

Contributions à l'analyse de sensibilité

Christian Lavergne

Institut de Mathématiques et de Modélisation de Montpellier

I3M, UMR CNRS 5149

juillet 2006

- Contexte
 - modèles de type « boîte noire » (code de calcul informatique) :
 - déterministes,
 - entrées et sorties aléatoires,
 - modélisation de phénomènes complexes.
- Objectifs principaux
 - améliorer la prédiction (réduction de l'incertitude),
 - simplifier ou alléger le modèle (fixer les variables les moins influentes),
 - améliorer la compréhension du phénomène.
- Domaines d'applications
 - ingénierie nucléaire,
 - écologie et environnement, chimie ...

- 1 Analyse de sensibilité
 - Locale et globale
 - Indices de sensibilité : ex à 1,2 ou 3 variables
 - Indices de sensibilité
 - Estimation
- 2 Compléments
- 3 Incertitude de modèle et analyse de sensibilité
 - Prise en compte d'une mutation de modèle
 - Utilisation d'un modèle simplifié
- 4 Modèles à entrées non indépendantes
 - Le problème...
 - Indices de sensibilité : 2 variables dépendantes
 - Indices de sensibilité multidimensionnels
 - Applications au code de calcul Stay'SL

- 1 Analyse de sensibilité
 - Locale et globale
 - Indices de sensibilité : ex à 1,2 ou 3 variables
 - Indices de sensibilité
 - Estimation
- 2 Compléments
- 3 Incertitude de modèle et analyse de sensibilité
 - Prise en compte d'une mutation de modèle
 - Utilisation d'un modèle simplifié
- 4 Modèles à entrées non indépendantes
 - Le problème...
 - Indices de sensibilité : 2 variables dépendantes
 - Indices de sensibilité multidimensionnels
 - Applications au code de calcul Stay'SL

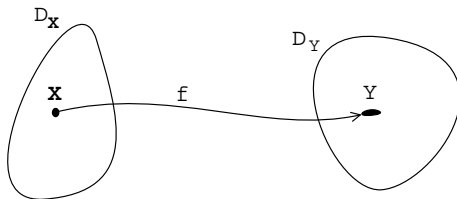
- 1 Analyse de sensibilité
 - Locale et globale
 - Indices de sensibilité : ex à 1,2 ou 3 variables
 - Indices de sensibilité
 - Estimation
- 2 Compléments
- 3 Incertitude de modèle et analyse de sensibilité
 - Prise en compte d'une mutation de modèle
 - Utilisation d'un modèle simplifié
- 4 Modèles à entrées non indépendantes
 - Le problème...
 - Indices de sensibilité : 2 variables dépendantes
 - Indices de sensibilité multidimensionnels
 - Applications au code de calcul Stay'SL

- 1 Analyse de sensibilité
 - Locale et globale
 - Indices de sensibilité : ex à 1,2 ou 3 variables
 - Indices de sensibilité
 - Estimation
- 2 Compléments
- 3 Incertitude de modèle et analyse de sensibilité
 - Prise en compte d'une mutation de modèle
 - Utilisation d'un modèle simplifié
- 4 Modèles à entrées non indépendantes
 - Le problème...
 - Indices de sensibilité : 2 variables dépendantes
 - Indices de sensibilité multidimensionnels
 - Applications au code de calcul Stay'SL

- 1 Analyse de sensibilité
 - Locale et globale
 - Indices de sensibilité : ex à 1,2 ou 3 variables
 - Indices de sensibilité
 - Estimation
- 2 Compléments
- 3 Incertitude de modèle et analyse de sensibilité
 - Prise en compte d'une mutation de modèle
 - Utilisation d'un modèle simplifié
- 4 Modèles à entrées non indépendantes
 - Le problème...
 - Indices de sensibilité : 2 variables dépendantes
 - Indices de sensibilité multidimensionnels
 - Applications au code de calcul Stay'SL

Analyse de sensibilité

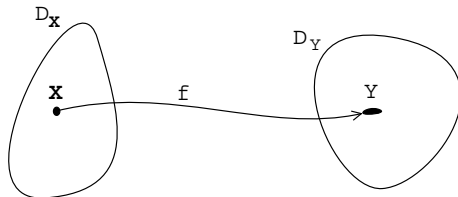
Soit le modèle $Y = f(\mathbf{X})$



L'analyse de sensibilité étudie comment des perturbations sur \mathbf{X} engendrent des perturbations sur Y

