Introduction aux statistiques spatiales.
Le krigeage

Denis Allard

Biostatistique et Processus Spatiaux (BioSP), INRA, Avignon

19 décembre 2012

Définition du krigeage

Echantillon $z(x_1), z(x_2), \ldots, z(x_n)$ de $\{Z(x), x \in D \subset \mathbb{R}^d\}$. Krigeage = Best

Linear Unbiased Predictor de $g(Z(\cdot))$, $x \in D$, où g est une fonctionnelle linéaire de $Z(\cdot)$

- Linéaire
- Autorisé (important si stationnarité intrinsèque)
- Sans biais
- Variance de prédiction minimale

Exemples :

- $Z(x_0)$, x_0 non observé : krigeage simple ou ordinaire
- $1/|B| \int_B Z(s)ds, \quad B \subset D$: krigeage de bloc
- avec covariables déterministes
- vecteur multivarié

Notations : $Z_\alpha = Z(x_\alpha); \quad C_{\alpha\beta} = \mathrm{Cov}(Z_\alpha, Z_\beta)\,; \alpha, \beta = 1, \ldots, n.$

(Le nom a été donné par G. Matheron en l'honneur de D. Krige, ingénieur des mines sud-africain.)

Dérivées vectorielles

On devra manipuler

$$\sum_{i=1}^{n} \lambda_i Z_i = \mathbf{\Lambda}^t \mathbf{Z} \ \text{ et } \ \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \lambda_j C_{ij} = \mathbf{\Lambda}^t \mathbf{C} \mathbf{\Lambda}$$

et leurs dérivées par rapport aux λ_i.

Queques résultats utiles

1. $$\frac{\partial}{\partial \lambda_k} \sum_{i=1}^{n} \lambda_i Z_i = Z_k \Rightarrow \frac{\partial \mathbf{\Lambda}^t \mathbf{Z}}{\partial \mathbf{\Lambda}} = \mathbf{Z}$$

   [la dérivée d'un scalaire par un vecteur est un vecteur]

2. $$\frac{\partial}{\partial \lambda_k} \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \lambda_j C_{ij} = 2 \sum_{i=1}^{n} \lambda_i C_{ik} \Rightarrow \frac{\partial \mathbf{\Lambda}^t \mathbf{C} \mathbf{\Lambda}}{\partial \mathbf{\Lambda}} = \mathbf{\Lambda}^t \mathbf{C} + \mathbf{C} \mathbf{\Lambda}$$

3. Mais, attention

   $$\frac{\partial^2 \mathbf{\Lambda}^t \mathbf{Z} \mathbf{\Lambda}}{\partial \mathbf{\Lambda} \partial \mathbf{\Lambda}^t} = 2\mathbf{C}$$

   [la dérivée second d'un scalaire par un vecteur est une matrice]

Krigeage simple

$Z(\cdot)$ est sta–deux ; $\mu = E[Z(\cdot)]$ est connue. Krigeage de Z_0.

- Prédicteur linéaire : $Z_0^* = \mu + \sum_{\alpha=1}^{n} \lambda_\alpha (Z(x_\alpha) - \mu)$
- Autorisé car sta–2
- Sans biais, car $E[Z_0^*] = \mu + \sum_{\alpha=1}^{n} \lambda_\alpha E[(Z(x_\alpha) - \mu)] = \mu = E[Z_0]$
- Minimiser

$$\mathrm{Var}(Z_0^* - Z_0) = \sum_\alpha \sum_\beta \lambda_\alpha \lambda_\beta C_{\alpha\beta} - 2 \sum_\alpha \lambda_\alpha C_{0\alpha} + C_{00} = \mathbf{\Lambda}^t \mathbf{C} \mathbf{\Lambda} - 2\mathbf{\Lambda}^t C_0 + C_{00}$$

Le vecteur $\mathbf{\Lambda} = (\lambda_1, \ldots, \lambda_n)$ qui minimise la variance de prédiction est

$$\mathbf{\Lambda} = \mathbf{C}^{-1} C_0$$

où $\mathbf{C} = [C_{\alpha\beta}]_{\alpha,\beta=1,\ldots,n}$ et $C_0 = (C_{0\alpha})_{\alpha=1,\ldots,n}$.

