
Examen de rattrapage – majeure Data Science – UP4

3 mars 2016 – Durée 1h30 – Documents interdits

Exercice 1 (5 pts)

Après avoir résumé les principes de fonctionnement des principaux algorithmes d'optimisation globale vus en cours, vous expliquerez dans quel cas utiliser un algorithme plutôt qu'un autre en *justifiant* par des raisons techniques votre explication.

Exercice 2 (4 pts)

Let us consider the 2-dimensional function :

$$f(x_1, x_2) = x_1 + x_1^2 x_2^2$$

Assume that X_1, X_2 are independent and identically distributed random variables, with for $i = 1, 2$: $E(X_i) = 0$, $\text{var}(X_i) = 1$, and denote $v = \text{var}(X_i^2) < +\infty$.

1. Compute the Sobol-Hoeffding decomposition of $f(X_1, X_2)$.
2. Let D the global variance of $f(X_1, X_2)$. Express the Sobol indices as a function of D and v .
Indication : Start by $S_2, S_{1,2}$ and deduce S_1 .

Exercice 3 (3 pts)

You would like to perform a global sensitivity analysis of a time-consuming d -dimensional function f , and your computational budget is limited to $n = 10d$ function evaluations. What kind of method can you use directly on f ? What other strategy can you propose? Discuss briefly the advantages/drawbacks of the two methods.

Exercice 4 (4 pts)

Soient Y et Z deux processus Gaussiens centrés indépendants sur \mathbb{R} , de fonctions de covariance k_Y et k_Z . De plus, on introduit le processus W qui est défini par $W(x) = Y(x) + Z(x)$.

1. Donnez l'expression du noyau de W .
2. Donnez l'expression de la moyenne conditionnelle, de la variance conditionnelle et de la fonction de covariance conditionnelle de W sachant qu'il prend les valeurs $F = (f_1, \dots, f_n)$ aux points $X = (x_1, \dots, x_n)$.

On part maintenant du principe que le processus Y correspond au signal que l'on souhaite prédire et que Z correspond à du bruit d'observation (mais pas forcément un bruit blanc). Comme pour les questions précédentes, on dispose d'observations de W : $W(X) = F$.

3. Proposez un modèle sous forme d'un processus conditionnel. Donnez l'expression de la moyenne conditionnelle et de la variance de prédiction.

Exercice 5 (4 pts)

1. Détaillez les différentes étapes pour construire un plan d'expérience de type LHS, et pour l'optimiser par rapport au critère maximin.
2. Détaillez les différentes étapes permettant d'obtenir un plan d'expérience de type tessellations centroidales de Voronoï.