Analyse de sensibilité

Soit le modèle $Y = f(\mathbf{X})$



L'analyse de sensibilité étudie comment des perturbations sur \mathbf{X} engendrent des perturbations sur Y

AS locale $\mathbf{x} \dashrightarrow Y$ valeur

AS globale $D_{\mathbf{X}} \dashrightarrow D_Y$ variabilité

Analyse de sensibilité globale ...

... consiste à déterminer la part de variabilité de la réponse du modèle due à un sous ensemble de variables d'entrée.

Indicateurs utilisés :

les indices de sensibilité

- 1 Analyse de sensibilité
 - Locale et globale
 - Indices de sensibilité : ex à 1,2 ou 3 variables
 - Indices de sensibilité
 - Estimation
- 2 Compléments
- 3 Incertitude de modèle et analyse de sensibilité
 - Prise en compte d'une mutation de modèle
 - Utilisation d'un modèle simplifié
- 4 Modèles à entrées non indépendantes
 - Le problème...
 - Indices de sensibilité : 2 variables dépendantes
 - Indices de sensibilité multidimensionnels
 - Applications au code de calcul Stay'SL

Indices de sensibilité à 1,2 ou 3 variables

- 1 variable (inutile) :

$$Y = f(X) \text{ alors } Y \text{ est aussi } = E(f(X)/X)$$

$$\text{et donc } V(Y) = V(E(Y/X)).$$

$$\text{Rappel : } V(Y) = V(E(Y/X)) + E(V(Y/X))$$

Indices de sensibilité à 1,2 ou 3 variables

- 2 variables indépendantes ($X_1 \perp X_2$) :

$$Y = f(X_1, X_2) = E(Y/X_1, X_2).$$

Alors Y peut se décomposer sous la forme :

$$\begin{aligned} Y &= E(Y) && \rightarrow f_0 \\ &+ E(Y/X_1) - E(Y) && \rightarrow f_1 \\ &+ E(Y/X_2) - E(Y) && \rightarrow f_2 \\ &+ E(Y/X_1, X_2) - E(Y/X_1) - E(Y/X_2) + E(Y) && \rightarrow f_{12} \end{aligned}$$

Indices de sensibilité à 1,2 ou 3 variables

f_0 est une constante ;

f_1 , f_2 et f_{12} sont 3 variables aléatoires d'espérance nulle et \perp .

Donc $V(Y) = V(f_1) + V(f_2) + V(f_{12})$ mais

$$V(f_1) = V(E(Y/X_1))$$

donc

$$\frac{V(E(Y/X_1))}{V(Y)}$$

sera donc la sensibilité au 1^{er} ordre de X_1 sur la variance de Y .

Indices de sensibilité à 1,2 ou 3 variables

- 3 variables indépendantes ($X_1 \perp X_2 \perp X_3$) :

$$Y = f(X_1, X_2, X_3) = E(Y/X_1, X_2, X_3)$$

Alors Y peut se décomposer sous la forme :

$$Y = f_0 + f_1 + f_2 + f_3 + f_{12} + f_{13} + f_{23} + f_{123}$$

avec :

$$f_i = E(Y/X_i) - E(Y);$$

$$f_{ij} = E(Y/X_i, X_j) - E(Y/X_i) - E(Y/X_j) + E(Y);$$

$$f_{123} = E(Y/X_1, X_2, X_3) - E(Y/X_1, X_2) - E(Y/X_1, X_3) - E(Y/X_2, X_3) + E(Y/X_1) + E(Y/X_2) + E(Y/X_3) - E(Y)$$

Indices de sensibilité à 1,2 ou 3 variables

Tous les f_i sont variables aléatoires d'espérance nulle et \perp .

donc :

$\frac{V(f_i)}{V(Y)}$ sera donc la sensibilité au 1^{er} ordre de la variable X_i sur la variance de Y ;

$\frac{V(f_{ij})}{V(Y)}$ sera donc la sensibilité de l'interaction des variables X_i et X_j sur la variance de Y .