La variance de krigeage $=$ la variance de prédiction minimisée :

$$\sigma_{KS}^2(s_0) = \sigma^2 - C_0^t \mathbf{C}^{-1} C_0$$

Krigeage simple : propriétés

- Dans le cas où $Z(\cdot)$ est une fonction aléatoire gaussienne, le KS est l'espérance conditionnelle de $Z_0 \mid Z_1, \ldots, Z_n$, cf. cas Gaussien général.
- Krigeage simple est dans tous les cas l'interpolateur linéaire "optimal" *à fonction de covariance connue*
- En un point de donnée x_i, on peut vérifier $Z_i^* = Z_i$. Le krigeage est un interpolateur exact : KS passe par les données
- En interpolant en tous noeuds d'une grille, on peut construire une carte de krigeage et une carte de variance
- Les poids $\mathbf{\Lambda}$ dépendent explicitement
  - de la fonction de covariance (qui dépend indirectement des données, à travers le variogramme)
  - de la géométrie de l'échantillon et de la position de x_0
  - ne dépendent pas de la variance $\sigma^2 = C(0)$
  - ne dépendent pas explicitement des données
- Le prédicteur

$$Z_0^* = \mu + \mathbf{\Lambda}^t \mathbf{Z} = C_0^t \mathbf{C}^{-1}(\mathbf{Z} - \mu \mathbf{1})$$

dépend des valeurs Z_i

Krigeage simple (suite)

- Le KS est également l'unique solution de l'équation

$$\mathrm{Cov}(Z_0^*, Z_0^* - Z_0) = 0,$$

  où $Z_0^* =$ combinaison linéaire des données. En effet, ...
- Donc,

$$\mathrm{Var}(Z_0) = \mathrm{Var}(Z_0 - Z_0^* + Z_0^*) = ... = \sigma^2_{KS(s_0)} + \mathrm{Var}(Z_0^*),$$

  et donc,

$$\mathrm{Var}(Z_0^*) \leq \mathrm{Var}(Z_0).$$

Si $C(h) = \sigma^2 \delta_0(h)$, alors

- $\mathbf{C} = \sigma^2 \mathbf{I}$
- $C_0 = 0$
- $\boldsymbol{\Lambda} = \mathbf{0}$

$\Rightarrow Z_0^* = \mu$ et $\mathrm{Var}(Z_0^*) = \sigma^2$.

Pépite $\Leftrightarrow$ absence de spatialisation !

Krigeage de la moyenne

$Z(\cdot)$ est sta–deux ; $\mu = E[Z(\cdot)]$ est inconnue. Krigeage de μ.

- Prédicteur linéaire : $\mu^* = \sum_{\alpha=1}^{n} \lambda_\alpha Z(x_\alpha) = \mathbf{\Lambda}^t\mathbf{Z}$
- Autorisé car sta–2
- Sans biais si $\sum_{\alpha=1}^{n} \lambda_\alpha = 1$ car $E[\mu^*] = \sum_{\alpha=1}^{n} \lambda_\alpha E[Z(x_\alpha)] = 1.\mu$
- Minimiser
$$\text{Var}(\mu^*) = \sum_\alpha \sum_\beta \lambda_\alpha \lambda_\beta C_{\alpha\beta} = \mathbf{\Lambda}^t\mathbf{C}\mathbf{\Lambda}$$
sous la condition $\sum_{\alpha=1}^{n} \lambda_\alpha = \mathbf{\Lambda}^t\mathbf{1} = 1.$
- Utilisation des pondérateurs de Lagrange. On va minimiser
$$Q = \mathbf{\Lambda}^t\mathbf{C}\mathbf{\Lambda} - \nu(\mathbf{\Lambda}^t\mathbf{1} - 1)$$
- La solution est
$$\mathbf{\Lambda} = \frac{\mathbf{C}^{-1}\mathbf{1}}{\mathbf{1}^t\mathbf{C}^{-1}\mathbf{1}}$$
- Alors,
$$\mu^* = \frac{\mathbf{1}^t\mathbf{C}^{-1}\mathbf{Z}}{\mathbf{1}^t\mathbf{C}^{-1}\mathbf{1}} \quad \text{et} \quad \sigma^2_{KM} = \frac{1}{\mathbf{1}^t\mathbf{C}^{-1}\mathbf{1}}$$

