Derivative-based global sensitivity measures for interactions

Olivier Roustant

École Nationale Supérieure des Mines de Saint-Étienne

Joint work with J. Fruth, B. Iooss and S. Kuhnt

—

SAMO 13

Outline

1. **Background and motivation**
   - Variance-based and derivative-based sensitivity measures
   - 1st-order analysis. Screening : Total indices & DGSM.
   - 2nd-order analysis. Interaction screening. Crossed DGSM
2. **The main result : A link between superset importance and crossed DGSM**
3. **Applications**
   - A 6-dimensional example
   - When the gradient is supplied

Let $X = (X_1, \ldots, X_d)$ be a vector of independent input variables with distribution $\mu_1 \otimes \cdots \otimes \mu_d$, and $g : \Delta \subseteq \mathbb{R}^d \to \mathbb{R}$ such that $g(\mathbf{X}) \in L^2(\mu)$.

Sobol-Hoeffding decomposition [Sobol, 1993, Efron and Stein, 1981]

$$\begin{aligned} g(\mathbf{X}) &= g_0 + \sum_{i=1}^{d} g_i(X_i) + \sum_{1 \leq i < j \leq d} g_{i,j}(X_i, X_j) + \cdots + g_{1,\ldots,d}(X_1, \ldots, X_d) \\ &= \sum_{I \subseteq \{1,\ldots,d\}} g_I(\mathbf{X}_I) \end{aligned} \tag{1}$$

The g_I's are centered and orthogonal.

Global sensitivity measures

Variance-based measures

- Partial variances : $D_I = \mathrm{var}(g_I(X_I))$, and **Sobol indices** $S_I = D_I/D$

$$D := \mathrm{var}(g(\mathbf{X})) = \sum_I D_I, \qquad 1 = \sum_I S_I$$

Global sensitivity measures

Variance-based measures

- Partial variances : $D_I = \text{var}(g_I(X_I))$, and **Sobol indices** $S_I = D_I / D$

$$D := \text{var}(g(\mathbf{X})) = \sum_I D_I, \qquad 1 = \sum_I S_I$$

- **Total index** : $D_i^{\mathrm{T}} = \sum_{J \supseteq \{i\}} D_J, \qquad S_i^{\mathrm{T}} = \frac{D_i^{\mathrm{T}}}{D}$.

Global sensitivity measures

Variance-based measures

- Partial variances : $D_I = \mathrm{var}(g_I(X_I))$, and **Sobol indices** $S_I = D_I/D$

$$D := \mathrm{var}(g(\mathbf{X})) = \sum_I D_I, \qquad 1 = \sum_I S_I$$

- **Total index** : $D_i^{\mathrm{T}} = \sum_{J \supseteq \{i\}} D_J, \qquad S_i^{\mathrm{T}} = \frac{D_i^{\mathrm{T}}}{D}$.
- **Superset importance** [Liu and Owen, 2006] :
  $D_I^{\mathrm{super}} := \sum_{J \supseteq I} D_J, \qquad S_I^{\mathrm{super}} = \frac{D_I^{\mathrm{super}}}{D}$
  $\rightarrow D_{i,j}^{\mathrm{super}} := \sum_{J \supseteq \{i,j\}} D_J$ **"total interaction index"** [Fruth et al., 2013]

Global sensitivity measures

Variance-based measures

- Partial variances : $D_I = \text{var}(g_I(X_I))$, and **Sobol indices** $S_I = D_I/D$

$$D := \text{var}(g(\mathbf{X})) = \sum_I D_I, \qquad 1 = \sum_I S_I$$

- **Total index** : $D_i^{\text{T}} = \sum_{J \supseteq \{i\}} D_J, \qquad S_i^{\text{T}} = \frac{D_i^{\text{T}}}{D}.$
- **Superset importance** [Liu and Owen, 2006] :
  $D_I^{\text{super}} := \sum_{J \supseteq I} D_J, \qquad S_I^{\text{super}} = \frac{D_I^{\text{super}}}{D}$
  $\rightarrow D_{i,j}^{\text{super}} := \sum_{J \supseteq \{i,j\}} D_J$ **"total interaction index"** [Fruth et al., 2013]

Derivative-based measures

- In sensitivity analysis [Sobol and Gresham, 1995] :
  $\nu_i = \int \left(\frac{\partial g(\mathbf{x})}{\partial x_i} \right)^2 d\mu(\mathbf{x})$, called **DGSM**

