output Y more or less variable when input are fixed?

$\mathrm{Var}(Y|X_i = x_i)$, how to choose x_i? $\Rightarrow E[V(Y|X_i)]$

the smaller this quantity, (i.e. fixing X_i), the smaller is the variance of Y when fixing the ith input: variable X_i has a strong impact.

Theorem (Total variance)

$\mathrm{Var}(Y) = \mathrm{Var}[\mathbb{E}(Y|X_i)] + \mathbb{E}[\mathrm{Var}(Y|X_i)].$

Definition (First order Sobol' Index)

$i = 1, \dots, d$

$$0 \leq S_i = \frac{V[E(Y|X_i)]}{\mathrm{Var}(Y)} \leq 1$$

ex. : linear output $Y = \sum_{i=1}^{d} \beta_i X_i$, we get $S_i = \frac{\beta_i^2 \mathrm{Var}(X_i)}{\mathrm{Var}(Y)} = \rho_i^2$.

Towards Sobol sensitivity indices

Is the output Y more or less variable when input are fixed?

$\mathrm{Var}\,(Y|X_i = x_i)$, how to choose x_i? $\Rightarrow E\,[V\,(Y|X_i)]$

the smaller this quantity, (i.e. fixing X_i), the smaller is the variance of Y when fixing the ith input: variable X_i has a strong impact.

Theorem (Total variance)

$\mathrm{Var}(Y) = \mathrm{Var}\left[\mathbb{E}\,(Y|X_i)\right] + \mathbb{E}\left[\mathrm{Var}\,(Y|X_i)\right].$

Definition (First order Sobol' Index)

$i = 1, \ldots, d$

$$0 \leq S_i = \frac{V\,[E\,(Y|X_i)]}{\mathrm{Var}(Y)} \leq 1$$

ex. : linear output $Y = \sum_{i=1}^{d} \beta_i X_i$, we get $S_i = \frac{\beta_i^2 \mathrm{Var}(X_i)}{\mathrm{Var}(Y)} = \rho_i^2$.

Toy case:

$$Y = X_1^2 + X_2 \quad X_i \sim \mathcal{U}([0,1]) \quad X_1 \perp\!\!\!\perp X_2$$

$$\mathbb{E}(Y|X_1) = X_1^2 + \mathbb{E}(X_2) \Rightarrow \mathrm{Var}[\mathbb{E}(Y|X_1)] = \mathrm{Var}(X_1^2) = \tfrac{4}{45}$$

$$\mathbb{E}(Y|X_2) = \mathbb{E}(X_1^2) + X_2 \Rightarrow \mathrm{Var}[\mathbb{E}(Y|X_2)] = \mathrm{Var}(X_2) = \tfrac{1}{12}$$

$$\mathrm{Var}(Y) = \mathrm{Var}(X_1^2) + \mathrm{Var}(X_2) = \tfrac{31}{180}$$

$$S_1 = \frac{16}{31} \approx 0,516\,,\ S_2 = \frac{15}{31} \approx 0,484$$

$S_1 + S_2 = 1$, additive model

Toy case:

$$Y = X_1^2 + X_2 \qquad X_i \sim \mathcal{U}([0,1]) \qquad X_1 \perp\!\!\!\perp X_2$$

$$\mathbb{E}(Y|X_1) = X_1^2 + \mathbb{E}(X_2) \Rightarrow \mathrm{Var}[\mathbb{E}(Y|X_1)] = \mathrm{Var}(X_1^2) = \frac{4}{45}$$

$$\mathbb{E}(Y|X_2) = \mathbb{E}(X_1^2) + X_2 \Rightarrow \mathrm{Var}[\mathbb{E}(Y|X_2)] = \mathrm{Var}(X_2) = \frac{1}{12}$$

$$\mathrm{Var}(Y) = \mathrm{Var}(X_1^2) + \mathrm{Var}(X_2) = \frac{31}{180}$$

$$S_1 = \frac{16}{31} \approx 0,516\,,\ S_2 = \frac{15}{31} \approx 0,484$$

$S_1 + S_2 = 1$, additive model

Toy case:

$$Y = X_1^2 + X_2 \quad X_i \sim \mathcal{U}([0,1]) \quad X_1 \perp\!\!\!\perp X_2$$

$$\mathbb{E}(Y|X_1) = X_1^2 + \mathbb{E}(X_2) \Rightarrow \mathrm{Var}\left[\mathbb{E}(Y|X_1)\right] = \mathrm{Var}(X_1^2) = \frac{4}{45}$$

$$\mathbb{E}(Y|X_2) = \mathbb{E}(X_1^2) + X_2 \Rightarrow \mathrm{Var}\left[\mathbb{E}(Y|X_2)\right] = \mathrm{Var}(X_2) = \frac{1}{12}$$

$$\mathrm{Var}(Y) = \mathrm{Var}(X_1^2) + \mathrm{Var}(X_2) = \frac{31}{180}$$

$$S_1 = \frac{16}{31} \approx 0,516\,, \; S_2 = \frac{15}{31} \approx 0,484$$

$S_1 + S_2 = 1$, additive model

Toy case:

$$Y = X_1^2 + X_2 \quad X_i \sim \mathcal{U}([0,1]) \quad X_1 \perp\!\!\!\perp X_2$$

$$\mathbb{E}(Y|X_1) = X_1^2 + \mathbb{E}(X_2) \Rightarrow \mathrm{Var}\left[\mathbb{E}(Y|X_1)\right] = \mathrm{Var}(X_1^2) = \frac{4}{45}$$

$$\mathbb{E}(Y|X_2) = \mathbb{E}(X_1^2) + X_2 \Rightarrow \mathrm{Var}\left[\mathbb{E}(Y|X_2)\right] = \mathrm{Var}(X_2) = \frac{1}{12}$$

$$\mathrm{Var}(Y) = \mathrm{Var}(X_1^2) + \mathrm{Var}(X_2) = \frac{31}{180}$$

$$S_1 = \frac{16}{31} \approx 0,516\,, \; S_2 = \frac{15}{31} \approx 0,484$$

$S_1 + S_2 = 1$, additive model

Toy case:

$$Y = X_1^2 + X_2 \qquad X_i \sim \mathcal{U}([0,1]) \quad X_1 \perp\!\!\!\perp X_2$$

$$\mathbb{E}(Y|X_1) = X_1^2 + \mathbb{E}(X_2) \Rightarrow \mathrm{Var}[\mathbb{E}(Y|X_1)] = \mathrm{Var}(X_1^2) = \tfrac{4}{45}$$

$$\mathbb{E}(Y|X_2) = \mathbb{E}(X_1^2) + X_2 \Rightarrow \mathrm{Var}[\mathbb{E}(Y|X_2)] = \mathrm{Var}(X_2) = \tfrac{1}{12}$$

