Prédiction spatiale : le krigeage

Marie-Pierre Etienne

https://marieetienne.github.io/statspat

2026-11-12

Introduction

Sources

Ce cours est construit principalement à partir de celui de Liliane Bel et des livres de Gaetan et al. (2010) et Cressie (2015).

Champs aléatoires et processus

Un processus stochastique est une famille de variables aléatoires indexée par un ensemble D :

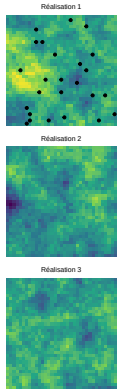
$$\{Z(s),\ s \in D\}$$

- $D \subset \mathbb{R}$ (le temps) : on parle de **processus**,
- $D \subset \mathbb{R}^d$, $d \geq 2$ (l'espace) : on parle de **champ aléatoire**.

Une réalisation est une fonction $s \mapsto z(s)$ sur une carte.

La loi du champ est caractérisée par ses lois fini-dimensionnelles : celles de $(Z(s_1), \dots, Z(s_n))$ pour tous n et tous $s_1, \dots, s_n$.

Champ gaussien : toutes ces lois sont gaussiennes, donc la loi est entièrement déterminée par la moyenne $\mu(s)$ et la covariance $c(s, s')$.



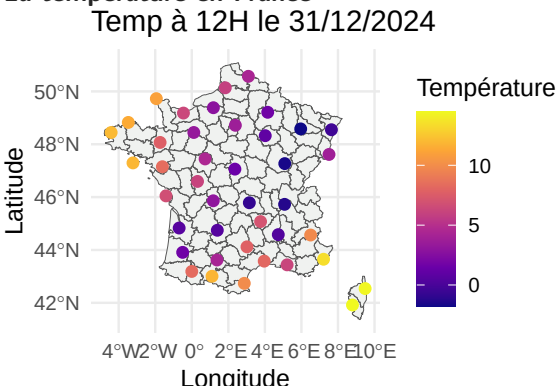
Le problème

On observe $z(s_1), \dots, z(s_n)$ aux sites $s_1, \dots, s_n$ d'un domaine D et on veut :

- prédire $Z(s_0)$ en un site s_0 non observé,
- construire une carte de prédiction sur tout D,
- quantifier l'incertitude associée à cette carte.

Exemple

La température en France



Pourquoi ne pas simplement interpoler ?

Interpolateur par inverse de la distance :

$$\widehat{Z}(s_0) = \sum_{i=1}^{n} \lambda_i Z(s_i), \qquad \lambda_i \propto \frac{1}{\|s_0 - s_i\|^p}$$

Limites

- Aucune mesure d'incertitude.
- Le poids d'un site ne dépend pas de la position des autres sites (deux sites collés comptent double).
- Les poids ne s'adaptent pas à la structure de dépendance des données (portée, régularité).

Idée du krigeage : choisir les poids à partir d'un **modèle de la dépendance spatiale**.

Cadre : champ aléatoire du second ordre

Ce dont on a besoin pour prédire

On modélise Z comme un champ aléatoire du second ordre :

- moyenne $\mu(s) = \mathbb{E}(Z(s))$,
- covariance $c(s, s') = \mathbb{C}\text{ov}(Z(s), Z(s'))$.

Notations

- $\mathbf{Z} = (Z(s_1), \dots, Z(s_n))^T$, $\widehat{Z}_0 = \widehat{Z}(s_0)$
- Σ : matrice $n \times n$, $\Sigma_{ij} = c(s_i, s_j)$
- c_0 : vecteur de taille n, $c_{0,i} = c(s_0, s_i)$
- $\sigma_0^2 = \mathbb{V}\text{ar}(Z(s_0))$

Dans ce cours : on suppose la moyenne et la covariance **connues**. La question de leur estimation est repoussée à la fin.

> Les formules du krigeage ne font intervenir que Σ, c_0 et σ_0^2, quelle que soit la forme de la covariance $c(s, s')$.

Covariance utilisée dans les illustrations

Une covariance qui décroît avec la distance entre les sites :

$$c(s, s') = \sigma^2 \exp\left(-\|s - s'\|/\rho\right) \; + \; \tau^2 \mathbb{1}_{s=s'}$$

- σ^2 : variance du signal,
- ρ : vitesse de décroissance de la corrélation avec la distance (« portée »),
- τ^2 : bruit de mesure (« pépite »).