Global sensitivity measures

Variance-based measures

- Partial variances : $D_I = \mathrm{var}(g_I(X_I))$, and **Sobol indices** $S_I = D_I / D$

$$D := \mathrm{var}(g(\mathbf{X})) = \sum_I D_I, \qquad 1 = \sum_I S_I$$

- **Total index** : $D_i^{\mathrm{T}} = \sum_{J \supseteq \{i\}} D_J, \qquad S_i^{\mathrm{T}} = \frac{D_i^{\mathrm{T}}}{D}.$
- **Superset importance** [Liu and Owen, 2006] :
  $D_I^{\text{super}} := \sum_{J \supseteq I} D_J, \qquad S_I^{\text{super}} = \frac{D_I^{\text{super}}}{D}$
  $\rightarrow D_{i,j}^{\text{super}} := \sum_{J \supseteq \{i,j\}} D_J$ **"total interaction index"** [Fruth et al., 2013]

Derivative-based measures

- In sensitivity analysis [Sobol and Gresham, 1995] :
  $\nu_i = \int \left(\frac{\partial g(\mathbf{x})}{\partial x_i} \right)^2 d\mu(\mathbf{x})$, called **DGSM**
- In statistical learning [Friedman and Popescu, 2008] :
  $\nu_{i,j} = \int \left(\frac{\partial^2 g(\mathbf{x})}{\partial x_i \partial x_j} \right)^2 d\mu(\mathbf{x}), \quad \nu_I = \int \left(\frac{\partial^{|I|} g(\mathbf{x})}{\partial \mathbf{x}_I} \right)^2 d\mu(\mathbf{x}).$ **"crossed DGSM"**.

First-order analysis considers single variables. One aim is **screening** : detection of non-influential input variables.

First-order analysis considers single variables. One aim is **screening** : detection of non-influential input variables.

Screening with total indices or DGSMs

- If either $D_i^{\mathrm{T}} = 0$ or $\nu_i = 0$, than X_i is non influential.

First-order analysis considers single variables. One aim is **screening** : detection of non-influential input variables.

Screening with total indices or DGSMs

- If either $D_i^{\mathrm{T}} = 0$ or $\nu_i = 0$, than X_i is non influential.
- There is a Poincaré-type inequality between total indices and DGSMs

$$D_i \leq D_i^{\mathrm{T}} \leq C(\mu_i)\nu_i$$

$\rightarrow$ *Proved by [Sobol and Kucherenko, 2009] for the uniform and normal distributions, [Lamboni et al., 2013] for the general case .*

Poincaré inequality (1-dimensional case)

A distribution μ satisfies a Poincaré inequality if for all h in $L^2(\mu)$ such that $\int h(x)d\mu(x) = 0$, and $h'(x) \in L^2(\mu)$:

$$\int h(x)^2 d\mu(x) \leq C(\mu) \int h'(x)^2 d\mu(x)$$

The best constant is denoted $C_{\text{opt}}(\mu)$.

Poincaré inequality (1-dimensional case)

A distribution μ satisfies a Poincaré inequality if for all h in $L^2(\mu)$ such that $\int h(x) d\mu(x) = 0$, and $h'(x) \in L^2(\mu)$:

$$\int h(x)^2 d\mu(x) \leq C(\mu) \int h'(x)^2 d\mu(x)$$

The best constant is denoted $C_{\text{opt}}(\mu)$.

Examples ([Sobol and Kucherenko, 2009], [Ané et al., 2000], [Lamboni et al., 2013], [Bobkov and Houdré, 1997, Bobkov, 1999])

Distribution	$C_{\text{opt}}(\mu)$	*A case of equality*
Uniform $\mathcal{U}[a,b]$	$(b-a)^2/\pi^2$	$g(x) = \cos\left(\frac{\pi(x-a)}{b-a}\right)$
Normal $\mathcal{N}(\mu,\sigma^2)$	σ^2	$g(x) = x - \mu$

Poincaré inequality (1-dimensional case)

A distribution μ satisfies a Poincaré inequality if for all h in $L^2(\mu)$ such that $\int h(x)d\mu(x) = 0$, and $h'(x) \in L^2(\mu)$:

$$\int h(x)^2 d\mu(x) \leq C(\mu) \int h'(x)^2 d\mu(x)$$

The best constant is denoted $C_{\text{opt}}(\mu)$.