Krigeage de la moyenne (suite)

$$\mu^* = \frac{\mathbf{1}^t\mathbf{C}^{-1}\mathbf{1}}{\mathbf{1}^t\mathbf{C}^{-1}\mathbf{1}} \quad \text{et} \quad \sigma^2_{KM} = \frac{1}{\mathbf{1}^t\mathbf{C}^{-1}\mathbf{1}}$$

Si $C(h) = \sigma^2\delta_0(h)$, $\Rightarrow \mathbf{C} = \sigma^2\mathbf{I}$ et $\mathbf{1}^t\mathbf{C}^{-1}\mathbf{1} = n/\sigma^2$. D'où

- $\mathbf{\Lambda} = (1/n, \dots, 1/n)$
- $\mu^* = \bar{Z}$
- $\sigma^2_{KM} = C(0) = \sigma^2/n$

Si $C(h) \neq \delta_0(h)$, $\Rightarrow \mathbf{C} = \sigma^2\mathbf{R}$

- $\mathbf{\Lambda} \neq (1/n, \dots, 1/n)$
- $NEDI := \mathbf{1}^t\mathbf{R}^{-1}\mathbf{1} \leq n$
- $\sigma^2_{KM} = C(0)/NEDI$

Krigeage ordinaire

$Z(\cdot)$ est sta–deux ; $\mu = E[Z(\cdot)]$ est inconnue. Krigeage de Z_0.

- Prédicteur linéaire : $Z_0^* = \sum_{\alpha=1}^n \lambda_\alpha Z(x_\alpha)$
- Autorisé car sta–2
- Sans biais si $\sum_{\alpha=1}^n \lambda_\alpha = 1$
  car $E[Z_0^*] = \sum_{\alpha=1}^n \lambda_\alpha E[Z(x_\alpha)] = 1.\mu = E[Z_0]$
- Minimiser

$$\mathrm{Var}(Z_0^* - Z_0) = \sum_\alpha \sum_\beta \lambda_\alpha \lambda_\beta C_{\alpha\beta} - 2 \sum_\alpha \lambda_\alpha C_{0\alpha} + C_{00} = \mathbf{\Lambda}^t \mathbf{C} \mathbf{\Lambda} - 2\mathbf{\Lambda}^t C_0 + C_{00}$$

sous la condition $\sum_{\alpha=1}^n \lambda_\alpha = \mathbf{\Lambda}^t \mathbf{1} = 1$.

Krigeage ordinaire (suite)

- Pondérateurs de Lagrange. On va minimiser

$$Q = \mathbf{\Lambda}^t \mathbf{C} \mathbf{\Lambda} - 2\mathbf{\Lambda}^t C_0 + C_{00} + \nu(\mathbf{\Lambda}^t \mathbf{1} - 1)$$

- La solution est

$$\mathbf{\Lambda} = \mathbf{C}^{-1} C_0 + \frac{\mathbf{C}^{-1}\mathbf{1}}{\mathbf{1}^t \mathbf{C}^{-1} \mathbf{1}} (1 - \mathbf{1}^t \mathbf{C}^{-1} C_0)$$

- Alors,

$$Z_0^* = C_0^t \mathbf{C}^{-1} \mathbf{Z} + \frac{\mathbf{1}^t \mathbf{C}^{-1} \mathbf{Z}}{\mathbf{1}^t \mathbf{C}^{-1} \mathbf{1}} (1 - \mathbf{1}^t \mathbf{C}^{-1} C_0)$$

  et

$$\mathrm{Var}(Z_0^* - Z_0) = C(0) - C_0^t \mathbf{C}^{-1} C_0 + \frac{(1 - \mathbf{1}^t \mathbf{C}^{-1} C_0)^2}{\mathbf{1}^t \mathbf{C}^{-1} \mathbf{1}}$$

Krigeage ordinaire : propriétés

En gros, mêmes propriétés que le krigeage simple, pour les combinaisons linéaires $\sum_{i=1}^n \lambda_i = 1$.