$$\mathrm{Var}(Y) = \mathrm{Var}(X_1^2) + \mathrm{Var}(X_2) = \tfrac{31}{180}$$

$$S_1 = \frac{16}{31} \approx 0,516\,, \; S_2 = \frac{15}{31} \approx 0,484$$

$S_1 + S_2 = 1$, additive model

Toy case:

$$Y = X_1^2 + X_2 \quad X_i \sim \mathcal{U}([0,1]) \quad X_1 \perp\!\!\!\perp X_2$$

$$\mathbb{E}(Y|X_1) = X_1^2 + \mathbb{E}(X_2) \Rightarrow \mathrm{Var}[\mathbb{E}(Y|X_1)] = \mathrm{Var}(X_1^2) = \frac{4}{45}$$

$$\mathbb{E}(Y|X_2) = \mathbb{E}(X_1^2) + X_2 \Rightarrow \mathrm{Var}[\mathbb{E}(Y|X_2)] = \mathrm{Var}(X_2) = \frac{1}{12}$$

$$\mathrm{Var}(Y) = \mathrm{Var}(X_1^2) + \mathrm{Var}(X_2) = \frac{31}{180}$$

$$S_1 = \frac{16}{31} \approx 0,516\,, \; S_2 = \frac{15}{31} \approx 0,484$$

$S_1 + S_2 = 1$, additive model

More generally,

Theorem (Hoeffding decomposition)

$\mathcal{M} : [0,1]^d \to \mathbb{R}$, $\int_{[0,1]^d} \mathcal{M}^2(x)dx < \infty$

$\mathcal{M}$ has an unique decomposition

$\mathcal{M}_0 + \sum_{i=1}^d \mathcal{M}_i(x_i) + \sum_{1 \leq i < j \leq d} \mathcal{M}_{i,j}(x_i, x_j) + \ldots + \mathcal{M}_{1,\ldots,d}(x_1, \ldots, x_d)$

under the constraint

- *$\mathcal{M}_0$ constant,*
- $\forall\, 1 \leq s \leq d$, $\forall\, 1 \leq i_1 < \ldots < i_s \leq d$, $\forall\, 1 \leq p \leq s$

$$\int_0^1 \mathcal{M}_{i_1,\ldots,i_s}(x_{i_1}, \ldots, x_{i_s}) dx_{i_p} = 0$$

Consequences : $\mathcal{M}_0 = \int_{[0,1]^d} \mathcal{M}(x)dx$ and the terms of the decomposition are orthogonal.

The computation of each term in the decomposition writes:

- $\mathcal{M}_i(x_i) = \int_{[0,1]^{d-1}} \mathcal{M}(x) \Pi_{p \neq i} dx_p - \mathcal{M}_0$
- $i \neq j$
  $\mathcal{M}_{i,j}(x_i, x_j) = \int_{[0,1]^{d-2}} \mathcal{M}(x) \Pi_{p \neq i,j} dx_p - \mathcal{M}_0 - \mathcal{M}_i(x_i) - \mathcal{M}_j(x_j)$
- ...

$\Rightarrow$ computation of multiple integrals.

Consequences : $\mathcal{M}_0 = \int_{[0,1]^d} \mathcal{M}(x)dx$ and the terms of the decomposition are orthogonal.

The computation of each term in the decomposition writes:

- $\mathcal{M}_i(x_i) = \int_{[0,1]^{d-1}} \mathcal{M}(x) \Pi_{p \neq i} dx_p - \mathcal{M}_0$
- $i \neq j$
  $\mathcal{M}_{i,j}(x_i, x_j) = \int_{[0,1]^{d-2}} \mathcal{M}(x) \Pi_{p \neq i,j} dx_p - \mathcal{M}_0 - \mathcal{M}_i(x_i) - \mathcal{M}_j(x_j)$
- ...

$\Rightarrow$ computation of multiple integrals.

Consequences : $\mathcal{M}_0 = \int_{[0,1]^d} \mathcal{M}(x)dx$ and the terms of the decomposition are orthogonal.

The computation of each term in the decomposition writes:

- $\mathcal{M}_i(x_i) = \int_{[0,1]^{d-1}} \mathcal{M}(x) \Pi_{p \neq i} dx_p - \mathcal{M}_0$
- $i \neq j$
  $\mathcal{M}_{i,j}(x_i, x_j) = \int_{[0,1]^{d-2}} \mathcal{M}(x) \Pi_{p \neq i,j} dx_p - \mathcal{M}_0 - \mathcal{M}_i(x_i) - \mathcal{M}_j(x_j)$
- ...

$\Rightarrow$ computation of multiple integrals.

<u>Consequences :</u> $\mathcal{M}_0 = \int_{[0,1]^d} \mathcal{M}(x)dx$ and the terms of the decomposition are orthogonal.

The computation of each term in the decomposition writes:

- $\mathcal{M}_i(x_i) = \int_{[0,1]^{d-1}} \mathcal{M}(x) \Pi_{p \neq i} dx_p - \mathcal{M}_0$
- $i \neq j$
  $\mathcal{M}_{i,j}(x_i, x_j) = \int_{[0,1]^{d-2}} \mathcal{M}(x) \Pi_{p \neq i,j} dx_p - \mathcal{M}_0 - \mathcal{M}_i(x_i) - \mathcal{M}_j(x_j)$
- $\ldots$

$\Rightarrow$ computation of multiple integrals.

Variance decomposition : $X_1, \dots, X_d$ i.i.d. $\sim \mathcal{U}([0,1])$

$$Y = \mathcal{M}(X) = \mathcal{M}_0 + \sum_{i=1}^{d} \mathcal{M}_i(X_i) + \ldots + \mathcal{M}_{1,\dots,d}(X_1, \dots, X_d)$$

- $\mathcal{M}_0 = \mathbb{E}(Y)$,
- $\mathcal{M}_i(X_i) = \mathbb{E}(Y|X_i) - \mathbb{E}(Y)$,
- $i \neq j$ $\mathcal{M}_{i,j}(X_i, X_j) = \mathbb{E}(Y|X_i, X_j) - \mathbb{E}(Y|X_i) - \mathbb{E}(Y|X_j) + \mathbb{E}(Y)$,
- $\ldots$

$$\mathrm{Var}(Y) = \sum_{i=1}^{d} \mathrm{Var}(\mathcal{M}_i(X_i)) + \ldots + \mathrm{Var}(\mathcal{M}_{1,\dots,d}(X_1, \dots, X_d))$$

<u>Variance decomposition :</u> $X_1, \ldots, X_d$ i.i.d. $\sim \mathcal{U}([0,1])$

$$Y = \mathcal{M}(X) = \mathcal{M}_0 + \sum_{i=1}^{d} \mathcal{M}_i(X_i) + \ldots + \mathcal{M}_{1,\ldots,d}(X_1, \ldots, X_d)$$