Examples ([Sobol and Kucherenko, 2009], [Ané et al., 2000], [Lamboni et al., 2013], [Bobkov and Houdré, 1997, Bobkov, 1999])

Distribution	$C_{\text{opt}}(\mu)$	*A case of equality*
Uniform $\mathcal{U}[a, b]$	$(b-a)^2/\pi^2$	$g(x) = \cos\left(\frac{\pi(x-a)}{b-a}\right)$
Normal $\mathcal{N}(\mu, \sigma^2)$	σ^2	$g(x) = x - \mu$

Properties of μ	*A Poincaré constant* $C(\mu)$
Continuous	$4\left[\sup_{x\in\mathbb{R}} \frac{\min(F(x), 1-F(x))}{f(x)}\right]^2$
log-concave	$1/f(m)^2$
log-concave, truncated on $[a, b]$	$(F(b) - F(a))^2 / f\left(q\left(\frac{F(a)+F(b)}{2}\right)\right)^2$

Second-order analysis considers *pairs* of variables. It allows *interaction screening* : To detect $\{X_i, X_j\}$ that do not interact together ($D_{i,j}^{\text{super}} = 0$).

Second-order analysis considers *pairs* of variables. It allows *interaction screening* : To detect $\{X_i, X_j\}$ that do not interact together ($D_{i,j}^{\text{super}} = 0$).

2nd-order analysis and additive structures

If either $\nu_{i,j} = 0$ or $D_{i,j}^{\text{super}} = 0$, then g can be written as a sum of two functions, one that does not depend on x_i, the other that does not depend on x_j [Hooker, 2004, Friedman and Popescu, 2008] :

$$g(\mathbf{x}) = g_{-i}(\mathbf{x}_{-i}) + g_{-j}(\mathbf{x}_{-j})$$

Second-order analysis considers *pairs* of variables. It allows *interaction screening* : To detect $\{X_i, X_j\}$ that do not interact together ($D_{i,j}^{\text{super}} = 0$).

2nd-order analysis and additive structures

If either $\nu_{i,j} = 0$ or $D_{i,j}^{\text{super}} = 0$, then g can be written as a sum of two functions, one that does not depend on x_i, the other that does not depend on x_j [Hooker, 2004, Friedman and Popescu, 2008] :

$$g(\mathbf{x}) = g_{-i}(\mathbf{x}_{-i}) + g_{-j}(\mathbf{x}_{-j})$$

- [Hooker, 2004] uses it in machine learning
- [Muehlenstaedt et al., 2012] use it in computer experiments (see after).

An application of 2nd-order analysis

Here, the estimated interaction structure is used to define a suitable covariance kernel for the Ishigami function

$$f(x_1, x_2, x_3) = \sin(x_1) + 7\sin(x_2)^2 + 0.1x_3^4\sin(x_1)$$

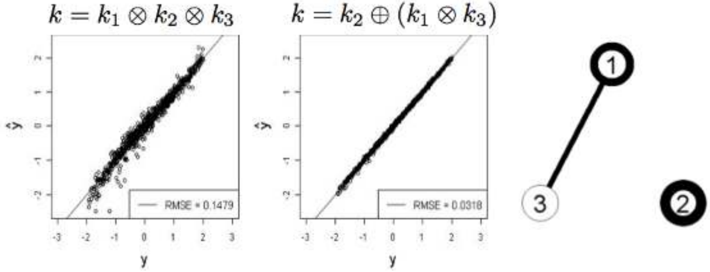

FIGURE: Left : Kriging with standard kernel ; Middle : Kriging with a block-additive structure estimated from the FANOVA graph (right). See [Muehlenstaedt et al., 2012] for more details.

Theorem - A link between superset importance and crossed DGSM

Assume that all μ_i ($i = 1, \dots, d$) satisfy a Poincaré inequality. Then for all pairs $\{i, j\}$ ($1 \leq i, j \leq n$),

$$D_{i,j} \leq D_{i,j}^{\text{super}} \leq C(\mu_i)C(\mu_j)\nu_{i,j}.$$

and $C_{\text{opt}}(\mu_i)C_{\text{opt}}(\mu_j)$ is the best constant. Generalizes to more than pairs.