- interpolateur linéaire optimal
- interpolateur exact
- unique solution de

$$\mathrm{Cov}(Z_0 - \sum_{i=1}^n \lambda_i Z_i, \sum_j \nu_j Z_j) = 0, \ \ \forall \boldsymbol{\nu}, \text{ t.q. } \sum_{j=1}^n \nu_j = 0$$

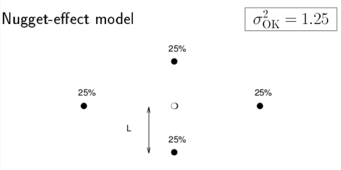
Si $C(h) = \sigma^2 \delta_0(h)$, alors

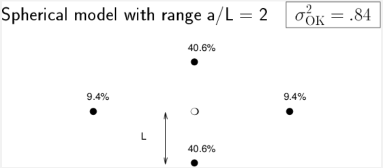
- $\mathbf{C} = \mathbf{I}$, $\mathbf{1}^t\mathbf{C}^{-1}\mathbf{1} = n$, et $C_0 = 0$
- $\boldsymbol{\Lambda} = 1/n$

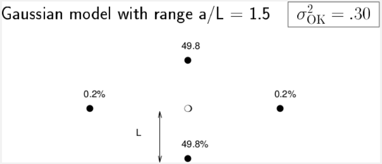
$\Rightarrow Z_0^* = \bar{Z}$ et $\mathrm{Var}(Z_0^*) = C(0) = \sigma^2$.

Pépite $\Leftrightarrow$ absence de spatialisation !

Illustrations du krigeage (1)

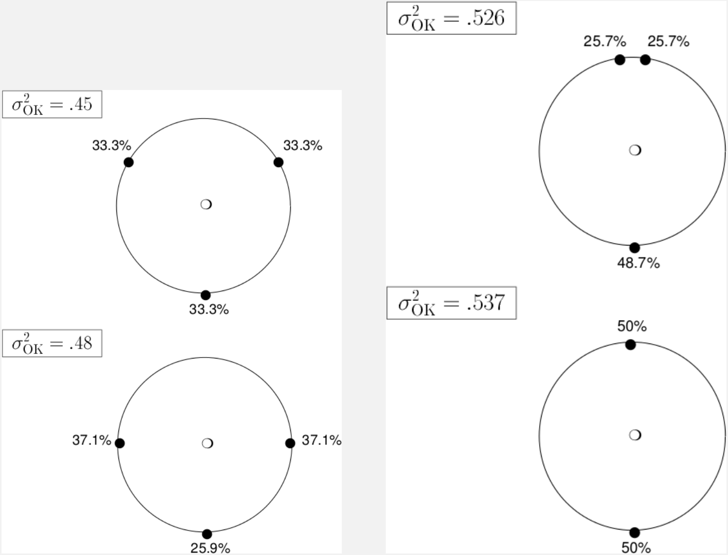




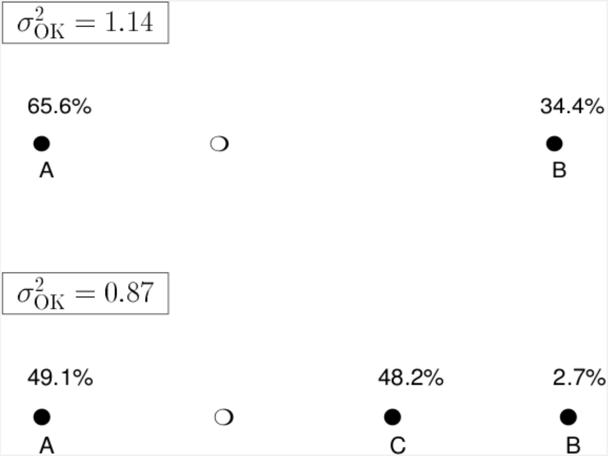


En haut : pépite pure. Au milieu : modèle sphérique($a = 2L$). En bas : modèle Gaussien ($a = 1.5L$

