Statistique en grande dimension
M2SSD

METAMODELISATION PAR PROCESSUS GAUSSIEN - TP

<u>EXERCICE 1 :</u> Simulation des trajectoires d'un processus gaussien de moyenne et covariance données

1.a) Soit $Z(t)_{t\in R}$ un processus gaussien centré stationnaire de covariance $C(t,t')=\sigma^2 R(t-t')$.

- <u>covariance gaussienne :</u> simuler des réalisations de *Z(t)* de N = 200 points sur *[0 ;1]* avec $\sigma^2 = 0.5$ et $R(h) = e^{-\frac{h}{\theta}}$. Prendre *θ=0.05* puis *θ=0.2*.
- <u>covariance exponentielle</u> : simuler des réalisations de *Z(t)* de N = 200 points sur *[0 ;1]* avec $\sigma^2 = 0.5$ et $R(h) = e^{-\frac{|h|}{\theta}}$. Prendre *θ=0.05* puis *θ=0.2*.
- <u>covariance gaussienne avec effet de pépite:</u> simuler des réalisations de *Z(t)* de N = 200 points sur *[0 ;1]* avec $\sigma^2 = 0.5$ et $R(h) = e^{-\left(\frac{h}{\theta}\right)^2} + \lambda\delta$. Prendre *λ=0.2* et *θ=0.05* puis *θ=0.2*.

1.b) Soit $Z(x)_{x\in R^2}$ un processus gaussien centré stationnaire de covariance $C(x,x') = \sigma^2 R(x-x')$.

- <u>covariance gaussienne isotrope :</u> simuler des réalisations de N = 50×50 points sur *[0 ;1]×[0 ;1]* avec $\sigma^2 = 0.5$ et $R(h) = e^{-\left\|\frac{h}{\theta}\right\|^2}$. Prendre *θ = 0.1*.
- <u>covariance gaussienne anisotrope :</u> simuler des réalisations de N = 50×50 points sur *[0 ;1]×[0 ;1]* avec $\sigma^2 = 0.5$ et $R(h) = e^{-\sum_{i=1}^{2}\left(\frac{h_i}{\theta_i}\right)^2}$. Prendre $\theta_1=0.1$ et $\theta_2=0.03$.
- <u>covariance exponentielle anisotrope :</u> simuler des réalisations de N = 50×50 points sur *[0 ;1]×[0 ;1]* avec $\sigma^2 = 0.5$ et $R(h) = e^{-\sum_{i=1}^{2}\frac{|h_i|}{\theta_i}}$. Prendre $\theta_1=0.1$ et $\theta_2=0.03$.

<u>EXERCICE 2 :</u> Construction d'un métamodèle PG à partir d'une base d'apprentissage en dimension 1

On considère la fonction analytique 1D sur [0 ;1] :

$$f(x) = \sin(30(x-0.9)^4)\cos(2(x-0.9)) + \frac{x-0.9}{2}$$

Appliquer la méthodologie suivante :

- **Etape 0 : Représentation de la fonction sur [0 ;1] => constitution de la base de test**
  - Evaluer f sur 100 points équirépartis et tracé de f

- **Etape 1 : Construction d'un plan d'expériences et évaluation de *f* sur ce plan => constitution de la base d'apprentissage**
  - Faire varier la taille du plan de la base de N = 10 à 30 points
  - Faire varier le type de plan : points équirépartis sur [0 ;1], tirage aléatoire uniforme

- **Etape 2 : Estimation des paramètres du métamodèle PG**

  Soit $Z(x)_{x\in[0;1]}$ un processus gaussien de moyenne constante et de covariance stationnaire de type Matern5/2.

  Construire le métamodèle basé sur Z conditionnellement aux points de la base d'apprentissage
  - Estimation des paramètres du métamodèle
  - Calcul et tracé du prédicteur sur la base de test
  - Calcul et tracé du MSE sur la base de test
  - Calcul du Q^2 sur la base de test

  -

- **Etape 3 (optionnelle) : Estimation des paramètres du métamodèle PG**

  Estimation des hyperparamètres par maximum de vraisemblance :

  tracer $\varphi(\theta) = \left|\Sigma_{s,\theta}\right|^{-\frac{1}{N}} \sigma^{2^*}(\theta)$ et retrouver la valeur optimale de θ.

- **Etape 4 (optionnelle) : planification adaptative**

  Trouver le point x où le MSE est maximal et l'ajouter à la base d'apprentissage. Mettre à jour le métamodèle (estimation des hyperparamètres, construction du prédicteur et du MSE). Tracé du prédicteur et du MSE comme à l'étape 2.

<u>EXERCICE 3 :</u> Construction d'un métamodèle PG à partir d'une base d'apprentissage en dimension 2

On considère la fonction analytique schwefel2D sur $[-200\ ;200]^2$:

$$f(x_1, x_2) = -x_1 \sin\left(\sqrt{|x_1|}\right) - x_2 \sin\left(\sqrt{|x_2|}\right)$$

Appliquer la méthodologie suivante :

- **Etape 0 : Représentation de la fonction sur $[-200\ ;200]^2$ => constitution de la base de test**
  - Evaluer f sur une grille de 70x70 points équirépartis et tracé de f

- **Etape 1 : Construction d'un plan d'expériences et évaluation de *f* sur ce plan => constitution de la base d'apprentissage**
  - Faire varier la taille du plan de la base de N = 70 à 100 points
  - Faire varier le type de plan : tirage aléatoire uniforme, plan hypercubes latins

- **Etape 2 : Estimation des paramètres du métamodèle PG**

  Soit $Z(x)_{x\in[0;1]}$ un processus gaussien centré stationnaire de covariance Matern3/2.

  Construire le métamodèle basé sur Z conditionnellement aux points de la base d'apprentissage.
  - Calcul et tracé du prédicteur sur la base de test
  - Tracé de l'erreur (valeur absolue) sur la base de test
  - Calcul du Q^2 sur la base de test

<u>EXERCICE 4 :</u> Utilisation d'un métamodèle PG pour l'analyse de sensibilité

On considère la fonction analytique Ishigami :

$$f(X_1, X_2, X_3) = \sin X_1 + 7(\sin X_2)^2 + 0.1 X_3^4 \sin X_1$$

Avec X_i uniforme sur [- π; π], pour i = 1,..,3.

- **Etape 0 : Calcul analytique des indices de sensibilité**
    - Calculer les indices de Sobol du 1^{er} et du 2^{nd} ordre théoriques.
    - Calculer les indices de Sobol totaux

- **Etape 1 : Construction d'un plan d'expériences et évaluation de *f* sur ce plan => constitution de la base d'apprentissage**
    - Construire un plan LHS de N = 70 à 100 points

- **Etape 2 : Construction du métamodèle PG et contrôle de sa qualité de prédiction**

  Construire le métamodèle basé sur Z conditionnellement aux points de la base d'apprentissage.
    - Calcul du Q^2 sur une base de test
    - Calcul du Q^2 par validation croisée sur la base d'apprentissage

- **Etape 3 : Calcul des indices de sensibilité**

  Estimer les indices de Sobol du 1^{er} ordre et totaux en utilisant le métamodèle

  Comparer les valeurs estimées avec le métamodèle et les valeurs théoriques