On note $C(h) = \sigma^2 e^{-h/\rho}$ ($+ \tau^2$ si $h = 0$), de sorte que $c(s, s') = C(\|s - s'\|)$: la covariance ne dépend que de la distance. Ce choix n'est pas anodin, on y reviendra.

```
cov_exp <- function(h, sigma2 = 1, rho = 1) sigma2 * exp(-h / rho)
cov_gauss <- function(h, sigma2 = 1, rho = 1) sigma2 * exp(-(h / rho)^
```

Krigeage simple : le cas gaussien

Hypothèses

- μ **connue** et C **connue**,
- Z est un champ gaussien : tout vecteur $(Z(s_0), Z(s_1), \dots, Z(s_n))$ est gaussien.

$$\begin{pmatrix} Z(s_0) \\ \mathbf{Z} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mu \\ \mu\mathbf{1} \end{pmatrix}, \begin{pmatrix} \sigma_0^2 & c_0^T \\ c_0 & \Sigma \end{pmatrix} \right)$$

Question

Quelle est la loi de $Z(s_0)$ sachant $\mathbf{Z} = \mathbf{z}$?

Conditionnement gaussien

$$Z(s_0) \mid \mathbf{Z} = \mathbf{z} \sim \mathcal{N}\left(\mu + {c_0}^T \Sigma^{-1}(\mathbf{z} - \mu\mathbf{1}),\ \sigma_0^2 - {c_0}^T \Sigma^{-1} c_0\right)$$

Prédicteur de krigeage simple

$$\widehat{Z}_{KS}(s_0) = \mathbb{E}(Z(s_0) \mid \mathbf{Z}) = \mu + {c_0}^T \Sigma^{-1}(\mathbf{Z} - \mu\mathbf{1})$$

Variance de krigeage

$$\sigma^2_{KS}(s_0) = \mathbb{V}\mathsf{ar}(Z(s_0) \mid \mathbf{Z}) = \sigma_0^2 - {c_0}^T \Sigma^{-1} c_0$$

Bonus du cadre gaussien : on a la loi prédictive complète, donc des intervalles de prédiction $\widehat{Z}_{KS}(s_0) \pm 1.96\, \sigma_{KS}(s_0)$, et toute quantité $P(Z(s_0) > \zeta \mid \mathbf{Z})$.

Interprétation

$$\widehat{Z}_{KS}(s_0) = \mu + \sum_{i=1}^{n} \lambda_i (Z(s_i) - \mu), \qquad \lambda = \Sigma^{-1} c_0$$

- On part de la moyenne μ et on la **corrige** en fonction des écarts observés à la moyenne.
- Les poids λ_i dépendent de la covariance entre s_0 et les sites (c_0) **et** de la redondance entre les sites (Σ^{-1}).
- Les données font **baisser** la variance : $\sigma^2_{KS}(s_0) \leq \sigma^2_0$.

Exercice : une observation

On observe un seul site s_1. On note $r = C(\|s_0 - s_1\|)/C(0)$ la corrélation entre $Z(s_0)$ et $Z(s_1)$.

1. Donner $\widehat{Z}_{KS}(s_0)$ et $\sigma^2_{KS}(s_0)$.
2. Que se passe-t-il quand $s_0 \to s_1$? Quand $\|s_0 - s_1\| \to \infty$?

Exercice : une observation

On observe un seul site s_1. On note $r = C(\|s_0 - s_1\|)/C(0)$ la corrélation entre $Z(s_0)$ et $Z(s_1)$.

1. Donner $\widehat{Z}_{KS}(s_0)$ et $\sigma^2_{KS}(s_0)$.
2. Que se passe-t-il quand $s_0 \to s_1$? Quand $\|s_0 - s_1\| \to \infty$?

Solution

$$\widehat{Z}_{KS}(s_0) = \mu + r\,(Z(s_1) - \mu), \qquad \sigma^2_{KS}(s_0) = C(0)\,(1 - r^2)$$

- $s_0 \to s_1$ (sans pépite) : $r \to 1$, on retrouve $Z(s_1)$ et une variance nulle.
- s_0 loin : $r \to 0$, on prédit μ avec la variance $C(0)$.