- $\mathcal{M}_0 = \mathbb{E}(Y)$,
- $\mathcal{M}_i(X_i) = \mathbb{E}(Y|X_i) - \mathbb{E}(Y)$,
- $i \neq j$ $\mathcal{M}_{i,j}(X_i, X_j) = \mathbb{E}(Y|X_i, X_j) - \mathbb{E}(Y|X_i) - \mathbb{E}(Y|X_j) + \mathbb{E}(Y)$,
- $\ldots$

$$\mathrm{Var}(Y) = \sum_{i=1}^{d} \mathrm{Var}(\mathcal{M}_i(X_i)) + \ldots + \mathrm{Var}(\mathcal{M}_{1,\ldots,d}(X_1, \ldots, X_d))$$

Definition (Sobol' indices)

$$\forall\, i = 1, \ldots, d \;\; S_i = \frac{\mathrm{Var}(\mathcal{M}_i(X_i))}{\mathrm{Var}(Y)} = \frac{\mathrm{Var}[\mathbb{E}(Y|X_i)]}{\mathrm{Var}(Y)}$$

$$\forall\, i \neq j \;\; S_{i,j} = \frac{\mathrm{Var}\big(\mathcal{M}_{i,j}(X_i, X_j)\big)}{\mathrm{Var}(Y)} = \frac{\mathrm{Var}\big[\mathbb{E}\big(Y|X_i, X_j\big)\big] - \mathrm{Var}[\mathbb{E}(Y|X_i)] - \mathrm{Var}\big[\mathbb{E}\big(Y|X_j\big)\big]}{\mathrm{Var}(Y)}$$

$\ldots$

$$1 = \sum_{i=1}^{d} S_i + \sum_{i \neq j} S_{i,j} + \ldots + S_{1,\ldots,d}$$

Definition (Total indices)

$$i = 1, \ldots, d \quad S_{T_i} = \sum_{\mathbf{u} \subset \{1,\ldots,d\}\,,\, \mathbf{u} \neq \emptyset\,,\, i \in \mathbf{u}} S_{\mathbf{u}}\,.$$

Definition (Sobol' indices)

$$\forall\, i = 1,\ldots,d \quad S_i = \frac{\mathrm{Var}(\mathcal{M}_i(X_i))}{\mathrm{Var}(Y)} = \frac{\mathrm{Var}[\mathbb{E}(Y|X_i)]}{\mathrm{Var}(Y)}$$

$$\forall\, i \neq j \quad S_{i,j} = \frac{\mathrm{Var}\big(\mathcal{M}_{i,j}(X_i,X_j)\big)}{\mathrm{Var}(Y)} = \frac{\mathrm{Var}\big[\mathbb{E}\big(Y|X_i,X_j\big)\big] - \mathrm{Var}[\mathbb{E}(Y|X_i)] - \mathrm{Var}\big[\mathbb{E}\big(Y|X_j\big)\big]}{\mathrm{Var}(Y)}$$

$$\ldots$$

$$1 = \sum_{i=1}^{d} S_i + \sum_{i \neq j} S_{i,j} + \ldots + S_{1,\ldots,d}$$

Definition (Total indices)

$$i = 1,\ldots,d \qquad S_{T_i} = \sum_{\mathbf{u} \subset \{1,\ldots,d\}\,,\ \mathbf{u} \neq \emptyset\,,\ i \in \mathbf{u}} S_{\mathbf{u}}\ .$$

Sobol' indices :

Definition (Total indices)

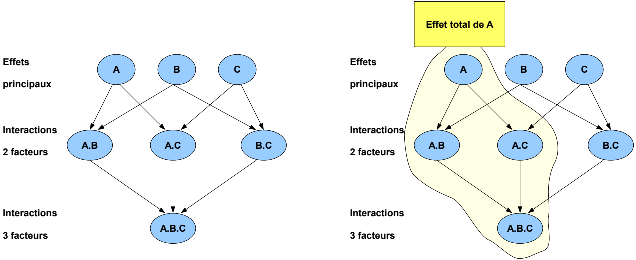
$$i = 1, \dots, d \qquad S_{T_i} = \sum_{\mathbf{u} \subset \{1,\dots,d\}\,,\, \mathbf{u} \neq \emptyset\,,\, i \in \mathbf{u}} S_{\mathbf{u}} \,.$$

$X_{(-i)} = (X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_d)$

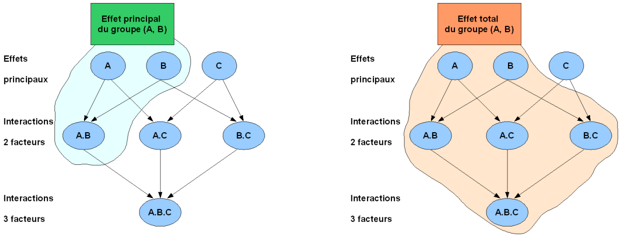
Using the theorem of the total variance,

$$S_{T_i} = \frac{\mathbb{E}\left[\mathrm{Var}\left(Y | X_{(-i)}\right)\right]}{\mathrm{Var}(Y)} = 1 - \frac{\mathrm{Var}\left[\mathbb{E}\left(Y | X_{(-i)}\right)\right]}{\mathrm{Var}(Y)} \,.$$

Indices with factor:



Indices with groupe of factors:



Fact : Analytical expressions of Sobol' indices, with integrals in high dimensional spaces, are rarely available.

Two inferential approaches

1) Monte-Carlo type (hypothesis $\mathbb{L}^2$ with the model);
2) spectral techniques (additional hypotheses of regularity).

If the model is too costly to assess, we fit a metamodel before applying these techniques.

ex.: parametric and non-parametric regressions, Gaussian metamodel (see practical session), . . .

Fact : Analytical expressions of Sobol' indices, with integrals in high dimensional spaces, are rarely available.

Two inferential approaches

1) Monte-Carlo type (hypothesis $\mathbb{L}^2$ with the model);
2) spectral techniques (additional hypotheses of regularity).

If the model is too costly to assess, we fit a metamodel before applying these techniques.

ex.: parametric and non-parametric regressions, Gaussian metamodel (see practical session), . . .

Fact : Analytical expressions of Sobol' indices, with integrals in high dimensional spaces, are rarely available.

Two inferential approaches

1) Monte-Carlo type (hypothesis $\mathbb{L}^2$ with the model);
2) spectral techniques (additional hypotheses of regularity).

If the model is too costly to assess, we fit a metamodel before applying these techniques.

ex.: parametric and non-parametric regressions, Gaussian metamodel (see practical session), . . .

Monte-Carlo type Approaches : (Sobol' 93, Saltelli 02, Mauntz, ...)

Idea : $X'_{(-i)}$ indep. copy of $X_{(-i)}$, $Y = \mathcal{M}(X_i, X_{(-i)})$, $Y^i = \mathcal{M}(X_i, X'_{(-i)})$

We have $S_i = \frac{\mathrm{Cov}(Y, Y^i)}{\mathrm{Var}(Y)}$, the idea is based on empirical formulas.