Theorem - A link between superset importance and crossed DGSM

Assume that all μ_i ($i = 1, \dots, d$) satisfy a Poincaré inequality. Then for all pairs $\{i, j\}$ ($1 \leq i, j \leq n$),

$$D_{i,j} \leq D_{i,j}^{\text{super}} \leq C(\mu_i)C(\mu_j)\nu_{i,j}.$$

and $C_{\text{opt}}(\mu_i)C_{\text{opt}}(\mu_j)$ is the best constant. Generalizes to more than pairs.

Sketch of Proof. Denote $g_{i,j}^{\text{super}}(\mathbf{x}) := \sum_{J \supseteq \{i,j\}} g_J(\mathbf{x}_J)$. Then :

Theorem - A link between superset importance and crossed DGSM

Assume that all μ_i ($i = 1, \dots, d$) satisfy a Poincaré inequality. Then for all pairs $\{i, j\}$ ($1 \leq i, j \leq n$),

$$D_{i,j} \leq D_{i,j}^{\text{super}} \leq C(\mu_i)C(\mu_j)\nu_{i,j}.$$

and $C_{\text{opt}}(\mu_i)C_{\text{opt}}(\mu_j)$ is the best constant. Generalizes to more than pairs.

Sketch of Proof. Denote $g_{i,j}^{\text{super}}(\mathbf{x}) := \sum_{J \supseteq \{i,j\}} g_J(\mathbf{x}_J)$. Then :

1. $D_{i,j}^{\text{super}} = \text{var}(g_{i,j}^{\text{super}}(\mathbf{x})) = \int \left(g_{i,j}^{\text{super}}(\mathbf{x})\right)^2 d\mu(\mathbf{x})$

Theorem - A link between superset importance and crossed DGSM

Assume that all μ_i ($i = 1, \dots, d$) satisfy a Poincaré inequality. Then for all pairs $\{i, j\}$ ($1 \leq i, j \leq n$),

$$D_{i,j} \leq D_{i,j}^{\text{super}} \leq C(\mu_i)C(\mu_j)\nu_{i,j}.$$

and $C_{\text{opt}}(\mu_i)C_{\text{opt}}(\mu_j)$ is the best constant. Generalizes to more than pairs.

Sketch of Proof. Denote $g_{i,j}^{\text{super}}(\mathbf{x}) := \sum_{J \supseteq \{i,j\}} g_J(\mathbf{x}_J)$. Then :

1. $D_{i,j}^{\text{super}} = \text{var}(g_{i,j}^{\text{super}}(\mathbf{x})) = \int \left(g_{i,j}^{\text{super}}(\mathbf{x})\right)^2 d\mu(\mathbf{x})$
2. $\nu_{i,j} = \int \left(\frac{\partial^2 g(\mathbf{x})}{\partial x_i \partial x_j}\right)^2 d\mu(\mathbf{x}) = \int \left(\frac{\partial^2 g_{i,j}^{\text{super}}(\mathbf{x})}{\partial x_i \partial x_j}\right)^2 d\mu(\mathbf{x})$

Theorem - A link between superset importance and crossed DGSM

Assume that all μ_i ($i = 1, \dots, d$) satisfy a Poincaré inequality. Then for all pairs $\{i, j\}$ ($1 \leq i, j \leq n$),

$$D_{i,j} \leq D_{i,j}^{\text{super}} \leq C(\mu_i)C(\mu_j)\nu_{i,j}.$$

and $C_{\text{opt}}(\mu_i)C_{\text{opt}}(\mu_j)$ is the best constant. Generalizes to more than pairs.

Sketch of Proof. Denote $g_{i,j}^{\text{super}}(\mathbf{x}) := \sum_{J \supseteq \{i,j\}} g_J(\mathbf{x}_J)$. Then :

1. $D_{i,j}^{\text{super}} = \text{var}(g_{i,j}^{\text{super}}(\mathbf{x})) = \int \left(g_{i,j}^{\text{super}}(\mathbf{x})\right)^2 d\mu(\mathbf{x})$
2. $\nu_{i,j} = \int \left(\frac{\partial^2 g(\mathbf{x})}{\partial x_i \partial x_j}\right)^2 d\mu(\mathbf{x}) = \int \left(\frac{\partial^2 g_{i,j}^{\text{super}}(\mathbf{x})}{\partial x_i \partial x_j}\right)^2 d\mu(\mathbf{x})$