Exercice : deux observations

On se place en 1D : $s_0 = 0$, $s_1 = 1$, $s_2 = 2$, $\mu = 0$.

Calculer les poids λ_1, λ_2 pour

1. une covariance exponentielle,
2. une covariance gaussienne.

```
poids_ks <- function(s0, s, covf, ...) {
  Sigma <- covf(as.matrix(dist(s)), ...)
  c0 <- covf(abs(s - s0), ...)
  solve(Sigma, c0)
}
round(poids_ks(0, c(1, 2), cov_exp, rho = 1), 3)
```

```
    1     2
0.368 0.000
```

```
round(poids_ks(0, c(1, 2), cov_gauss, rho = 1), 3)
```

```
     1      2
 0.418 -0.135
```

Effet d'écran et poids négatifs

- Effet d'écran : le site s_1 « masque » s_2. En 1D avec une covariance exponentielle, $\lambda_2 = 0$ exactement (propriété de Markov du processus d'Ornstein-Uhlenbeck).
- Poids négatifs : avec une covariance très régulière (gaussienne), le poids de s_2 est négatif. Le krigeage **extrapole** la tendance locale.

> Contrairement à l'inverse de la distance, on n'a pas forcément $0 \leq \lambda_i \leq 1$: le prédicteur peut sortir de l'intervalle des valeurs observées.

Krigeage simple sans hypothèse gaussienne

Le meilleur prédicteur linéaire (BLUP)

On ne suppose plus que Z est gaussien. Seuls les deux premiers moments sont connus.

On cherche un prédicteur **linéaire** $\widehat{Z}(s_0) = \alpha + \lambda^T \mathbf{Z}$ qui minimise le risque quadratique

$$R(\alpha, \lambda) = \mathbb{E}\left[(Z(s_0) - \alpha - \lambda^T \mathbf{Z})^2\right].$$

Exercice

1. Montrer que $R(\alpha, \lambda) = (\mu - \alpha - \mu\lambda^T \mathbf{1})^2 + \sigma_0^2 - 2\lambda^T c_0 + \lambda^T \Sigma \lambda$.
2. En déduire α et λ optimaux.

Solution

$$\lambda = \Sigma^{-1} c_0, \qquad \alpha = \mu(1 - \lambda^T \mathbf{1})$$

On retrouve exactement le même prédicteur et la même variance d'erreur :

$$\widehat{Z}_{KS}(s_0) = \mu + {c_0}^T \Sigma^{-1}(\mathbf{Z} - \mu \mathbf{1}), \qquad \mathbb{E}\left[(Z(s_0) - \widehat{Z}_{KS}(s_0))^2\right] = \sigma_0^2 - {c_0}^T \Sigma^{-1} c_0$$

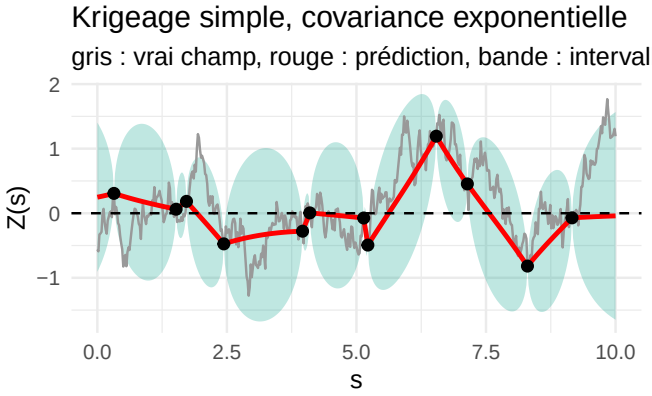
Gaussien ou pas : ce qui change

	Cas gaussien	Second ordre seulement
Prédicteur	$\widehat{Z}_{KS}(s_0)$	$\widehat{Z}_{KS}(s_0)$
Optimal parmi	**tous** les prédicteurs	les prédicteurs **linéaires**
$\sigma^2_{KS}(s_0)$ est	la variance conditionnelle	l'erreur quadratique moyenne
Intervalle $\pm 1.96\,\sigma_{KS}$	exact	approximatif
$P(Z(s_0) > \zeta \mid \mathbf{Z})$	disponible	non disponible