- 1 Analyse de sensibilité
 - Locale et globale
 - Indices de sensibilité : ex à 1,2 ou 3 variables
 - **Indices de sensibilité**
 - Estimation
- 2 Compléments
- 3 Incertitude de modèle et analyse de sensibilité
 - Prise en compte d'une mutation de modèle
 - Utilisation d'un modèle simplifié
- 4 Modèles à entrées non indépendantes
 - Le problème...
 - Indices de sensibilité : 2 variables dépendantes
 - Indices de sensibilité multidimensionnels
 - Applications au code de calcul Stay'SL

Indices de sensibilité pour modèle linéaire

Soit le modèle linéaire $Y = \sum_{i=1}^p \alpha_i X_i$ avec $X_i \perp X_j \quad \forall i \neq j$.

La variance de Y s'écrit $V(Y) = \sum_{i=1}^p \alpha_i^2 V(X_i)$.

La **part de variance** due à X_i est $\alpha_i^2 V(X_i)$.

⇒ on obtient une décomposition de la variance de Y en fonction des X_j .

La sensibilité de Y à X_i est quantifiée par l'indice Standardized Regression Coefficient :

$$\alpha_i^2 \frac{V(X_i)}{V(Y)} = SRC_i$$

Si le modèle n'est pas linéaire

Lorsqu'il n'est possible de faire **aucune hypothèse** sur le modèle (cas général), on définit des indices de sensibilité à partir d'une décomposition de la variance de Y .

Définition des indices de sensibilité 1/3

Soit le modèle $Y = f(X_1, \dots, X_p)$ avec $X_i \in \mathbb{R}$ et $X_i \perp X_j$.
Si f intégrable, elle peut se décomposer (Sobol[1993]) en

$$f(x_1, \dots, x_p) = f_0 + \sum_{i=1}^p f_i(x_i) + \sum_{1 \leq i < j \leq p} f_{ij}(x_i, x_j) + \dots + f_{1\dots p}(x_1, \dots, x_p) \quad (1)$$

avec

$$\begin{aligned} E_{i_k}[f_{i_1 \dots i_s}] &= 0 & \forall & 1 \leq i_1 \leq \dots \leq i_k \leq \dots \leq i_s \leq p \\ E[f_{i_1 \dots i_s} f_{i_1 \dots i_t}] &= 0 & \text{si } & (i_1 \dots i_s) \neq (i_1 \dots i_t) \end{aligned} \quad (2)$$

En écrivant (1) pour des variables aléatoires, et en utilisant (2),
on obtient :

$$\begin{aligned} f_0 &= E[Y] \\ f_i(X_i) &= E[Y|X_i] - f_0 \\ f_{ij}(X_i, X_j) &= E[Y|X_i, X_j] - f_i - f_j - f_0 \\ &\dots \end{aligned}$$

Définition des indices de sensibilité 2/3

ainsi, en utilisant cette décomposition, la variance de Y se décompose en :

$$V(Y) = D = \sum_{i=1}^p D_i + \sum_{1 \leq i < j \leq p} D_{ij} + \dots + D_{1\dots p}$$

où

$$D_i = V(f_i) = V(E[Y|X_i])$$

$$D_{ij} = V(f_{ij}) = V(E[Y|X_i, X_j] - E[Y|X_i] - E[Y|X_j])$$

$$\begin{aligned} D_{ijk} &= V(f_{ijk}) \\ &= V(E[Y|X_i, X_j, X_k] - E[Y|X_i, X_j] - E[Y|X_i, X_k] \\ &\quad - E[Y|X_j, X_k] - E[Y|X_i] - E[Y|X_j] - E[Y|X_k]) \end{aligned}$$

...

Définition des indices de sensibilité 3/3

on définit alors les indices de sensibilité :

- de premier ordre (sensibilité de Y à X_i)

$$S_i = \frac{D_i}{D} = \frac{V(E[Y|X_i])}{V(Y)}$$

- de deuxième ordre (sensibilité de Y à l'interaction entre X_i et X_j)

$$\begin{aligned} S_{ij} &= \frac{D_{ij}}{D} = \frac{V(E[Y|X_i, X_j]) - E[Y|X_i] - E[Y|X_j]}{V(Y)} \\ &\stackrel{\perp}{=} \frac{V(E[Y|X_i, X_j]) - D_i - D_j}{V(Y)} \end{aligned}$$

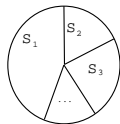
- etc ...

