
TP cours 2 : Modélisation par krigeage

Ecole-chercheur Mexico, La Rochelle

N. Durrande - V. Picheny

Cette session est dédiée à la construction d'un modèle de krigeage pour approximer la fonction `compute_wls` qui correspond à l'erreur du simulateur numérique. Ce metamodelle sera utilisé par la suite pour l'étape de calibration du simulateur.

1 Construction d'un modèle 1D

Q1. Un jeu de données jouet vous est fourni dans le fichier `toy_data.Rdata`. Charger le à l'aide de la fonction `load` et représenter graphiquement Y contre X . Que pouvez-vous intuiter de la régularité de la fonction f ?

Q2. Quelques lignes de codes vous sont données pour générer des trajectoires non conditionnelles à partir du package `DiceKriging`. Modifier la valeur des paramètres `covtype`, `coef.trend`, `coef.var` et `coef.cov` afin de comprendre l'influence de ces paramètres (il n'est pas interdit de consulter la documentation !). Trouver un jeu de paramètre qui vous semble adapté au jeu de données.

Q3. Simuler et représenter des trajectoires conditionnelles (option `cond` de la fonction `simulate`). Tracer la moyenne et des intervalles de confiance à l'aide de la fonction `sectionview` du package `DiceView`.

Q4. Changer les paramètres d'appel de la fonction `km` afin d'estimer les paramètres de covariance par maximum de vraisemblance. Vous pouvez ensuite utiliser la commande `print(m)` pour afficher des détails sur le modèle. Les paramètres trouvés sont-ils proche de ceux que vous aviez intuités ?

Q5. La fonction `leaveOneOut.km` est bien pratique pour calculer les prédictions du modèle aux points du plan d'expérience en utilisant tous les points du plan sauf celui où l'on effectue la prédiction. Utiliser ce vecteur de prédiction pour calculer les résidus standardisés (erreur de prédiction divisée par l'écart type de prédiction). Les comparer avec une distribution normale centrée réduite. Le résultat vous paraît-il satisfaisant ?

Q6. Utiliser la fonction `predict` pour prédire la valeur de f au point $x = 5$. Pouvez-vous expliquer (et critiquer) le résultat obtenu ? Créer un nouveau modèle en choisissant la valeur par défaut de `coef.trend` et jouer sur la formule (premier argument) de la fonction `km` pour obtenir un modèle plus satisfaisant.

2 Cas test du volcan

Q7. Récupérer le meilleur plan d'expérience de 100 points en dimension 5 obtenu lors de la séance d'hier. Si vous n'êtes pas arrivé à un résultat convainquant vous pouvez utiliser la correction présente sur le dropbox.

Q8. Reprendre les questions Q4, Q5 et Q6 (en les adaptant si nécessaire) sur les données de calibration du volcan pour obtenir un modèle qui vous semble probant.