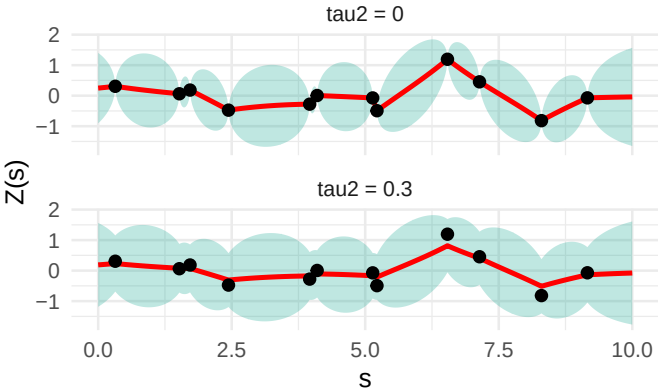
À retenir : l'hypothèse gaussienne ne change pas la formule, elle change ce qu'on peut **en dire**.

Propriétés du krigeage

Une illustration en 1D

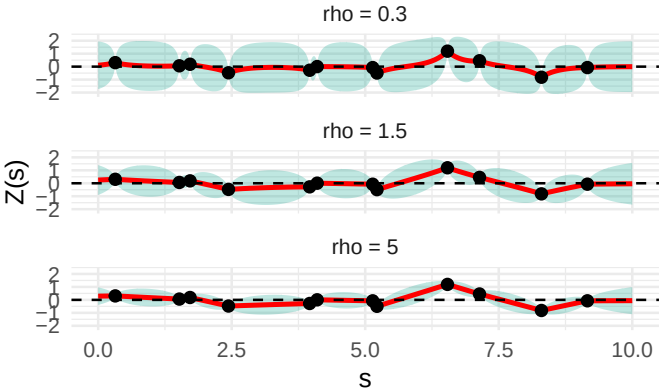


Interpolation exacte et effet pépite



- Sans pépite : $\widehat{Z}(s_i) = Z(s_i)$ et $\sigma^2_{KS}(s_i) = 0$, le krigeage est un interpolateur exact.
- Avec pépite : le krigeage **lisse**, on ne fait plus confiance à 100% aux observations.

Effet de la portée



Loin des données (à l'échelle de la portée) : $\widehat{Z}(s_0) \to \mu$ et $\sigma^2_{KS}(s_0) \to C(0)$.

La variance de krigeage ne dépend pas des valeurs observées

$$\sigma^2_{KS}(s_0) = \sigma^2_0 - {c_0}^T \Sigma^{-1} c_0$$

ne dépend que :

- de la **position** des sites $s_1, \ldots, s_n$ et de s_0,
- du **modèle** de covariance.

Conséquences

- On peut calculer la carte d'incertitude avant de collecter les données : c'est un outil de **plan d'échantillonnage** (où placer une nouvelle station ?).
- L'incertitude ne réagit pas à des valeurs surprenantes, si le modèle est faux.

Krigeage ordinaire

Moyenne inconnue

On suppose maintenant μ inconnue (constante), C connue.

On cherche $\widehat{Z}(s_0) = \lambda^T \mathbf{Z}$:

- **sans biais** pour tout μ : $\mathbb{E}(\widehat{Z}(s_0)) = \mu \Leftrightarrow \lambda^T \mathbf{1} = 1$,
- de **variance d'erreur minimale** : $\min\ \sigma_0^2 - 2\lambda^T c_0 + \lambda^T \Sigma \lambda$.

Exercice

Résoudre ce problème par la méthode du multiplicateur de Lagrange.

Solution

$$\lambda_{KO} = \Sigma^{-1}(c_0 + m\mathbf{1}), \qquad m = \frac{1 - \mathbf{1}^T\Sigma^{-1}c_0}{\mathbf{1}^T\Sigma^{-1}\mathbf{1}}$$

Relecture

Avec $\hat{\mu} = \dfrac{\mathbf{1}^T\Sigma^{-1}\mathbf{Z}}{\mathbf{1}^T\Sigma^{-1}\mathbf{1}}$ l'estimateur des moindres carrés généralisés de μ (c'est aussi le meilleur estimateur linéaire sans biais, BLUE) :

$$\widehat{Z}_{KO}(s_0) = \hat{\mu} + {c_0}^T\Sigma^{-1}(\mathbf{Z} - \hat{\mu}\mathbf{1})$$