Interprétation

Pour un modèle à p variables d'entrées indépendantes :

$$\sum_{i=1}^p S_i + \sum_{1 \leq i < j \leq p} S_{ij} + \dots + S_{1\dots p} = 1$$

ce qui permet de les interpréter facilement,



plus l'indice est grand, plus la variable ou le groupe de variables est important vis à vis de la variance de Y

Indices de sensibilité totaux

Le nombre d'indices de sensibilité, $2^p - 1$, devient vite grand.
On introduit alors l'indice de sensibilité total

S_{T_i} = somme de tous les indices relatifs à X_i

qui exprime la sensibilité de Y à X_i sous toutes ses formes, i.e. à X_i seule et en interaction avec d'autres variables.

Ex : $p = 3$ $Y = f(X_1, X_2, X_3)$

$$S_{T_1} = S_1 + S_{12} + S_{13} + S_{123}$$

- 1 Analyse de sensibilité
 - Locale et globale
 - Indices de sensibilité : ex à 1,2 ou 3 variables
 - Indices de sensibilité
 - Estimation
- 2 Compléments
- 3 Incertitude de modèle et analyse de sensibilité
 - Prise en compte d'une mutation de modèle
 - Utilisation d'un modèle simplifié
- 4 Modèles à entrées non indépendantes
 - Le problème...
 - Indices de sensibilité : 2 variables dépendantes
 - Indices de sensibilité multidimensionnels
 - Applications au code de calcul Stay'SL

Méthodes d'estimation

- McKay
- FAST (*Fourier Amplitude Sensitivity Test*)
- Sobol
 - basée sur des estimations de Monte Carlo,
 - des améliorations existent :
 - échantillonnage stratifié,
 - LHS sampling,
 - Quasi Monte Carlo (séquences de Sobol $LP\tau$).

Estimation de Sobol

Indice de premier ordre : $S_i = \frac{D_i}{D} = \frac{V(E[Y|X_i])}{V(Y)}$

$D = V(Y)$ est estimée classiquement par

$$D \simeq \frac{1}{N} \sum_{k=1}^N f(x_{k1}^1, \dots, x_{kp}^1)^2 - f_0^2$$

où $(x_{k1}^1, \dots, x_{kp}^1)_{k=1 \dots N}$ est un premier échantillon de simulations des variables d'entrée,

et où f_0 est l'estimateur classique de l'espérance de Y .

Estimation de Sobol

Indice de premier ordre : $S_i = \frac{D_i}{D} = \frac{V(E[Y|X_i])}{V(Y)}$

$D = V(Y)$ est estimée classiquement par

$$D \simeq \frac{1}{N} \sum_{k=1}^N f(x_{k1}^1, \dots, x_{ki-1}^1, x_{ki}^1, x_{ki+1}^1, \dots, x_{kp}^1) f(x_{k1}^1, \dots, x_{ki-1}^1, x_{ki}^1, x_{ki+1}^1, \dots, x_{kp}^1) - f_0^2$$

et D_i est estimée par

$$D_i \simeq \frac{1}{N} \sum_{k=1}^N f(x_{k1}^1, \dots, x_{ki-1}^1, x_{ki}^1, x_{ki+1}^1, \dots, x_{kp}^1) f(x_{k1}^2, \dots, x_{ki-1}^2, x_{ki}^1, x_{ki+1}^2, \dots, x_{kp}^2) - f_0^2$$

où $(x_{k1}^2, \dots, x_{kp}^2)_{k=1 \dots N}$ est un second échantillon de simulations des variables d'entrée.