Illustrations du krigeage (2)



Illustrations du krigeage (3)

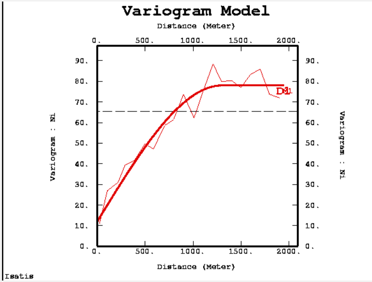


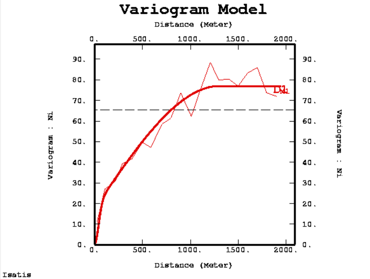
Modèle sphérique($a = 2L$)

Pratique du krigeage

1. Si l'espérance est connue, centrer la variable
2. Calculer un variogramme expérimental.
3. En déduire des hypothèses et une famille paramétrique pour le variogramme
4. Faire l'estimation des paramètres du variogramme
5. Si $n \leq 5000$ (environ)
   5.1 Construire $\mathbf{C}$ et l'inverser : $\mathbf{D} = \mathbf{C}^{-1}$ (fonction `solve` dans `R`)
   5.2 En chaque point x_0 d'un grille :
      5.2.1 Calculer C_0 ; en déduire $\mathbf{\Lambda}$ et σ_K^2 selon les équations du krigeage (produits scalaires entre $\mathbf{D}$ et C_0)
      5.2.2 Calculer $Z_0^* = \sum_{i=1}^n \lambda_i Z_i$
6. Valider (par validation croisée, p.ex.) le choix du modèle. Si critères de validation mauvais, retourner en 3.

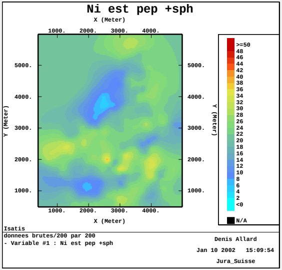
Illustrations : variogrammes

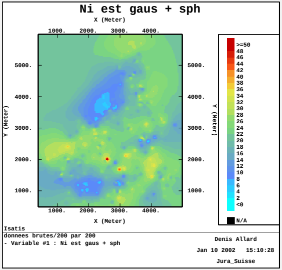




A gauche : modèle pépite + Sph. A droite : modèle Gaus. + Sph.

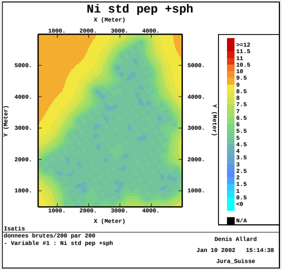
Illustrations : krigeage

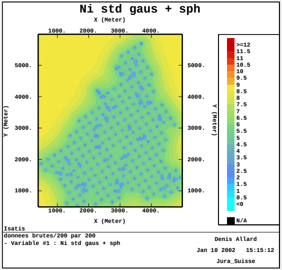




A gauche : modèle pépite + Sph. A droite : modèle Gaus. + Sph.

Illustrations : variance de krigeage





A gauche : modèle pépite + Sph. A droite : modèle Gaus. + Sph.