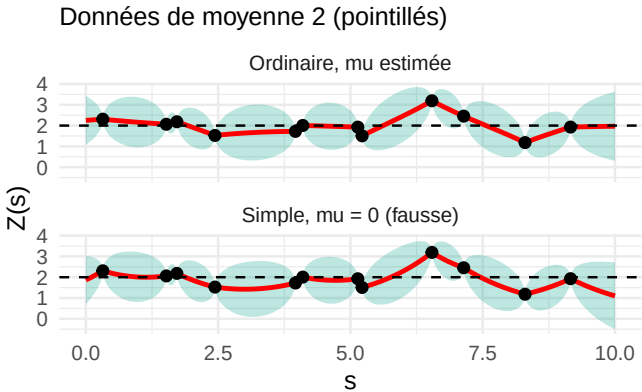
Krigeage ordinaire = krigeage simple dans lequel on remplace μ par son estimateur BLUE.

Variance de krigeage ordinaire

$$\sigma^2_{KO}(s_0) = \underbrace{\sigma_0^2 - {c_0}^T \Sigma^{-1} c_0}_{\sigma^2_{KS}(s_0)} + \underbrace{\frac{(1 - \mathbf{1}^T \Sigma^{-1} c_0)^2}{\mathbf{1}^T \Sigma^{-1} \mathbf{1}}}_{\text{coût de l'estimation de } \mu}$$

- $\sigma^2_{KO}(s_0) \geq \sigma^2_{KS}(s_0)$.
- Près des données, $\mathbf{1}^T \Sigma^{-1} c_0$ est proche de 1 et le surcoût est faible.
- Loin des données, on prédit $\hat{\mu}$ et on paye $\mathbb{V}\text{ar}(\hat{\mu}) = 1/(\mathbf{1}^T \Sigma^{-1} \mathbf{1})$.

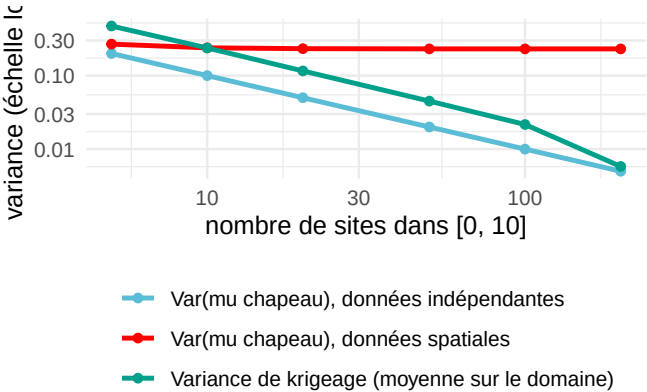
Simple vs ordinaire



Estimer la moyenne : une tâche ingrate

Le second terme de σ^2_{KO} fait apparaître $\mathbb{V}\text{ar}(\hat{\mu}) = 1/(\mathbf{1}^T\Sigma^{-1}\mathbf{1})$.

Que se passe-t-il quand on ajoute des stations **dans le même domaine** ?



- Densifier les observations n'apprend presque rien de plus sur μ : des données corrélées n'apportent pas n informations indépendantes (taille d'échantillon effective).

Krigeage universel

Une moyenne qui dépend de covariables

On suppose maintenant

$$Z(s) = X(s)^T\beta + \varepsilon(s), \qquad \mathbb{E}(\varepsilon(s)) = 0, \quad \mathbb{C}\text{ov}(\varepsilon(s), \varepsilon(s')) = c(s, s') \text{ connue,}$$

avec $X(s) \in \mathbb{R}^p$ connu en **tout** site (y compris s_0) et β inconnu.

- Covariables : altitude, distance à la rivière... (« dérive externe »)
- Coordonnées : $X(s) = (1, x, y)$ pour une tendance spatiale (« dérive interne »)

Notations

$\mathbf{X}$: matrice $n \times p$ des covariables aux sites observés, $X_0 = X(s_0)$.

Le problème d'optimisation

On cherche $\widehat{Z}(s_0) = \lambda^T \mathbf{Z}$:

- **sans biais pour tout** β : $\lambda^T \mathbf{X}\beta = X_0^T \beta \; \forall \beta \Leftrightarrow \mathbf{X}^T \lambda = X_0$ (p contraintes),
- de **variance d'erreur minimale** : $\min \; \sigma_0^2 - 2\lambda^T c_0 + \lambda^T \Sigma \lambda$.