- 1 Analyse de sensibilité
 - Locale et globale
 - Indices de sensibilité : ex à 1,2 ou 3 variables
 - Indices de sensibilité
 - Estimation
- 2 Compléments
- 3 Incertitude de modèle et analyse de sensibilité
 - Prise en compte d'une mutation de modèle
 - Utilisation d'un modèle simplifié
- 4 Modèles à entrées non indépendantes
 - Le problème...
 - Indices de sensibilité : 2 variables dépendantes
 - Indices de sensibilité multidimensionnels
 - Applications au code de calcul Stay'SL

Incertitude de modèle

- Que deviennent les indices de sensibilité estimés si le **modèle change** (mutation) ?
- Comment prendre en compte, dans les résultats de sensibilité, l'utilisation d'un **modèle simplifié** ?

Modèles à entrées corrélées

- Comment réaliser (interpréter) une analyse de sensibilité lorsque les variables d'**entrée** ne sont **pas indépendantes** ?

- 1 Analyse de sensibilité
 - Locale et globale
 - Indices de sensibilité : ex à 1,2 ou 3 variables
 - Indices de sensibilité
 - Estimation
- 2 Compléments
- 3 **Incertitude de modèle et analyse de sensibilité**
 - **Prise en compte d'une mutation de modèle**
 - Utilisation d'un modèle simplifié
- 4 Modèles à entrées non indépendantes
 - Le problème...
 - Indices de sensibilité : 2 variables dépendantes
 - Indices de sensibilité multidimensionnels
 - Applications au code de calcul Stay'SL

Mutation de modèle

Que deviennent les indices de sensibilité si le modèle change ?

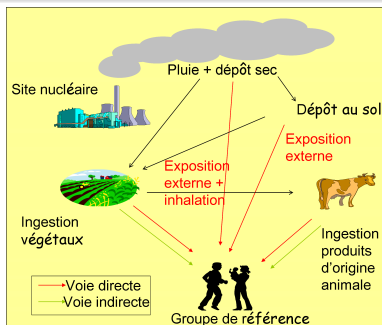
- une analyse de sensibilité a été réalisée sur un modèle M_1 ,
- ce modèle doit être modifié (nouvelle information, changement dans le processus modélisé...) en M_2 ,
- peut-on déduire les indices de sensibilité de M_2 à partir de ceux de M_1 (de façon exacte) ?
(et ainsi économiser le coût d'une nouvelle analyse)

Méthodologie :

- Un listing exhaustif des mutations n'est pas envisageable
- pour chaque mutation, existe-t-il un lien entre indices de sensibilité **avant** et **après** mutations ?

Le code de calcul GASCON

GASCON quantifie les transferts dans l'environnement et son impact sur l'homme suite à un rejet atmosphérique continu de radionucléides.



Temps de calcul trop long \Rightarrow utilisation de surfaces de réponse*

$$Y - \alpha_0 = \alpha_1 X_1 X_2 X_3 X_4 X_5 + \alpha_2 X_1 X_2 X_3 X_4 X_5^2 + \alpha_3 X_1 X_2 X_3 X_6 X_5 + \alpha_4 X_1 X_2 X_3 X_6 X_5^2 + \alpha_5 X_1 X_7 X_8 X_9 X_5 + \alpha_6 X_1 X_7 X_8 X_9 X_5^2 + \alpha_7 X_1 X_7 X_8 X_{10} X_5$$

* modèle de régression multiple ajusté sur une base de données de 1000 simulations de GASCON.

Exemple de mutation de modèle dans GASCON

On suppose les variables X_1 et X_5 fixées.

On suppose avoir réalisé les AS séparément sur

$$Y_1 = \alpha_1 x_1 x_2 x_3 x_4 x_5 + \alpha_2 x_1 x_2 x_3 x_4 x_5^2 \\ + \alpha_3 x_1 x_2 x_3 x_6 x_5 + \alpha_4 x_1 x_2 x_3 x_6 x_5^2$$

et $Y_2 = \alpha_5 x_1 x_7 x_8 x_9 x_5 + \alpha_6 x_1 x_7 x_8 x_9 x_5^2 + \alpha_7 x_1 x_7 x_8 x_{10} x_5$

Ces deux modèles correspondent à deux chaînes alimentaires particulières (lait de brebis et lait de chèvre).

Typologie de la mutation :

soient $Y_1 = f_1(X_1, \dots, X_p)$ et $Y_2 = f_2(X_{p+1}, \dots, X_{p+q})$.
on s'intéresse à $Y = Y_1 + Y_2$.