Finally combine (by integrating) the two Poincaré inequalities :

$$\int \left(g_{i,j}^{\text{super}}(\mathbf{x})\right)^2 d\mu_i(x_i) \leq C(\mu_i) \int \left(\frac{\partial g_{i,j}^{\text{super}}(\mathbf{x})}{\partial x_i}\right)^2 d\mu_i(x_i)$$

$$\int \left(\frac{\partial g_{i,j}^{\text{super}}(\mathbf{x})}{\partial x_i}\right)^2 d\mu_j(x_j) \leq C(\mu_j) \int \left(\frac{\partial}{\partial x_j}\frac{\partial g_{i,j}^{\text{super}}(\mathbf{x})}{\partial x_i}\right)^2 d\mu_j(x_j)$$

Applications

In applications, we would base the results on the upper bounds :

$$U_i := C(\mu_i)\frac{\nu_i}{D} \geq S_i^T$$

$$U_{i,j} := C(\mu_i)C(\mu_j)\frac{\nu_{i,j}}{D} \geq S_{i,j}^{\text{super}}$$

Applications

In applications, we would base the results on the upper bounds :

$$\begin{aligned} U_i &:= C(\mu_i)\frac{\nu_i}{D} &\geq S_i^T \\ U_{i,j} &:= C(\mu_i)C(\mu_j)\frac{\nu_{i,j}}{D} &\geq S_{i,j}^{\text{super}} \end{aligned}$$

The estimation of D_i^T and $D_{i,j}^{\text{super}}$ can be done by MC (or QMC) from [Jansen, 1999] [Liu and Owen, 2006] :

$$D_i^T = \tfrac{1}{2}\int \left[f(\mathbf{x}) - f(z_i, \mathbf{x}_{-\mathbf{i}})\right]^2 d\mu(\mathbf{x}) d\mu_i(z_i)$$

$$D_{i,j}^{\text{super}} = \frac{1}{4}\int \Big[f(\mathbf{x}) - f(x_i, z_j, \mathbf{x}_{-\mathbf{i,j}}) - f(z_i, x_j, \mathbf{x}_{-\mathbf{i,j}}) + f(z_i, z_j, \mathbf{x}_{-\mathbf{i,j}})\Big]^2 d\mu(\mathbf{x}) d\mu_i(z_i) d\mu_j(z_j)$$

A 6-dimensional example

We consider the 6-dimensional function in L^2 ([Muehlenstaedt et al., 2012]) :

$$a(X_1, \dots, X_6) = \cos([1, X_1, X_5, X_3]\phi) + \sin([1, X_4, X_2, X_6]\gamma)$$

with $\phi = [-0.8, -1.1, 1.1, 1]^T$, $\gamma = [-0.5, 0.9, 1, -1.1]^T$, and where $X_1, \dots, X_6$ are assumed i.i.d uniform on $[-1, 1]$.

A 6-dimensional example

We consider the 6-dimensional function in L^2 ([Muehlenstaedt et al., 2012]) :

$$a(X_1, \dots, X_6) = \cos([1, X_1, X_5, X_3]\phi) + \sin([1, X_4, X_2, X_6]\gamma)$$

with $\phi = [-0.8, -1.1, 1.1, 1]^T$, $\gamma = [-0.5, 0.9, 1, -1.1]^T$, and where $X_1, \dots, X_6$ are assumed i.i.d uniform on $[-1, 1]$.

First-order analysis

Input	S_i	S_i^T	$\hat{S}_i^T$	sd	U_i	$\hat{U}_i$	sd
X_1	0.11	0.231	0.231	(0.012)	0.329	0.329	(0.007)
X_2	0.143	0.214	0.215	(0.009)	0.272	0.285	(0.005)
X_3	0.086	0.196	0.197	(0.01)	0.272	0.272	(0.006)
X_4	0.112	0.176	0.176	(0.008)	0.22	0.231	(0.004)
X_5	0.11	0.231	0.232	(0.011)	0.329	0.329	(0.007)
X_6	0.18	0.256	0.256	(0.011)	0.329	0.345	(0.007)

$\rightarrow$ *Screening does not discard any inputs here. Ranking is different (cf. [Sobol and Kucherenko, 2009])*