Two independent samples A and B (Monte-Carlo, LHS)

$$A = \begin{pmatrix} x^A_{1,1} & \cdots & x^A_{d,1} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ x^A_{1,n} & \cdots & x^A_{d,n} \end{pmatrix} \qquad B = \begin{pmatrix} x^B_{1,1} & \cdots & x^B_{d,1} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ x^B_{1,n} & \cdots & x^B_{d,n} \end{pmatrix}$$

From A and of B, we create d sampling matrices C_i, $i = 1, \ldots, d$.

$$C_i = \begin{pmatrix} x^A_{1,1} & \cdots & x^B_{i,1} & \cdots & x^A_{d,1} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ x^A_{1,n} & \cdots & x^B_{i,n} & \cdots & x^A_{d,n} \end{pmatrix}$$

We compute $(2+d) \times n$ the model $\mathcal{M}$:

$$y^A = \begin{pmatrix} y_1^A \\ \cdot \\ \cdot \\ \cdot \\ y_n^A \end{pmatrix} \quad y^B = \begin{pmatrix} y_1^B \\ \cdot \\ \cdot \\ \cdot \\ y_n^B \end{pmatrix} \quad \forall\, 1 \leq i \leq d \quad y^{C_i} = \begin{pmatrix} y_1^{C_i} \\ \cdot \\ \cdot \\ \cdot \\ y_n^{C_i} \end{pmatrix}$$

`sobolEff()` (Janon *et al.*, 2012 & 2013)

- $\hat{V}_i = \frac{1}{n}\sum_{k=1}^{n} y_k^B y_k^{C_i} - \left(\frac{1}{n}\sum_{k=1}^{n} \frac{y_k^B + y_k^{C_i}}{2}\right)^2$ numerator of the first-order index
- $\hat{V} = \frac{1}{n}\sum_{k=1}^{n} \frac{(y_k^B)^2 + (y_k^{C_i})^2}{2} - \left(\frac{1}{n}\sum_{k=1}^{n} \frac{y_k^B + y_k^{C_i}}{2}\right)^2$ denominator

asymptotic or bootstrap confidence intervals (see practical session)

asymptotic efficiency of the estimator

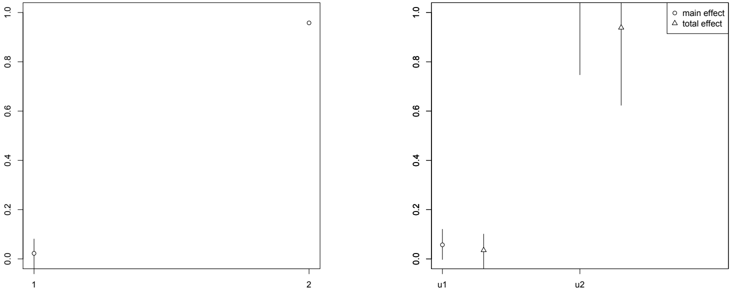
Remark :

We can also replace the MC ou LHS samplings with QMC (hyp. of regular variations).

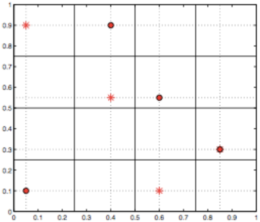
The g-function of Sobol' : $f(x) = f_1(x_1) * f_2(x_2)$ with $f_i(x_i) = \frac{|4x_i - 2| + a_i}{1 + a_i}$, $a_1 = 9$, $a_2 = 1$.

$S_1 \approx 0.038$, $S_2 \approx 0.958$.

$n = 1000$, $b = 100$, IC(0.95) `sobolEff` (left), `sobol2007` (right)



Replicate latin hypercubes: (Tissot *et al.*)



Definition (Replicated Latin Hypercube Sampling)

$k = 1, \dots, n$

$$\mathbf{x}_k = \left(\frac{\pi_1(k) - U_{1,\pi_1(k)}}{n}, \dots, \frac{\pi_d(k) - U_{d,\pi_d(k)}}{n} \right)$$

$$\tilde{\mathbf{x}}_k = \left(\frac{\tilde{\pi}_1(k) - U_{1,\tilde{\pi}_1(k)}}{n}, \dots, \frac{\tilde{\pi}_d(k) - U_{d,\tilde{\pi}_d(k)}}{n} \right)$$

We have two matrices B and $\widetilde{B}$ at our disposal

$$B = \begin{pmatrix} x_{1,1} & \dots & x_{d,1} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ x_{1,n} & \dots & x_{d,n} \end{pmatrix} \qquad \widetilde{B} = \begin{pmatrix} \tilde{x}_{1,1} & \dots & \tilde{x}_{d,1} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \tilde{x}_{1,n} & \dots & \tilde{x}_{d,n} \end{pmatrix}$$

We compute the model $\mathcal{M}$ 2n times (on the n lines of B and the n lines of $\widetilde{B}$).

permut. of lines :

$$\begin{cases} \widetilde{B} = (\tilde{x}_{k,l})_{1\leq k\leq n, 1\leq l\leq d} & \rightarrow \quad \widetilde{B}_i = (\tilde{x}^i_{k,l})_{1\leq k\leq n, 1\leq l\leq d} \\ L_k & \mapsto \quad L_{\pi_i^{-1}\circ\tilde{\pi}_i(k)}, \; k = 1, \dots, n \end{cases}$$

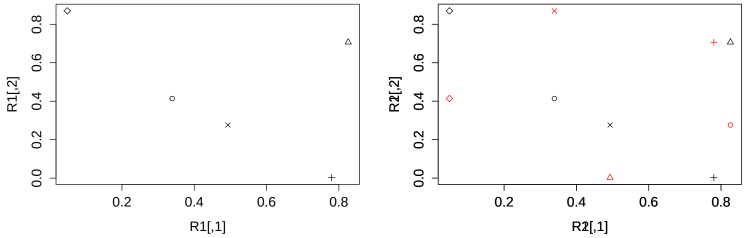
Then, $\tilde{x}^i_{k,i} = \tilde{x}_{\tilde{\pi}_i^{-1}\circ\pi_i(k),i} = x_{k,i}$, $k = 1, \dots, n$.

To estimate S_i, we replace C_i with $\widetilde{B}_i$ (same column number i).

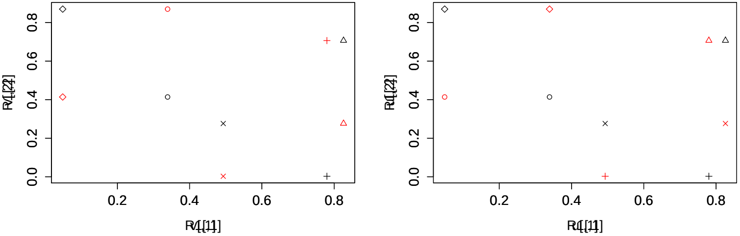
Caption:

point 1 $\circ$ point 2 $\triangle$ point 3 $+$ point 4 $\times$ point 5 $\diamond$

Design B (left), B and $\widetilde{B}$ (right, black and red)



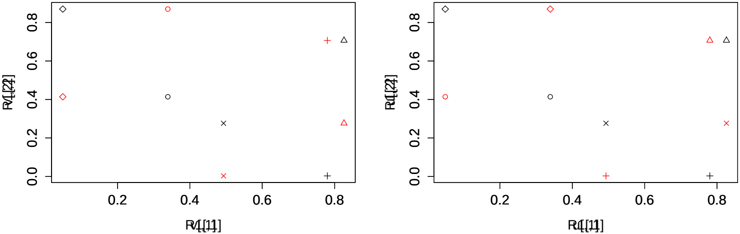
Design $\widetilde{B}_1$ (left), $\widetilde{B}_2$ (right)



Asymptotic confidence intervals with variance smaller than for MC.