Mutation : somme de deux modèles

$$\text{hyp : } Y = Y_1 + Y_2 = f_1(X_1, \dots, X_p) + f_2(X_{p+1}, \dots, X_{p+q}).$$

Comme

$$\begin{aligned} E[Y|X_j] &= E[f_1(X_1, \dots, X_p) + f_2(X_{p+1}, \dots, X_{p+q})|X_j] \\ &= \begin{cases} E[Y_1|X_j] & \text{si } 1 \leq j \leq p \\ E[Y_2|X_j] & \text{si } p+1 \leq j \leq p+q \end{cases} \end{aligned}$$

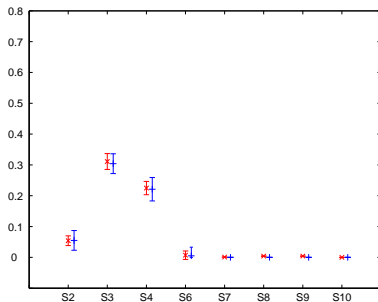
on a

$$S_j = \begin{cases} S_j^{\{1\}} \times \frac{V(Y_1)}{V(Y_1)+V(Y_2)} & \text{si } 1 \leq j \leq p \\ S_j^{\{2\}} \times \frac{V(Y_2)}{V(Y_1)+V(Y_2)} & \text{si } p+1 \leq j \leq p+q \end{cases}$$

où $S_j^{\{1\}}$ et $S_j^{\{2\}}$ sont les indices de Y_1 et Y_2 .

Rq : idem pour les indices d'ordre supérieur et totaux.

Résultats



Indices de sensibilité totaux du modèle somme.

rouge : mutation, bleu : nouvelle analyse

- 1 Analyse de sensibilité
 - Locale et globale
 - Indices de sensibilité : ex à 1,2 ou 3 variables
 - Indices de sensibilité
 - Estimation
- 2 Compléments
- 3 **Incertitude de modèle et analyse de sensibilité**
 - Prise en compte d'une mutation de modèle
 - **Utilisation d'un modèle simplifié**
- 4 Modèles à entrées non indépendantes
 - Le problème...
 - Indices de sensibilité : 2 variables dépendantes
 - Indices de sensibilité multidimensionnels
 - Applications au code de calcul Stay'SL

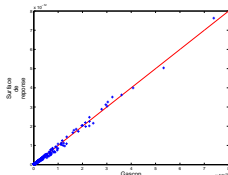
Prise en compte de l'utilisation d'un modèle simplifié

Comment prendre en compte, dans les résultats de sensibilité, l'utilisation d'un modèle simplifié ?

- le modèle de référence M_1 est trop lourd en temps de calcul pour être utilisé pour l'analyse de sensibilité,
- un modèle simplifié M_2 (régression multiple, splines, réseaux de neurones, ondelettes) est utilisé pour réaliser l'AS et ainsi approcher les indices de M_1 ,
- que dire de cette approximation ?

Méthodologie

- test de la validité de la régression (étude des résidus, test sur coefficient de détermination...)



modèle simplifié vs base de données Gascon (régression valide)

- si la régression n'est pas correcte
 - définition d'un schéma itératif d'amélioration de la qualité de l'ajustement du modèle simplifié,
 - estimation des indices du nouveau modèle simplifié à partir des travaux sur les mutations.

- 1 Analyse de sensibilité
 - Locale et globale
 - Indices de sensibilité : ex à 1,2 ou 3 variables
 - Indices de sensibilité
 - Estimation
- 2 Compléments
- 3 Incertitude de modèle et analyse de sensibilité
 - Prise en compte d'une mutation de modèle
 - Utilisation d'un modèle simplifié
- 4 **Modèles à entrées non indépendantes**
 - **Le problème...**
 - Indices de sensibilité : 2 variables dépendantes
 - Indices de sensibilité multidimensionnels
 - Applications au code de calcul Stay'SL

Le problème...

Soit le modèle $Y = f(X_1, \dots, X_i, \dots, X_j, \dots, X_p)$
où X_i et X_j sont non indépendantes.

Dans ce cas, les indices

$$S_i = \frac{V(E[Y|X_i])}{V(Y)} \qquad S_j = \frac{V(E[Y|X_j])}{V(Y)}$$

expriment une information redondante,
puisque $E[Y|X_i]$ est fonction de X_i mais aussi de X_j .

