#

Objectifs : - pouvoir anticiper si le cokrigeage peut être utile;

 - vérifier la qualité et l’admissibilité d’un modèle linéaire de corégionalisation;

 - être capable d’interpréter les résultats d’un cokrigeage.

Important : répondez brièvement aux questions; soyez directs. Répondez dans le fichier word en mettant vos réponses en évidence.

 **Formez des équipes de deux ou trois personnes.**

|  |
| --- |
| Équipe : |
| Noms des étudiants :  |
|  |
|  |

Données : Walker Lake (USA) : Données topographiques du sud-ouest américain à partir desquelles 2 variables sont construites : Z et Y. *Copiez tous les fichiers situés sur k:\geostat\glq3401\2020\tp8 dans votre répertoire usager*. *Choisissez votre répertoire usager ou vous avez copié les fichiers comme répertoire de travail dans Matlab.*

A- Initialiser certaines variables et lire les données: **>> go**

B- Calculer les statistiques élémentaires : **>> statelem** (examinez les figures 2, 3 et 4. Sur la figure 3, la taille des symboles est proportionnelle à la valeur).

Q1. **(1pts)** D’après ce que vous observez, a-t-on raison de croire que le cokrigeage pourrait améliorer la précision de l’estimation par rapport au krigeage? Justifiez.

C- Calculer et modéliser les covariances et covariances croisées : **>> covariance** (voir figure 9). Note : On a supposé qu’un modèle isotrope (et donc symétrique en « h ») pouvait s’appliquer. Les paramètres du modèle (courbes rouges) apparaissent dans les titres.

*Q2.* **(2pts)** *Spécifiez le modèle linéaire de corégionalisation représenté sur la figure (type, paramètres). Le modèle est-il admissible? Faites le calcul pour justifier votre réponse.*

*Q3.* **(1pts)** *Quelle est la corrélation théorique décrite par le modèle entre Z et Y à la distance 0 ? Comparez à la corrélation expérimentale obtenue aux 50 points où Z et Y ont été observés simultanément.*

Q4. **(1pts)** Observez le nombre de paires utilisées pour les différentes fonctions de covariances. Combien de paires aurait-on eu sur le variogramme croisé pour les deux premiers points (excluant le 0) ?

D- Faites la validation croisée du modèle (krigeage de Z et cokrigeage de Z par Z et Y) : **>>validation** (l’exécution prend quelques secondes; les statistiques d’erreur sont affichées dans la fenêtre Matlab; notez ces valeurs pour la question 6).

Note : 2 validations croisées différentes sont effectuées pour le cokrigeage. Dans un premier temps on n’enlève que la variable principale au point à valider (cok1). Ceci simule le cas où l’on connaîtrait la variable Y en chaque point du domaine (ex. une variable secondaire tirée d’une image satellite). Dans un 2e temps, on enlève les 2 variables au point à valider (cok2), ce qui correspond davantage au jeu de données actuel.

*Toutes ces statistiques sont calculées seulement avec les 50 données Z connues. Complétez le tableau suivant :*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | kri | cok1 | cok2 |  | kri | cok1 | cok2 |
| erreurs^2 |  |  |  | erreurs\_normalisées^2 |  |  |  |

Q5. **(1pts)** Discutez les résultats obtenus (krigeage vs cok1 vs cok2)? Les résultats semblent-ils justifier l’utilisation du cokrigeage? Le modèle ajusté semble-t-il entièrement adéquat? Justifiez.

E- Faites le krigeage et le cokrigeage de la variable Z sur la grille 50 x 50 : **>>kriger**

F- Visualisez les résultats du krigeage : **>>visualisation**

Ici la réalité est connue pour Z et Y (voir figures 17 et 18). La surface krigée est représentée (figure 19) ainsi que la surface cokrigée (figure 20). Les statistiques d’erreur entre les surfaces estimées et les surfaces réelles sont aussi données dans les titres des figures.

Q6 **(1pts)** Comparez les statistiques d’erreur obtenues sur l’ensemble de la surface par krigeage et par cokrigeage avec les statistiques obtenues par validation croisée. Commentez (la validation croisée permet-elle de prévoir assez bien la performance réelle des 2 méthodes d’estimation) ?

G- Examinez la figure suivante qui montre les histogrammes des variances de krigeage et de cokrigeage aux 150 points où seul Y est connu ainsi que sur l’ensemble des points de la surface.

Q7. **(1pts)** Se basant sur cette figure, si l’on pouvait mesurer « Y » en tout point et à faible coût, devrait-on le faire? Justifiez.

|  |  |  |
| --- | --- | --- |
|  | Histogrammes des variances de krigeage/cokrigeage aux 150 points où seul Y est connu | Histogrammes des variances de krigeage/cokrigeage sur l’ensemble des points |
| krigeagecokrigeage |  |  |

Cette fois, on suppose que la variable Y est connue en tout point de la grille. On va vérifier si cette connaissance accrue de Y permet d’améliorer l’estimation de Z. Notez **les statistiques des figures 19 et 20. La figure 20 de la question 6 est recopiée sur la figure 120 pour vous aider à répondre.**

Q8. **(2pts)** Lancez cette analyse en faisant >>**go2**. Le programme repasse toutes les étapes précédentes, mais cette fois avec 2500 valeurs Y au lieu de 200 comme précédemment. Comparez les statistiques de validation croisée de la Q.5 (cok1) et celles indiquées sur la nouvelle figure 20. La validation croisée effectuée à la Q.5 permet-elle de prédire approximativement le gain du cokrigeage lorsque Y est connu partout? L’ajout des informations sur « Y » a-t-il été bénéfique comme le laissait présager les histogrammes de la question 7 ?

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | kri | cok1 | cok2 |  | kri | cok1 | cok2 |
| erreurs^2 |  |  |  | erreurs\_normalisées^2 |  |  |  |