COURS 9 - Interpolations spatiales probabilistes

Statistiques spatiales et SIG

Arlette Antoni

Université de Bretagne Sud

Année Universitaire 2021 -2022

Les méthodes déterministes ne prennent pas en compte la structure spatiale globale
mais
de manière très locale via des moyennes pondérées sur un voisinage.
Le Krigeage :

- Krige (1951) et Matheron (1962)
- Cressie (1990)

Le Krigeage

$$\hat{z}(s0) = \sum_{r \in Vois(s_0)} \lambda_r z(s_r)$$

Mais les λ_r calculés en vue :

- d'une prévision non biaisée :

$$E(\hat{Z}_{s0} - Z_{s0})$$

- de variance minimale de l'erreur :

$$var(\hat{Z}_{s0} - Z_{s0})$$

- en fonction de la localisation des observations

de leur structure de dépendance spatiale.

- la meilleure prévision linéaire sans bais (cf BLUE)

C'est une méthode souple qui peut être soit locale soit globale en fonction du voisinage.

Modèle de base

régression

$$Z_s = \mu_s + \delta_s$$

- μ partie déterministe (espérance)
- l'erreur δ stationnaire
- d'espérance nulle
- et de covariance connue

Types de modèles

selon la nature de μ

- **Krigeage simple**

$$\mu_s = m$$

avec m connue

- **Krigeage ordinaire**

$$\mu_s = \mu$$

avec μ inconnue

- **Krigeage universel**

$$\mu_s = \sum_{j=1}^{p} \beta_j f_j(s)$$

- directement estimable à partir des $z(s)$

δ est inconnu :

- **appel au "variogramme"**
- **introduit une erreur de mesure**
- **introduction de postulats pour pallier l'unicité de l'observation**

Stationnarité de δ :
Stationnarité forte : la loi de probabilité de la fonction aléatoire est invariante par translation

- **Faible en krigeage**
- **à l'ordre 1**

$$E(\delta(s)) = m = 0$$

- **à l'ordre 2**

$$Cov(\delta(s), \delta(s+h)) = C(h)$$

- **Covariogramme**

Postulats

Stationnarité intrinsèque

Porte sur les accroissements

- **à l'ordre 1**

$$E(\delta(s+h)-\delta(s)) = m = 0$$

- **à l'ordre 2**

$$var(\delta(s)-\delta(s+h)) = \gamma(h)$$

- **Variogramme**
- si second ordre alors intrinséque

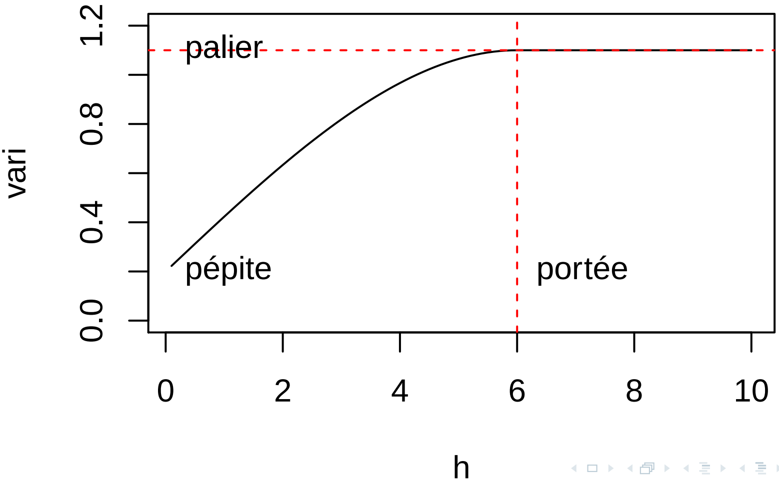
$$2\gamma(h) = 2(C(0)-C(h))$$

- plus facile car pas besoin d'estimer l'espérance.
- On travaille avec le semi variogramme.