Second-order analysis

Inputs pair	$S_{i,j}$	$S_{i,j}^{\text{super}}$	$\hat{S}_{i,j}^{\text{super}}$	*sd*	$U_{i,j}$	$\hat{U}_{i,j}$	*sd*
$X_1 : X_2$	0	0	0	(0)	0	0	(0)
$X_1 : X_3$	0.043	0.067	0.067	(0.005)	0.133	0.132	(0.003)
$X_1 : X_4$	0	0	0	(0)	0	0	(0)
$X_1 : X_5$	0.055	0.078	0.08	(0.006)	0.161	0.16	(0.004)
$X_1 : X_6$	0	0	0	(0)	0	0	(0)
$X_2 : X_3$	0	0	0	(0)	0	0	(0)
$X_2 : X_4$	0.018	0.04	0.039	(0.004)	0.085	0.085	(0.002)
$X_2 : X_5$	0	0	0	(0)	0	0	(0)
$X_2 : X_6$	0.031	0.053	0.052	(0.005)	0.127	0.127	(0.003)
$X_3 : X_4$	0	0	0	(0)	0	0	(0)
$X_3 : X_5$	0.043	0.067	0.067	(0.005)	0.133	0.132	(0.003)
$X_3 : X_6$	0	0	0	(0)	0	0	(0)
$X_4 : X_5$	0	0	0	(0)	0	0	(0)
$X_4 : X_6$	0.024	0.046	0.045	(0.004)	0.103	0.103	(0.002)
$X_5 : X_6$	0	0	0	(0)	0	0	(0)

Visualization of the second-order analysis

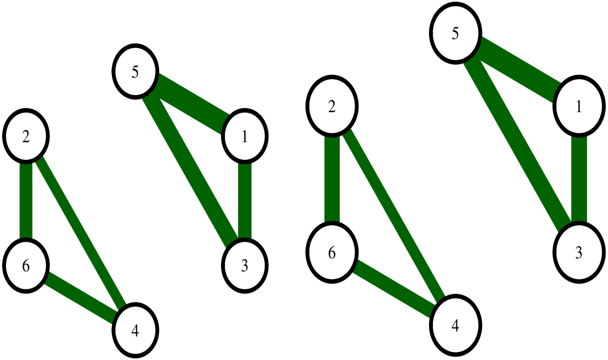

FIGURE: FANOVA graphs of $\hat{S}_{i,j}^{\text{super}}$ (left) and $\hat{U}_{i,j}$ (right).

An advantageous situation : When the gradient is supplied

Suppose that one run gives *both* $f(\mathbf{x})$ *and* $\nabla f(\mathbf{x}) = \left(\frac{\partial f(\mathbf{x})}{\partial x_i}\right)_{1 \leq i \leq d}$.
Then crossed DGSM estimation requires $\approx d/2$ fewer evaluations :

	Function only	Gradient supplied
Total indices	$(d+1)N$	no change
DGSMs	$(d+1)N$	N
superset importance	$(d+1+\frac{d(d-1)}{2})N$	no change
crossed DGSMs	$(d+1+\frac{d(d-1)}{2})N$	$(d+1)N$

TABLE: Computational cost for Monte Carlo estimation. N is the sample size.

Convergence study when the gradient is supplied

Example for the 6-dimensional function ‘a’

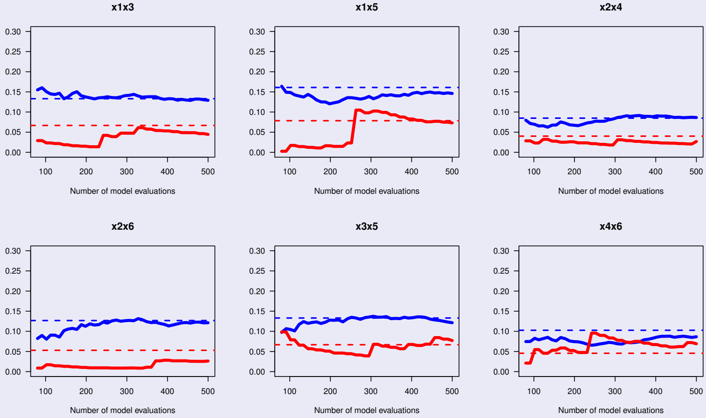

FIGURE: Blue : $U_{i,j}$ (upper bounds for crossed DGSM) ; Red : $S_{i,j}^{\text{super}}$.
Dotted : True value ; Solid : MC estimates.

Conclusion

This work is about *2nd-order analysis*, which considers *pairs* of inputs.