Krigeage avec variogramme non borné (KO)

- ex : variogramme puissance $\gamma(h) = a||h||^{\alpha}, \quad 0 < \alpha < 2$
- pas d'espérance mathématique $\Rightarrow$ pas de krigeage simple, ni de krigeage de la moyenne
- Il faut manier des combinaisons linéaires autorisées :

$$Z_0 - Z_0^* = Z_0 - \sum_{i=1}^{n} \lambda_i Z_i \text{ autorisée } \Rightarrow \sum_{i=1}^{n} \lambda_i = 1$$

  même condition que pour le non biais !
- Minimiser

$$\mathrm{Var}(Z_0^* - Z_0) = -\sum_{\alpha}\sum_{\beta} \lambda_\alpha \lambda_\beta \gamma_{\alpha\beta} + 2\sum_{\alpha} \lambda_\alpha \gamma_{0\alpha} + \gamma_{00} = -\mathbf{\Lambda}^t \mathbf{\Gamma} \mathbf{\Lambda} + 2\mathbf{\Lambda}^t \gamma_0$$

  sous la condition $\mathbf{\Lambda}^t \mathbf{1} = 1$.

$\Rightarrow$ Même équations que krigeage ordinaire, en remplaçant C_{ij} par γ_{ij}

Cas général : Krigeage avec régression

Modèle :

$$Z(\cdot) = \mathbf{a}^t \mathbf{X}(\cdot) + Y(\cdot)$$

où

- $\mathbf{a}$ est un K-vecteur de covariables connues en tous points
- $Y(\cdot)$ est sta–2, et $E[Y(\cdot)] = 0$

Krigeage de Z_0 :

- Prédicteur linéaire : $Z_0^* = \sum_{\alpha=1}^n \lambda_\alpha Z(x_\alpha)$
- Sans biais :

$$\begin{aligned} E[Z_0^* - Z_0] &= \sum_{i=1}^n \lambda_i \mathbf{a}^t \mathbf{X}(x_i) + \sum_{i=1}^n \lambda_i E[Y_i] - \mathbf{a}^t \mathbf{X}(x_0) \\ &= \mathbf{a}^t \left\{ \sum_{i=1}^n \lambda_i \mathbf{X}(x_i) - \mathbf{X}(x_0) \right\} \\ &= 0 \end{aligned}$$

quelque soit $\mathbf{a}$. Il faut donc K conditions de non biais :

$$\sum_{i=1}^n \lambda_i X_k(x_i) - X_k(x_0) \quad \forall k = 1, \ldots, K$$

Cas général : Krigeage avec régression (suite)

On minimise

$$\mathrm{Var}(Z_0^* - Z_0) = \mathrm{Var}(Y_0^* - Y_0)$$

sous les K conditions

$$\mathbf{\Lambda}^t \mathbf{X}(x_i) = \mathbf{X}(x_0)$$

On utilise K pondérateurs de Lagrange $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_K)$

Cela mène au système

$$\begin{pmatrix} \mathbf{C} & \mathbf{X} \\ \mathbf{X}^t & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{\Lambda} \\ \boldsymbol{\nu} \end{pmatrix} = \begin{pmatrix} C_0 \\ \mathbf{X}(x_0) \end{pmatrix}$$

où $\mathbf{X}$ est ma matrice $n \times K$ des covariables aux points de données.

Cas général : Krigeage avec régression (suite)

$$\begin{pmatrix} \mathbf{C} & \mathbf{X} \\ \mathbf{X}^t & \mathbf{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{\Lambda} \\ \boldsymbol{\nu} \end{pmatrix} = \begin{pmatrix} C_0 \\ \mathbf{X}(x_0) \end{pmatrix}$$

Un peu d'algèbre linéaire ! ! On trouve

$$\boldsymbol{\Lambda} = \mathbf{C}^{-1} C_0 - \mathbf{C}^{-1}\mathbf{X}(\mathbf{X}^t\mathbf{C}^{-1}\mathbf{X})^{-1}[\mathbf{X}^t\mathbf{C}^{-1}C_0 - \mathbf{X}(x_0)]$$

et

$$\sigma_K^2(x_0) = C(0) - C_0^t\mathbf{C}^{-1}C_0 + [\mathbf{X}^t\mathbf{C}^{-1}C_0 - \mathbf{X}(x_0)]^t(\mathbf{X}^t\mathbf{C}^{-1}\mathbf{X})^{-1}[\mathbf{X}^t\mathbf{C}^{-1}C_0 - \mathbf{X}(x_0)]$$

Note : si $\mathbf{X} = \mathbf{1}$, on retrouve le krigeage ordinaire.