La fonction est paire

$$\gamma(h) = \gamma(-h)$$

Caractéristique d'un semi variogramme



- **Si Isotrope**
- $\gamma(h)$ ne dépend que de $|h|$
- sans direction

$$\gamma(h) = 1/2(var(\delta(s+h) - \delta(s))$$

$$= 1/2E(Z(s) - Z(s+h))^2 - 1/2(\mu(s) - \mu(s+h))^2$$

- **Cas Simple et ordinaire**
- $\gamma(h)$ dépend de h

$$(\mu(s) - \mu(s+h))^2 = 0$$

car $\mu = Cte$

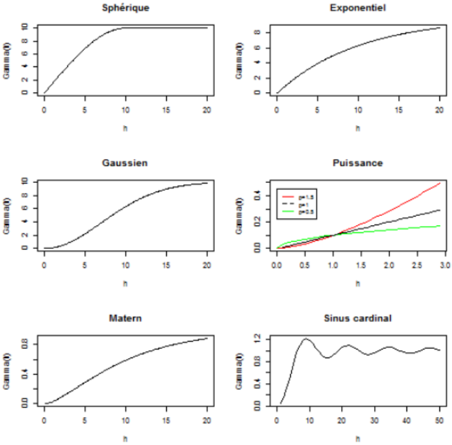
Formule la plus connue

$$\gamma(\hat{h}) = \frac{1}{2(|N(h)|)} \sum_{N(h)} (z(s_i) - z(s_j))^2$$

Avec $N(h)$ l'ensemble des couples de pointsi, j de distances $h + / - b_h$.
On ajoute b_h dans les cas où les points ne sont pas régulièrement répartis.

Variogramme théorique

Les plus courants



Variogramme théorique : estimation

Estimateurs

Ajustement de la courbe

- Par les moindres carrés
- par le maximum de vraisemblance
- par des méthodes bayésiennes

Le nombre d'observations pour calculer un point du variogramme expérimental est donc d'importance.

La courbe peut être de forme quand h>0

- sphérique : $c_0 + c_s(\frac{3}{2}(\frac{h}{a}) - \frac{1}{2}(\frac{h}{a})^3)$ pour $h \leq a$ et $c_0 + c_s$ si $h > a$
- exponentiel : $c_0 + c_s(1 - e^{-h/a})$
- gaussien : $c_0 + c_s(1 - e^{-(h/a)^2})$
- puissance : $c_0 + bh^p$
- matern : $c_s + (1 - \frac{h/a}{2^{\alpha-1}\Gamma(\alpha)}K_\alpha(h/a)$

avec c_0 la pépite, a : la portée, $c_0 + c_s$ le palier

Prédiction en un point :

$$Z(\hat{s_0}) = \sum_{i=1}^{n} \lambda_i Z(s_i)$$

On ajoute la contrainte $\sum_{i=1}^{n} \lambda_i$ sinon l'espérance ne serait pas m.
On minimise alors avec la contrainte et en dérivant par rapport aux λ_i :

$$E(\hat{Z}(s_0) - Z(s_0))^2$$

On trouve

$$\sum_{j=1}^{n} \lambda_j \gamma(s_j - s_j) + m = \gamma(s_0 - s_i)$$

On obtient les valeur des λ_j en inversant

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_N \\ m \end{pmatrix} = \begin{pmatrix} \gamma(s_1 - s_1) & \gamma(s_1 - s_2) & \dots & \gamma(s_1 - s_N) & 1 \\ \gamma(s_2 - s_1) & \gamma(s_2 - s_2) & \dots & \gamma(s_2 - s_N) & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \gamma(s_N - s_1) & \gamma(s_N - s_2) & \dots & \gamma(s_N - s_N) & 1 \\ 1 & 1 & \vdots & 1 & 0 \end{pmatrix}^{-1} x \begin{pmatrix} \gamma(s_0 - s_1) \\ \gamma(s_0 - s_2) \\ \vdots \\ \gamma(s_0 - s_N) \\ 1 \end{pmatrix}$$

Et en nommant ces matrices Λ_0 , $\Gamma_{i,j}$, Γ_0

${}^t\Gamma_0$
cad
On obtient une écriture matricielle simplifiée:

$$\Lambda_0 = \Gamma_{i,j}^{-1}\Gamma_0$$

Les γ sont récupérés du semi variogramme théorique.
L'erreur quadratique moyenne ou variance de krigeage est alors :$\Lambda_0^t\Gamma_0$ ou

$$\sum_{j=1}^{N} \lambda_j \gamma(s_0 - s_j)$$

en chaque point de la grille

ET on reconstitue pour trouver les interpolations en s_0 : $Z(\hat{s_0}) = \sum_{i=1}^{n} \lambda_i Z(s_i)$
Rq : ici l'interpolation au point est la valeur observée en ce point

$$RMSE = \sqrt{\sum_{j=1}^{N} (\hat{y}_j - y_j)^2 / N}$$