Ainsi, l'indice de sensibilité S_i n'exprime plus exclusivement la sensibilité à la variable X_i .

- 1 Analyse de sensibilité
 - Locale et globale
 - Indices de sensibilité : ex à 1,2 ou 3 variables
 - Indices de sensibilité
 - Estimation
- 2 Compléments
- 3 Incertitude de modèle et analyse de sensibilité
 - Prise en compte d'une mutation de modèle
 - Utilisation d'un modèle simplifié
- 4 **Modèles à entrées non indépendantes**
 - Le problème...
 - **Indices de sensibilité : 2 variables dépendantes**
 - Indices de sensibilité multidimensionnels
 - Applications au code de calcul Stay'SL

Indices de sensibilité : 2 variables dépendantes

- 2 variables dépendantes (X_1, X_2) :

$$Y = f(X_1, X_2) = E(Y/X_1, X_2).$$

et Y peut toujours se décomposer sous la forme :

$$Y = f_0 + f_1 + f_2 + f_{12}$$

f_0 est une constante ;

f_1 , f_2 et f_{12} sont toujours 3 variables aléatoires d'espérance nulle mais elles ne sont plus \perp .

Donc $V(Y) = V(f_1) + V(f_2) + V(f_{12}) + \text{toutes les covariances!}$
qui se ramènent à un seul terme :

$$V_{12}^{*(2)} = -2 * Cov(E(Y/X_1), E(Y/X_2))$$

Indices de sensibilité : 2 variables dépendantes

- 3 variables dont 2 dépendantes ($(X_1, X_2) \perp X_3$) :

$$Y = f(X_1, X_2, X_3) = E(Y/X_1, X_2, X_3)$$

Alors Y se décompose toujours sous la forme :

$$Y = f_0 + f_1 + f_2 + f_3 + f_{12} + f_{13} + f_{23} + f_{123}$$

Mais $V(Y) = V_1 + V_2 + V_3 + V_{12} + V_{13} + V_{23} + V_{123} + V_{12}^{*(3)}$
où $V_{12}^{*(3)}$ contient *toutes les covariances!*

Indices de sensibilité : 2 variables dépendantes

- p variables dont 2 dépendantes $((X_1, X_2) \perp X_3, \dots, X_p)$:
Dans la variance de Y se rajoute donc un terme $V_{12}^{*(p)}$
contenant *toutes les covariances!*
- $V_{12}^{*(p)} = 0$ si $X_1 \perp X_2$ mais ce n'est pas un nombre
nécessairement ≥ 0
- La corrélation entre (X_1, X_2) influe non seulement $V_{12}^{*(p)}$ mais
aussi la plupart des termes V_{\cdot} .
- $\frac{V_{12}^{*(p)}}{V(Y)}$ ne peut donc pas raisonnablement être interprété
comme un indice de la sensibilité du modèle à la corrélation
des variables d'entrée.

Une solution...

- considérer les 2 variables X_i et X_j comme une unique variable bidimensionnelle
- exprimer la sensibilité à cette variable (X_i, X_j) par

$$S_{\{i,j\}} = \frac{V(E[Y|X_i, X_j])}{V(Y)}$$

On généralise ainsi les indices de sensibilité classiques au cas **multidimensionnel** pour exprimer la sensibilité à des **groupes** de variables indépendants.

Rq : il est possible de définir des indices de sensibilité d'ordre supérieur et totaux sur ces variables multidimensionnelles.

Estimation des indices mutlidimensionnels

Indice de premier ordre :

$$S_{\{i,j\}} = \frac{V(E[Y|X_i, X_j])}{V(Y)} = \frac{D_{ij}}{D}$$

$D = V(Y)$ est estimée classiquement par

$$D \simeq \frac{1}{N} \sum_{k=1}^N f(x_{k1}^1, \dots, x_{kp}^1) f(x_{k1}^1, \dots, x_{ki-1}^1, x_{ki}^1, x_{ki+1}^1, \dots, x_{kj-1}^1, x_{kj}^1, x_{kj+1}^1, \dots, x_{kp}^1) - f_0^2$$

et D_{ij} est estimée par

$$D_{ij} \simeq \frac{1}{N} \sum_{k=1}^N f(x_{k1}^1, \dots, x_{kp}^1) f(x_{k1}^2, \dots, x_{ki-1}^2, x_{ki}^1, x_{ki+1}^2, \dots, x_{kj-1}^2, x_{kj}^1, x_{kj+1}^2, \dots, x_{kp}^2) - f_0^2$$