Exercice

1. Écrire le lagrangien avec un vecteur de multiplicateurs $\mathbf{m} \in \mathbb{R}^p$.
2. Montrer que $\lambda = \Sigma^{-1}(c_0 + \mathbf{X}\mathbf{m})$ et déterminer $\mathbf{m}$.
3. Vérifier que le krigeage ordinaire est le cas $X(s) = 1$.

Solution

$$\mathbf{m} = (\mathbf{X}^T\Sigma^{-1}\mathbf{X})^{-1}(X_0 - \mathbf{X}^T\Sigma^{-1}c_0)$$

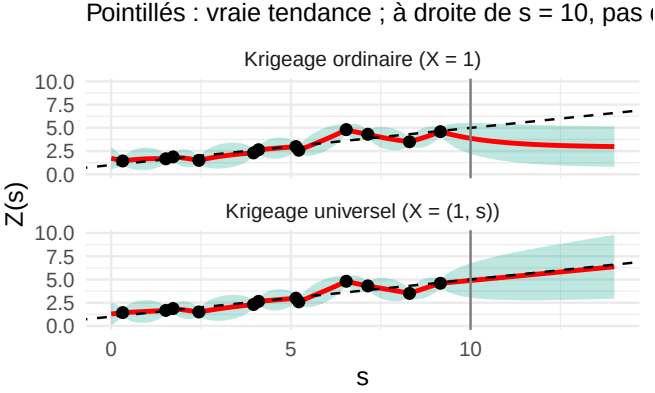
Relecture

Avec $\widehat{\beta} = (\mathbf{X}^T\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}^T\Sigma^{-1}\mathbf{Z}$ l'estimateur des moindres carrés généralisés :

$$\widehat{Z}_{KU}(s_0) = \underbrace{X_0^T\widehat{\beta}}_{\text{tendance}} + \underbrace{{c_0}^T\Sigma^{-1}(\mathbf{Z} - \mathbf{X}\widehat{\beta})}_{\text{krigeage simple des résidus}}$$

$$\sigma^2_{KU}(s_0) = \sigma^2_{KS}(s_0) + \underbrace{\left(X_0 - \mathbf{X}^T\Sigma^{-1}c_0\right)^T\left(\mathbf{X}^T\Sigma^{-1}\mathbf{X}\right)^{-1}\left(X_0 - \mathbf{X}^T\Sigma^{-1}c_0\right)}_{\text{coût de l'estimation de } \beta}$$

Illustration : tendance linéaire en 1D



Loin des données, le krigeage universel **suit la tendance**, avec une incertitude croissante (extrapolation de $\widehat{\beta}$).

Remarques

- Les covariables doivent être connues **partout** où l'on prédit (sur toute la grille).
- Cercle vicieux : pour estimer β par moindres carrés généralisés, il faut la covariance des résidus ; pour estimer la covariance des résidus, il faut β.
  - en pratique : $\widehat{\beta}$ par moindres carrés ordinaires, variogramme des résidus, puis krigeage universel (ce que fait `gstat`),
  - ou estimation conjointe par maximum de vraisemblance (restreint, REML).
- Séparer tendance et dépendance est en partie un **choix de modélisation** : une tendance mal spécifiée se retrouve dans le variogramme.

D'où vient la covariance ?

La question qui reste

En pratique, ni μ ni la covariance ne sont connues.

- Il faudrait connaître $c(s_i, s_j)$ pour **chaque paire** de sites, et $c(s_0, s_i)$ pour un site **jamais observé**.
- Or on n'a souvent qu'une seule réalisation du champ : un seul couple $(Z(s_i), Z(s_j))$ par paire de sites.

Discussion

- Peut-on estimer $\mathbb{C}\text{ov}(Z(s_i), Z(s_j))$ à partir d'une seule observation de $(Z(s_i), Z(s_j))$?
- Quelle hypothèse permettrait de « mutualiser » l'information entre paires de sites ?

Toute fonction ne peut pas être une covariance

Z est un processus du second ordre si $\forall s \in D,\ \mathbb{E}(Z(s)^2) < \infty$, de moyenne $m(s) = \mathbb{E}(Z(s))$ et de covariance $c(s,t) = \mathbb{C}\text{ov}(Z(s), Z(t))$.