Possible extension to indices of order two (via orthogonal arrays of strength 2).

Design $\widetilde{B}_1$ (left), $\widetilde{B}_2$ (right)



Asymptotic confidence intervals with variance smaller than for MC.

Possible extension to indices of order two (via orthogonal arrays of strength 2).

Spectral approaches: (case $d = 2$)
$Y = \sum_{\mathbf{k}=(k_1,k_2)\in\mathbb{Z}^2} c_{\mathbf{k}}(\mathcal{M})\Phi_{1,k_1}(X_1)\Phi_{2,k_2}(X_2)$

with , for all $i = 1, 2$, $(\Phi_{i,k})_{k\in\mathbb{Z}}$ is an orthonormal basis of $\mathbb{L}^2([0,1])$ and $\Phi_{i,0} \equiv 1$.

$\mathcal{M}_0 = c_0(\mathcal{M})$,

$\mathcal{M}_1(X_1) = \sum_{k_1\in\mathbb{Z}^*} c_{k_1,0}(\mathcal{M})\Phi_{1,k_1}(X_1)$,
$\mathcal{M}_2(X_2) = \sum_{k_2\in\mathbb{Z}^*} c_{0,k_2}(\mathcal{M})\Phi_{2,k_2}(X_2)$,

$\mathcal{M}_{1,2}(X_1,X_2) = \sum_{k_1\in\mathbb{Z}^*,k_2\in\mathbb{Z}^*} c_{k_1,k_2}(\mathcal{M})\Phi_{1,k_1}(X_1)\Phi_{2,k_2}(X_2)$.

We have with Parseval identity:

- $\mathrm{Var}(\mathcal{M}_1(X_1)) = \sigma_1^2 = \sum_{k_1\in\mathbb{Z}^*} |c_{k_1,0}(\mathcal{M})|^2$, (idem for σ_2^2),
- $\mathrm{Var}(\mathcal{M}_{1,2}(X_1,X_2)) = \sigma_{1,2}^2 = \sum_{k_1\in\mathbb{Z}^*,k_2\in\mathbb{Z}^*} |c_{k_1,k_2}(\mathcal{M})|^2$,
- $\mathrm{Var}(Y) = \sigma^2 = \sum_{(k_1,k_2)\in\mathbb{Z}\times\mathbb{Z},(k_1,k_2)\neq(0,0} |c_{k_1,k_2}(\mathcal{M})|^2$.

ex. : orthogonal polynomials, wavelet basis, Fourier basis.

Spectral approaches: (case $d = 2$)

$Y = \sum_{\mathbf{k}=(k_1,k_2)\in\mathbb{Z}^2} c_{\mathbf{k}}(\mathcal{M})\Phi_{1,k_1}(X_1)\Phi_{2,k_2}(X_2)$

with , for all $i = 1, 2$, $(\Phi_{i,k})_{k\in\mathbb{Z}}$ is an orthonormal basis of $\mathbb{L}^2([0,1])$ and $\Phi_{i,0} \equiv 1$.

$\mathcal{M}_0 = c_0(\mathcal{M})$,

$\mathcal{M}_1(X_1) = \sum_{k_1\in\mathbb{Z}^*} c_{k_1,0}(\mathcal{M})\Phi_{1,k_1}(X_1)$,
$\mathcal{M}_2(X_2) = \sum_{k_2\in\mathbb{Z}^*} c_{0,k_2}(\mathcal{M})\Phi_{2,k_2}(X_2)$,

$\mathcal{M}_{1,2}(X_1, X_2) = \sum_{k_1\in\mathbb{Z}^*,k_2\in\mathbb{Z}^*} c_{k_1,k_2}(\mathcal{M})\Phi_{1,k_1}(X_1)\Phi_{2,k_2}(X_2)$.

We have with Parseval identity:

- $\mathrm{Var}\left(\mathcal{M}_1(X_1)\right) = \sigma_1^2 = \sum_{k_1\in\mathbb{Z}^*} |c_{k_1,0}(\mathcal{M})|^2$, (idem for σ_2^2),
- $\mathrm{Var}\left(\mathcal{M}_{1,2}(X_1, X_2)\right) = \sigma_{1,2}^2 = \sum_{k_1\in\mathbb{Z}^*,k_2\in\mathbb{Z}^*} |c_{k_1,k_2}(\mathcal{M})|^2$,
- $\mathrm{Var}\left(Y\right) = \sigma^2 = \sum_{(k_1,k_2)\in\mathbb{Z}\times\mathbb{Z},(k_1,k_2)\neq(0,0} |c_{k_1,k_2}(\mathcal{M})|^2$.

ex. : orthogonal polynomials, wavelet basis, Fourier basis.

Spectral approaches: (case $d = 2$)
$Y = \sum_{\mathbf{k}=(k_1,k_2)\in\mathbb{Z}^2} c_{\mathbf{k}}(\mathcal{M})\Phi_{1,k_1}(X_1)\Phi_{2,k_2}(X_2)$

with , for all $i = 1, 2$, $(\Phi_{i,k})_{k\in\mathbb{Z}}$ is an orthonormal basis of $\mathbb{L}^2([0,1])$ and $\Phi_{i,0} \equiv 1$.

$\mathcal{M}_0 = c_0(\mathcal{M})$,

$\mathcal{M}_1(X_1) = \sum_{k_1\in\mathbb{Z}^*} c_{k_1,0}(\mathcal{M})\Phi_{1,k_1}(X_1)$,
$\mathcal{M}_2(X_2) = \sum_{k_2\in\mathbb{Z}^*} c_{0,k_2}(\mathcal{M})\Phi_{2,k_2}(X_2)$,

$\mathcal{M}_{1,2}(X_1, X_2) = \sum_{k_1\in\mathbb{Z}^*,k_2\in\mathbb{Z}^*} c_{k_1,k_2}(\mathcal{M})\Phi_{1,k_1}(X_1)\Phi_{2,k_2}(X_2)$.