- There is a Poincaré-type inequality between **superset importance** (total interaction index) & **crossed DGSM** :

$$D_{i,j} \leq D_{i,j}^{\text{super}} \leq C(\mu_i)C(\mu_j) \int \left(\frac{\partial^2 g(\mathbf{x})}{\partial x_i \partial x_j} \right)^2 d\mu(\mathbf{x})$$

Conclusion

This work is about *2nd-order analysis*, which considers *pairs* of inputs.

- There is a Poincaré-type inequality between **superset importance** (total interaction index) & **crossed DGSM** :

$$D_{i,j} \leq D_{i,j}^{\text{super}} \leq C(\mu_i)C(\mu_j) \int \left(\frac{\partial^2 g(\mathbf{x})}{\partial x_i \partial x_j} \right)^2 d\mu(\mathbf{x})$$

- Crossed DGSM can be used for ***interaction screening*** $\rightarrow$ *detection of additive structures.*

Conclusion

This work is about *2nd-order analysis*, which considers *pairs* of inputs.

- There is a Poincaré-type inequality between **superset importance** (total interaction index) & **crossed DGSM** :

$$D_{i,j} \leq D_{i,j}^{\text{super}} \leq C(\mu_i)C(\mu_j) \int \left(\frac{\partial^2 g(\mathbf{x})}{\partial x_i \partial x_j} \right)^2 d\mu(\mathbf{x})$$

- Crossed DGSM can be used for ***interaction screening*** $\rightarrow$ *detection of additive structures.*
- Crossed DGSM are especially **useful when the gradient is supplied**.

Conclusion

This work is about *2nd-order analysis*, which considers *pairs* of inputs.

- There is a Poincaré-type inequality between **superset importance** (total interaction index) & **crossed DGSM** :

$$D_{i,j} \leq D_{i,j}^{\text{super}} \leq C(\mu_i)C(\mu_j) \int \left(\frac{\partial^2 g(\mathbf{x})}{\partial x_i \partial x_j} \right)^2 d\mu(\mathbf{x})$$

- Crossed DGSM can be used for ***interaction screening*** $\rightarrow$ *detection of additive structures.*
- Crossed DGSM are especially **useful when the gradient is supplied**.
- **Limitations :** As for DGSM, crossed DGSM must NOT be used to rank interactions. They may be give poor results for functions with sharp variations, as well as when some of the $C(\mu_i)$'s are large.

Ané, C., Blachère, S., Chafaï, D., Fougères, P., Gentil, I., Malrieu, F., Roberto, C., and Scheffer, G. (2000).
Sur les inégalités de Sobolev logarithmiques, volume 10 of *Panoramas et Synthèses*.
Société Mathématique de France, Paris.

Bobkov, S. G. (1999).
Isoperimetric and analytic inequalities for log-concave probability measures.
The Annals of Probability, 27(4) :1903–1921.

Bobkov, S. G. and Houdré, C. (1997).
Isoperimetric constants for product probability measures.
The Annals of Probability, 25(1) :184–205.

Efron, B. and Stein, C. (1981).
The jackknife estimate of variance.
The Annals of Statistics, 9(3) :586–596.

Friedman, J. and Popescu, B. (2008).
Predictive Learning via Rule Ensembles.
The Annals of Applied Statistics, 2(3) :916–954.

Fruth, J., Roustant, O., and Kuhnt, S. (2013).
Total interaction index : A variance-based sensitivity index for interaction screening.
Submitted.

Hooker, G. (2004).
Discovering additive structure in black box functions.
In *Proceedings of KDD 2004*, pages 575–580. ACM DL.

Jansen, M. (1999).
Analysis of variance designs for model output.
Computer Physics Communication, 117 :25–43.

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

Liu, R. and Owen, A. (2006).
Estimating mean dimensionality of analysis of variance decompositions.
Journal of the American Statistical Association, 101(474) :712–721.

Muehlenstaedt, T., Roustant, O., Carraro, L., and Kuhnt, S. (2012).

Data-driven Kriging models based on FANOVA-decomposition.
Statistics & Computing, 22 :723–738.

Sobol, I. (1993).
Sensitivity estimates for non linear mathematical models.
Mathematical Modelling and Computational Experiments, 1 :407–414.

Sobol, I. and Gresham, A. (1995).
On an alternative global sensitivity estimators.
In *Proceedings of SAMO 1995*, pages 40–42, Belgirate.

Sobol, I. and Kucherenko, S. (2009).
Derivative-based global sensitivity measures and the link with global sensitivity indices.
Mathematics and Computers in Simulation, 79 :3009–3017.