Propriété Une fonction de covariance est semi-définie positive

$$\forall n,\ \forall s_1, \dots, s_n \in D,\ \forall a \in \mathbb{R}^n, \quad \sum_{i,j=1}^{n} a_i a_j\, c(s_i, s_j) \geq 0$$

Idée de preuve : regarder $\mathbb{V}\text{ar}\left(\sum_{i=1}^{n} a_i Z(s_i)\right)$.

C'est ce qui garantit que les variances de krigeage sont positives et que Σ est inversible (si définie positive). On ne peut donc pas choisir n'importe quelle fonction : on travaille avec des **familles de covariances valides**.

Stationnarité du second ordre

Définition

Un processus du second ordre Z est stationnaire sur $D \subset \mathbb{R}^d$ si

1. sa moyenne est constante : $\forall s \in D,\ m(s) = \mu$,
2. sa covariance est invariante par translation :

$$\forall s, s+h \in D, \quad c(s, s+h) = C(h).$$

Isotropie

Le processus est isotrope si $C(h)$ ne dépend que de $\|h\|$: $C(h) = C(\|h\|)$.
Conséquence : toutes les paires de sites séparées par h renseignent sur la même quantité $C(h)$. C'est le choix fait dans les illustrations précédentes.

Stationnarité des accroissements

Z est à accroissements stationnaires (ou intrinsèque) si

$$\mathbb{E}(Z(s+h) - Z(s)) = 0, \qquad \mathbb{V}\text{ar}(Z(s+h) - Z(s)) = 2\gamma(h).$$

- La stationnarité du second ordre implique la stationnarité des accroissements.
- La classe des processus intrinsèques est donc plus **large** que celle des processus stationnaires.

Exemple

Que dire du mouvement brownien sur $\mathbb{R}$?

Stationnarité des accroissements

Z est à accroissements stationnaires (ou intrinsèque) si

$$\mathbb{E}(Z(s+h) - Z(s)) = 0, \qquad \mathbb{V}\text{ar}(Z(s+h) - Z(s)) = 2\gamma(h).$$

- La stationnarité du second ordre implique la stationnarité des accroissements.
- La classe des processus intrinsèques est donc plus **large** que celle des processus stationnaires.

Exemple

Que dire du mouvement brownien sur $\mathbb{R}$?
$\mathbb{V}\text{ar}(B(t)) = t$: il n'est pas stationnaire. Mais $\mathbb{V}\text{ar}(B(t+h) - B(t)) = |h|$: il est intrinsèque, avec $\gamma(h) = |h|/2$.

Le variogramme

Définition

Le semi-variogramme d'un processus intrinsèque est

$$\gamma(h) = \frac{1}{2}\mathbb{V}\text{ar}(Z(s+h) - Z(s)) = \frac{1}{2}\mathbb{E}\left[(Z(s+h) - Z(s))^2\right].$$

Propriétés

- $\gamma(h) \geq 0$, $\gamma(0) = 0$, $\gamma(-h) = \gamma(h)$.
- Si Z est stationnaire du second ordre :

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

- Un variogramme **borné** correspond (à une constante aléatoire près) à un processus stationnaire du second ordre, et sa limite vaut $C(0)$.

Pourquoi le variogramme ? Il ne fait pas intervenir la moyenne on peut l'estimer sans estimer μ.

Modèles classiques de variogramme

1. **Pépitique** : $\gamma(h) = C\,\mathbb{1}_{h\neq 0}$
2. **Exponentiel** : $\gamma(h) = C(1-\exp(-\|h\|/\rho))$
3. **Sphérique** :

$$\gamma(h) = \begin{cases} C\left(\frac{3}{2}\frac{\|h\|}{\rho} - \frac{1}{2}\left(\frac{\|h\|}{\rho}\right)^3\right) & \text{si } \|h\| \leq \rho \\ C & \text{si } \|h\| > \rho \end{cases}$$

4. **Gaussien** : $\gamma(h) = C(1-\exp(-\|h\|^2/\rho^2))$
5. **Puissance** : $\gamma(h) = C\|h\|^\alpha$, $0 < \alpha < 2$ (non borné : processus intrinsèque non stationnaire)

La famille de Matérn

$$C(h) = \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\sqrt{2\nu}\,\|h\|}{\rho} \right)^{\nu} K_\nu \left(\frac{\sqrt{2\nu}\,\|h\|}{\rho} \right), \qquad \gamma(h) = \sigma^2 - C(h)$$