We have with Parseval identity:

- $\mathrm{Var}(\mathcal{M}_1(X_1)) = \sigma_1^2 = \sum_{k_1\in\mathbb{Z}^*} |c_{k_1,0}(\mathcal{M})|^2$, (idem for σ_2^2),
- $\mathrm{Var}(\mathcal{M}_{1,2}(X_1, X_2)) = \sigma_{1,2}^2 = \sum_{k_1\in\mathbb{Z}^*,k_2\in\mathbb{Z}^*} |c_{k_1,k_2}(\mathcal{M})|^2$,
- $\mathrm{Var}(Y) = \sigma^2 = \sum_{(k_1,k_2)\in\mathbb{Z}\times\mathbb{Z},(k_1,k_2)\neq(0,0} |c_{k_1,k_2}(\mathcal{M})|^2$.

ex. : orthogonal polynomials, wavelet basis, Fourier basis.

Spectral approaches: (case $d = 2$)

$Y = \sum_{\mathbf{k}=(k_1,k_2)\in\mathbb{Z}^2} c_{\mathbf{k}}(\mathcal{M})\Phi_{1,k_1}(X_1)\Phi_{2,k_2}(X_2)$

with , for all $i = 1, 2$, $(\Phi_{i,k})_{k\in\mathbb{Z}}$ is an orthonormal basis of $\mathbb{L}^2([0,1])$ and $\Phi_{i,0} \equiv 1$.

$\mathcal{M}_0 = c_0(\mathcal{M})$,

$\mathcal{M}_1(X_1) = \sum_{k_1\in\mathbb{Z}^*} c_{k_1,0}(\mathcal{M})\Phi_{1,k_1}(X_1)$,
$\mathcal{M}_2(X_2) = \sum_{k_2\in\mathbb{Z}^*} c_{0,k_2}(\mathcal{M})\Phi_{2,k_2}(X_2)$,

$\mathcal{M}_{1,2}(X_1, X_2) = \sum_{k_1\in\mathbb{Z}^*,k_2\in\mathbb{Z}^*} c_{k_1,k_2}(\mathcal{M})\Phi_{1,k_1}(X_1)\Phi_{2,k_2}(X_2)$.

We have with Parseval identity:

- $\mathrm{Var}\left(\mathcal{M}_1(X_1)\right) = \sigma_1^2 = \sum_{k_1\in\mathbb{Z}^*} |c_{k_1,0}(\mathcal{M})|^2$, (idem for σ_2^2),
- $\mathrm{Var}\left(\mathcal{M}_{1,2}(X_1, X_2)\right) = \sigma_{1,2}^2 = \sum_{k_1\in\mathbb{Z}^*,k_2\in\mathbb{Z}^*} |c_{k_1,k_2}(\mathcal{M})|^2$,
- $\mathrm{Var}\left(Y\right) = \sigma^2 = \sum_{(k_1,k_2)\in\mathbb{Z}\times\mathbb{Z},(k_1,k_2)\neq(0,0} |c_{k_1,k_2}(\mathcal{M})|^2$.

ex. : orthogonal polynomials, wavelet basis, Fourier basis.

III- Sobol' indices inference

Inference scheme:

If D is an experimental design with $[0,1]^2$, we propose the quadrature formula:

$$\hat{c}_{k_1,k_2}(\mathcal{M}, D) = \frac{1}{\mathrm{card} D} \sum_{\mathbf{x}=(x_1,x_2)\in D} \mathcal{M}(x) e^{-2i\pi(k_1 x_1 + k_2 x_2)} \,.$$

We then infer each part of variance with a truncation:

- $\hat{\sigma}_1^2(\mathcal{M}, K_1, D) = \sum_{k_1 \in K_1} |\hat{c}_{k_1,0}(\mathcal{M}, D)|^2$, with $K_1 \subset \mathbb{Z}^*$ of finite cardinal, (idem for $\hat{\sigma}_2^2$),
- $\hat{\sigma}_{1,2}^2(\mathcal{M}, K_{1,2}, D) = \sum_{(k_1,k_2)\in K_{1,2}} |\hat{c}_{k_1,k_2}(\mathcal{M}, D)|^2$, with $K_{1,2} \subset \mathbb{Z}^* \times \mathbb{Z}^*$ of finite cardinal.

We infer the total variance with
$\hat{\sigma}^2(\mathcal{M}, D) = \hat{c}_{0,0}(\mathcal{M}^2, D) - \hat{c}_{0,0}(\mathcal{M}, D)^2$.

The estimators of Sobol' indices can be written as:

$$\hat{S}_i = \frac{\hat{\sigma}_i^2}{\hat{\sigma}^2}\,,\ i = 1, 2\,, \quad S_{1,2} = \frac{\hat{\sigma}_{1,2}^2}{\hat{\sigma}^2}\,.$$

Inference scheme:

If D is an experimental design with $[0,1]^2$, we propose the quadrature formula:

$$\hat{c}_{k_1,k_2}(\mathcal{M}, D) = \frac{1}{\mathrm{card} D} \sum_{\mathbf{x}=(x_1,x_2)\in D} \mathcal{M}(x) e^{-2i\pi(k_1 x_1 + k_2 x_2)} .$$

We then infer each part of variance with a truncation:

- $\hat{\sigma}_1^2(\mathcal{M}, K_1, D) = \sum_{k_1\in K_1} |\hat{c}_{k_1,0}(\mathcal{M}, D)|^2$, with $K_1 \subset \mathbb{Z}^*$ of finite cardinal, (idem for $\hat{\sigma}_2^2$),
- $\hat{\sigma}_{1,2}^2(\mathcal{M}, K_{1,2}, D) = \sum_{(k_1,k_2)\in K_{1,2}} |\hat{c}_{k_1,k_2}(\mathcal{M}, D)|^2$, with $K_{1,2} \subset \mathbb{Z}^* \times \mathbb{Z}^*$ of finite cardinal.

We infer the total variance with $\hat{\sigma}^2(\mathcal{M}, D) = \hat{c}_{0,0}(\mathcal{M}^2, D) - \hat{c}_{0,0}(\mathcal{M}, D)^2$.

The estimators of Sobol' indices can be written as:

$$\hat{S}_i = \frac{\hat{\sigma}_i^2}{\hat{\sigma}^2}, \; i = 1, 2, \quad S_{1,2} = \frac{\hat{\sigma}_{1,2}^2}{\hat{\sigma}^2} .$$

Inference scheme:

If D is an experimental design with $[0,1]^2$, we propose the quadrature formula:

$$\hat{c}_{k_1,k_2}(\mathcal{M}, D) = \frac{1}{\text{card} D} \sum_{\mathbf{x}=(x_1,x_2)\in D} \mathcal{M}(x) e^{-2i\pi(k_1 x_1 + k_2 x_2)} .$$

We then infer each part of variance with a truncation:

- $\hat{\sigma}_1^2(\mathcal{M}, K_1, D) = \sum_{k_1 \in K_1} |\hat{c}_{k_1,0}(\mathcal{M}, D)|^2$, with $K_1 \subset \mathbb{Z}^*$ of finite cardinal, (idem for $\hat{\sigma}_2^2$),
- $\hat{\sigma}_{1,2}^2(\mathcal{M}, K_{1,2}, D) = \sum_{(k_1,k_2)\in K_{1,2}} |\hat{c}_{k_1,k_2}(\mathcal{M}, D)|^2$, with $K_{1,2} \subset \mathbb{Z}^* \times \mathbb{Z}^*$ of finite cardinal.