Validation

Peut-on valider un modèle ?

Deux familles d'approches

1. Le bootstrap paramétrique : on simule selon le modèle estimé et on compare les statistiques calculées sur les simulations à celles calculées sur les échantillons.
2. La validation croisée : on écarte à tour de rôle chaque donnée, et on l'estime par krigeage à l'aide des autres données. On compare cette estimation avec la vraie valeur.

Boostrap paramétrique

On a

- Variogramme empirique $\hat{\gamma}^{(0)}(h_k)$ calculés pour des classes de distance $(h_k)_{k=1,\dots,K}$
- Variogramme théorique $\gamma(h; \hat{\theta})$

Procédure

1. Réaliser N simulations $\mathbf{Z}^{(j)}$ aux points de données $x_1, \dots, x_n$, avec le variogramme $\gamma(h; \hat{\theta})$.
2. Pour chaque simulation j, calculer $\{\hat{\gamma}^{(j)}(h_k)\}$ pour les mêmes classes de distance $(h_k)_{k=1,\dots,K}$.
3. Calculer les enveloppes inférieures $\hat{\gamma}_{inf}(h_k) = \min_{j=1,\dots,N} \hat{\gamma}^{(j)}(h_k)$ et supérieures $\hat{\gamma}_{sup}(h_k) = \max_{j=1,\dots,N} \hat{\gamma}^{(j)}(h_k)$.
4. Ces deux enveloppes définissent un intervalle de confiance approximatif empirique de niveau $1 - 2/(N+1)$, car

$$P(\hat{\gamma}^{(0)}(h_k) < \hat{\gamma}_{inf}(h_k)) = P(\hat{\gamma}^{(0)}(h_k) > \hat{\gamma}_{sup}(h_k) \leq 1/(N+1).$$

Boostrap paramétrique (suite)

Si le variogramme empirique initial $\hat{\gamma}^{(0)}(h_k)$ sort de l'enveloppe $[\hat{\gamma}_{inf}(h_k), \hat{\gamma}_{sup}(h_k)]$ pour l'une des distances h_k, on rejette le modèle $\gamma(h;\theta)$ (et pas seulement l'estimation en $\hat{\theta}$).

Avantages

- Simple à comprendre
- Facile à mettre en œuvre

Inconvénients

- Vecteurs distribués selon une loi multigaussienne
- Tests multiples
- Ne permet pas de choisir parmi un ensemble de modèles valides
  $\Rightarrow$ validation croisée

Validation croisée

Principe général :

- écarter p données et les réestimer à l'aide des $n - p$ restantes
- répéter en modifiant l'ensemble des p valeur à réestimer
- en théorie, il faut recalculer le variogramme et réestimer θ
- souvent, $p = 1$ (leave-one-out).
  $\Rightarrow$ On fait l'approximation que le variogramme reste identique lorsqu'une donnée est retirée

Validation croisée (suite)

Procédure

1. Pour $i = 1, \dots, n$, calculer le krigeage Z_i^k et la variance de krigeage associée $(\sigma_i^K)^2$
2. Calculer les indicateurs

$$
\begin{aligned}
BIAIS &= 1/n \sum_{i=1}^{n} (Z_i^K - Z_i), \\
EQM &= 1/n \sum_{i=1}^{n} (Z_i^K - Z_i)^2, \\
EQNM &= 1/n \sum_{i=1}^{n} (Z_i^K - Z_i)^2 / \sigma_{K,i}^2.
\end{aligned}
$$

Validation croisée (suite)

Si on suppose

$$(Z_i^K - Z_i)/\sigma_{K,i} \sim \mathcal{N}(0,1),$$

on doit avoir

$$q(n, \alpha/2) \leq nEQNM \leq q(n, 1-\alpha/2) \qquad (*)$$

où $q(n, \beta)$ désigne le β-quantile d'une χ_n^2.

On écarte les modèles qui ne vérifient pas $(*)$.
Parmi les modèles qui vérifient $(*)$ on garde le modèle sans biais avec EQM minimum.