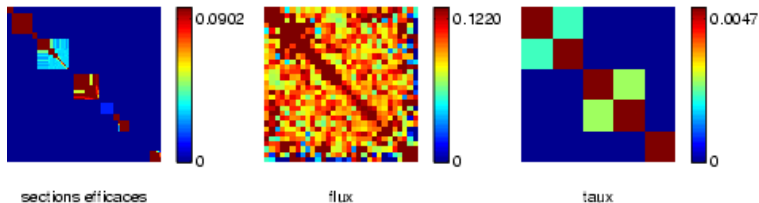
Plan de la partie Analyse de Sensibilité (AS)

- 1 Analyse de sensibilité
 - Locale et globale
 - Indices de sensibilité : ex à 1,2 ou 3 variables
 - Indices de sensibilité
 - Estimation
- 2 Compléments
- 3 Incertitude de modèle et analyse de sensibilité
 - Prise en compte d'une mutation de modèle
 - Utilisation d'un modèle simplifié
- 4 **Modèles à entrées non indépendantes**
 - Le problème...
 - Indices de sensibilité : 2 variables dépendantes
 - Indices de sensibilité multidimensionnels
 - Applications au code de calcul Stay'SL

Le code Stay'SL

Stay'SL est un code de calcul utilisé par le CEA pour l'estimation de la fluence neutronique reçue par un dispositif soumis à irradiation.

Entrées : 185 variables (taux de réactions, flux de neutrons, sections efficaces)



Sorties : 21 variables, dont

Y_{11} : somme des flux pondérés ayant une énergie $\geq 1\text{MeV}$.

Analyse de sensibilité multidimensionnelle de Stay'SL

On regroupe les 185 variables d'entrées en 9 groupes de variables indépendants :

- X_1 à X_5 : groupes de 30 sections efficaces (par groupe d'énergie), pour dosimètre Det1 à Det5,
- X_6 : groupe des 30 flux (par groupe d'énergie),
- X_7 : taux de réactions des dosimètres Det1 et Det2,
- X_8 : taux de réactions des dosimètres Det3 et Det4,
- X_9 : taux de réactions du dosimètre Det5.

Analyse de sensibilité multidimensionnelle de Stay'SL

Protocole d'estimation :

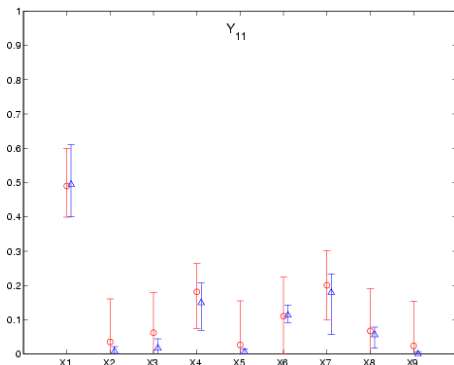
- taille des échantillons de MC : 100 000
(12h de calcul sur un PIV à 3GHz)
- on répète 7 fois l'analyse pour obtenir une idée de la variabilité des estimations

Résultats :

- indices de premier ordre et totaux

Analyse de sensibilité multidimensionnelle de Stay'SL

Indices de sensibilité **total** (o) et de **premier ordre** (Δ)



Analyse de sensibilité multidimensionnelle de Stay'SL

Conclusions de l'application :

- indices de premier ordre \simeq indices totaux
⇒ pas d'effet des interactions entre groupes de variables
- variables qui contribuent le plus à la variance de Y :
 - X_1 : sections efficaces (par groupe d'énergie) pour dosimètre Det1 ($\simeq 50\%$),
 - X_4 : sections efficaces (par groupe d'énergie) pour dosimètre Det4 ($\simeq 20\%$),
 - X_7 : taux de réactions des dosimètres Det1 et Det2.