We infer the total variance with $\hat{\sigma}^2(\mathcal{M}, D) = \hat{c}_{0,0}(\mathcal{M}^2, D) - \hat{c}_{0,0}(\mathcal{M}, D)^2$.

The estimators of Sobol' indices can be written as:

$$\hat{S}_i = \frac{\hat{\sigma}_i^2}{\hat{\sigma}^2}, \quad i = 1, 2, \quad S_{1,2} = \frac{\hat{\sigma}_{1,2}^2}{\hat{\sigma}^2} .$$

Inference scheme:

If D is an experimental design with $[0,1]^2$, we propose the quadrature formula:

$$\hat{c}_{k_1,k_2}(\mathcal{M}, D) = \frac{1}{\mathrm{card} D} \sum_{\mathbf{x}=(x_1,x_2)\in D} \mathcal{M}(x) e^{-2i\pi(k_1 x_1 + k_2 x_2)} .$$

We then infer each part of variance with a truncation:

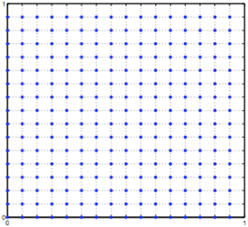
- $\hat{\sigma}_1^2(\mathcal{M}, K_1, D) = \sum_{k_1 \in K_1} |\hat{c}_{k_1,0}(\mathcal{M}, D)|^2$, with $K_1 \subset \mathbb{Z}^*$ of finite cardinal, (idem for $\hat{\sigma}_2^2$),
- $\hat{\sigma}_{1,2}^2(\mathcal{M}, K_{1,2}, D) = \sum_{(k_1,k_2)\in K_{1,2}} |\hat{c}_{k_1,k_2}(\mathcal{M}, D)|^2$, with $K_{1,2} \subset \mathbb{Z}^* \times \mathbb{Z}^*$ of finite cardinal.

We infer the total variance with
$\hat{\sigma}^2(\mathcal{M}, D) = \hat{c}_{0,0}(\mathcal{M}^2, D) - \hat{c}_{0,0}(\mathcal{M}, D)^2$.

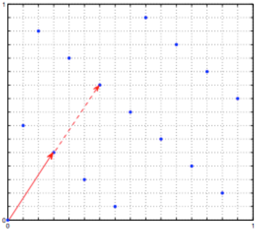
The estimators of Sobol' indices can be written as:

$$\hat{S}_i = \frac{\hat{\sigma}_i^2}{\hat{\sigma}^2} ,\ i = 1, 2 , \quad S_{1,2} = \frac{\hat{\sigma}_{1,2}^2}{\hat{\sigma}^2} .$$

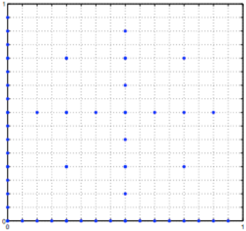
Classical designs:



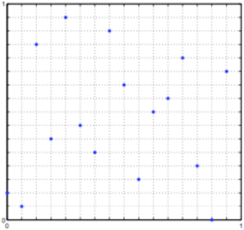
(a) grille régulière



(b) sous-groupe fini



(c) grille creuse



(d) tableau orthogonal

The performance of previous estimators is linked to the decreasing speed of Fourier spectrum (regularity) of $\mathcal{M}$. The techniques FAST and RBD are two particular cases of such approaches (after model regularisation).

FAST: (Cukier *et al.*, 78) *Fourier Amplitude Sensitivity Test*

• we fix $K_{\mathbf{u}}$ an ensemble of a priori non negligible frequencies;

• we chose D cyclic group (design (b)) in order to control the quadrature error.

Remarks:

- if $\mathcal{M}$ regular, we can obtain a speed of convergence $>> \sqrt{n}$;
- for the total indices `fast99()` (no confidence intervals in the function) Saltelli *et al.*, 99.

The performance of previous estimators is linked to the decreasing speed of Fourier spectrum (regularity) of $\mathcal{M}$. The techniques FAST and RBD are two particular cases of such approaches (after model regularisation).

FAST: (Cukier *et al.*, 78) *Fourier Amplitude Sensitivity Test*

• we fix $K_{\mathbf{u}}$ an ensemble of a priori non negligible frequencies;

• we chose D cyclic group (design (b)) in order to control the quadrature error.

Remarks:

- if $\mathcal{M}$ regular, we can obtain a speed of convergence $>> \sqrt{n}$;
- for the total indices `fast99()` (no confidence intervals in the function) Saltelli *et al.*, 99.

The performance of previous estimators is linked to the decreasing speed of Fourier spectrum (regularity) of $\mathcal{M}$. The techniques FAST and RBD are two particular cases of such approaches (after model regularisation).

FAST: (Cukier *et al.*, 78) *Fourier Amplitude Sensitivity Test*

• we fix $K_{\mathbf{u}}$ an ensemble of a priori non negligible frequencies;

• we chose D cyclic group (design (b)) in order to control the quadrature error.

Remarks:

- if $\mathcal{M}$ regular, we can obtain a speed of convergence $>> \sqrt{n}$;
- for the total indices `fast99()` (no confidence intervals in the function) Saltelli *et al.*, 99.

RBD: (Tarantola *et al.*, 06) *Random Balance Designs*

• we choose D a orthogonal array of strength 1 (design (d)), randomized by a random permutation ($D(\pi)$));

• $K_{\mathbf{u}}$ choice of a priori non negligible frequencies.

Remarks:

- these estimators are known to be biased;
- we can correct a part of this biais (Tissot *et al.*, 2012);
- if the function is not regular enough, the bias remains important.

RBD: (Tarantola *et al.*, 06) *Random Balance Designs*

• we choose D a orthogonal array of strength 1 (design (d)), randomized by a random permutation ($D(\pi)$));

• $K_{\mathbf{u}}$ choice of a priori non negligible frequencies.

Remarks:

- these estimators are known to be biased;
- we can correct a part of this biais (Tissot *et al.*, 2012);
- if the function is not regular enough, the bias remains important.

RBD: (Tarantola *et al.*, 06) *Random Balance Designs*

• we choose D a orthogonal array of strength 1 (design (d)), randomized by a random permutation ($D(\pi)$));

• $K_{\mathbf{u}}$ choice of a priori non negligible frequencies.

Remarks:

- these estimators are known to be biased;
- we can correct a part of this biais (Tissot *et al.*, 2012);
- if the function is not regular enough, the bias remains important.

Conclusions about the inference :

The spectral approach is interesting in terms of cost, but it requires more hypotheses in terms of regularity.

The Monte Carlo approach with replicated latin hypercubes is not as costly as the naive Monte Carlo approach.

see exercice 4 in the practical session

Conclusions about the inference :

The spectral approach is interesting in terms of cost, but it requires more hypotheses in terms of regularity.

The Monte Carlo approach with replicated latin hypercubes is not as costly as the naive Monte Carlo approach.

see exercice 4 in the practical session

Conclusions about the inference :

The spectral approach is interesting in terms of cost, but it requires more hypotheses in terms of regularity.

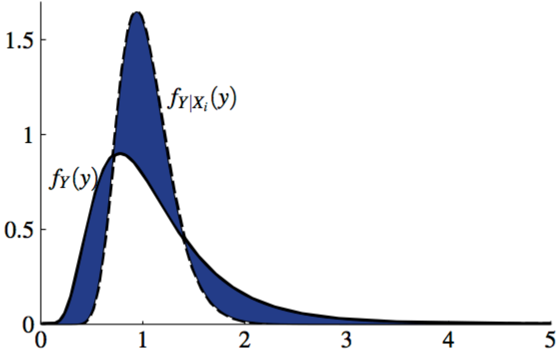
The Monte Carlo approach with replicated latin hypercubes is not as costly as the naive Monte Carlo approach.

see exercice 4 in the practical session

Borgonovo *et al.* $\geq$ 2007

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

<u>Remark:</u> extension to $\mathbf{u} \subset \{1, \ldots, d\}$



- ⋆ we can build metamodels when the initial model is too costly;
- ⋆ functionals inputs: MC approaches can be applied;
- ⋆ vectorial output: approaches component by component, or generalised indices GSI (Lamboni *et al.*;
- ⋆ functional output : we summarize the output by a vector (projection on an appropriate basis) or we make a movie (temporal output) or a map (spatial output) of indices;
- ⋆ correlated inputs: there is no an unique decomposition (like Hoeffding decomposition), hard to decouple interactions and correlation effects . . .

- we can build metamodels when the initial model is too costly;
- functionals inputs: MC approaches can be applied;
- vectorial output: approaches component by component, or generalised indices GSI (Lamboni *et al.*;
- functional output : we summarize the output by a vector (projection on an appropriate basis) or we make a movie (temporal output) or a map (spatial output) of indices;
- correlated inputs: there is no an unique decomposition (like Hoeffding decomposition), hard to decouple interactions and correlation effects . . .

- we can build metamodels when the initial model is too costly;
- functionals inputs: MC approaches can be applied;
- vectorial output: approaches component by component, or generalised indices GSI (Lamboni *et al.*;
- functional output : we summarize the output by a vector (projection on an appropriate basis) or we make a movie (temporal output) or a map (spatial output) of indices;
- correlated inputs: there is no an unique decomposition (like Hoeffding decomposition), hard to decouple interactions and correlation effects . . .

- ⋆ we can build metamodels when the initial model is too costly;
- ⋆ functionals inputs: MC approaches can be applied;
- ⋆ vectorial output: approaches component by component, or generalised indices GSI (Lamboni *et al.*;
- ⋆ functional output : we summarize the output by a vector (projection on an appropriate basis) or we make a movie (temporal output) or a map (spatial output) of indices;
- ⋆ correlated inputs: there is no an unique decomposition (like Hoeffding decomposition), hard to decouple interactions and correlation effects . . .

⋆ we can build metamodels when the initial model is too costly;

⋆ functionals inputs: MC approaches can be applied;

⋆ vectorial output: approaches component by component, or generalised indices GSI (Lamboni *et al.*;

⋆ functional output : we summarize the output by a vector (projection on an appropriate basis) or we make a movie (temporal output) or a map (spatial output) of indices;

⋆ correlated inputs: there is no an unique decomposition (like Hoeffding decomposition), hard to decouple interactions and correlation effects . . .

A short bibliography I

A. Antoniadis.
Analysis of variance on function spaces.
Math. Oper. Forsch. und Statist., series Statistics, 15(1):59–71, 1984.

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

R. I. Cukier, C. M. Fortuin, K. E. Shuler, A. G. Petschek, and J. H. Schaibly.
Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients: Theory.
Journal of Chemical Physics, 59:3873–3878, 1973.

R. I. Cukier, H. B. Levine, and K. E. Shuler.
Nonlinear sensitivity analysis of multiparameter model systems.
Journal of Computational Physics, 26:1–42, 1978.

R. I. Cukier, J. H. Schaibly, and K. E. Shuler.
Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients: Analysis of the approximations.
Journal of Chemical Physics, 63:1140–1149, 1975.

F. Gamboa, A. Janon, T. Klein, A. Lagnoux, and C. Prieur.
Statistical inference for sobol pick freeze monte carlo method.
http://hal.archives-ouvertes.fr/hal-00804668, 2013+.

W. F. Hoeffding.
A class of statistics with asymptotically normal distributions.
Annals of Mathematical Statistics, 19:293–325, 1948.

Bertrand Iooss.
Revue sur l'analyse de sensibilité globale de modèles numériques.
Journal de la Société Française de Statistique, 152(1):3–25, 2011.

M. Lamboni, B. Iooss, A.-L. Popelin, and F. Gamboa.
Derivative-based global sensitivity measures: general links with Sobol' indices and numerical tests.
Mathematics and Computers in Simulation, 87:45–54, 2013.

W. Mauntz.
Global sensitivity analysis of general nonlinear systems.
Master's Thesis, Imperial College. Supervisors: C. Pantelides and S. Kucherenko, 2002.

A. Saltelli.
Making best use of model evaluations to compute sensitivity indices.
Computer Physics Communications, 145:280–297, 2002.

A. Saltelli, K. Chan, and E. M. Scott.
Sensitivity Analysis.
John Wiley & Sons, 2000.

A short bibliography III

- I. M. Sobol'.
  Sensitivity analysis for nonlinear mathematical models.
  Mathematical Modeling and Computational Experiment, 1:407–414, 1993.
- I. M. Sobol' and A Gresham.
  On an alternative global sensitivity estimators.
  Proceedings of SAMO, Belgirate, pages 40–42, 1995.
- I. M. Sobol' and S. Kucherenko.
  Derivative based global sensitivity measures and the link with global sensitivity indices.
  Mathematics and Computers in Simulation, 79:3009–3017, 2009.
- S. Tarantola, D. Gatelli, and T. A. Mara.
  Random balance designs for the estimation of first-order global sensitivity indices.
  Reliability Engineering and System Safety, 91:717–727, 2006.
- J. Y. Tissot and C. Prieur.
  Bias correction for the estimation of sensitivity indices based on random balance designs.
  Reliability Engineering and System Safety, 107:205–213, 2012.
- J. Y. Tissot and C. Prieur.
  Variance-based sensitivity analysis using harmonic analysis.
  `http://hal.archives-ouvertes.fr/docs/00/68/07/25/PDF`, 2012+.

A short bibliography IV

J. Y. Tissot and C. Prieur.
A randomized orthogonal array-based procedure for the estimation of first- and second-order sobol' indices.
Journal of Statistical Computation and Simulation, 85:1358–1381, 2015.