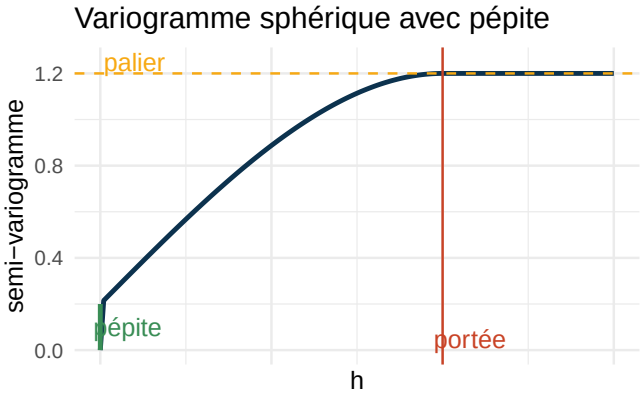
K_ν : fonction de Bessel modifiée de seconde espèce, d'ordre ν.

ν règle la régularité du processus (comportement du variogramme en 0) :

- $\nu = 1/2$: modèle exponentiel (trajectoires continues, non dérivables),
- $\nu \to \infty$: modèle gaussien (trajectoires infiniment dérivables).

C'est la famille utilisée par défaut dans INLA (approche SPDE).

Paramètres du variogramme



Interprétation des paramètres

- Palier $\lim_{h\to\infty}\gamma(h)$: s'il est fini, le processus est stationnaire et le palier vaut $\mathbb{V}\text{ar}(Z(s))$.
- Échelle ρ : vitesse à laquelle le variogramme rejoint le palier.
- Portée : distance au-delà de laquelle la corrélation est nulle ou négligeable (horizon de dépendance).
- Pépite $\tau^2 = \lim_{h\to 0^+}\gamma(h)$:
  - si $\tau^2 \neq 0$, variabilité à très petite échelle ou erreur de mesure (le krigeage lisse),
  - si $\tau^2 = 0$, la régularité du processus est donnée par le comportement de γ en 0 (paramètre ν de Matérn).

Krigeage et variogramme

Sous stationnarité, on remplace $C(h) = C(0) - \gamma(h)$ dans les formules.

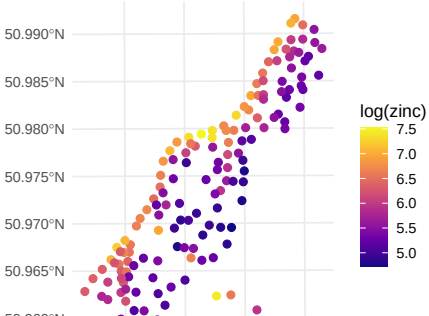
Remarque : le krigeage ordinaire (contrainte $\lambda^T \mathbf{1} = 1$) peut s'écrire **uniquement** avec γ. Il reste donc valable pour des processus intrinsèques non stationnaires. C'est la présentation historique de Matheron.

Estimation du variogramme

Exemple : les données Meuse

Concentrations en zinc (ppm) dans les sols de la plaine inondable de la Meuse, 155 sites (package sp).

```
library(sf)
library(gstat)
data(meuse, package = "sp")
meuse_sf <- st_as_sf(meuse, coords = c("x", "y"), crs = 28992)
ggplot(meuse_sf) + geom_sf(aes(col = log(zinc)), size = 2) +
  scale_color_viridis_c(option = "plasma")
```

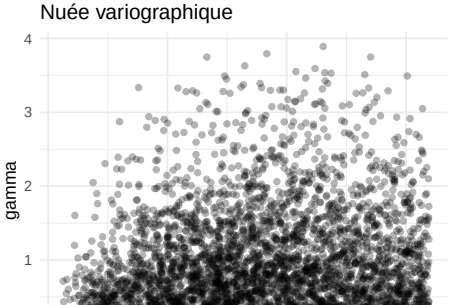


Nuée variographique

$$\gamma(h) = \frac{1}{2}\mathbb{E}\left[(Z(s+h) - Z(s))^2\right]$$

Pour chaque paire (i, j), on représente $\frac{1}{2}(Z(s_i) - Z(s_j))^2$ en fonction de $\|s_i - s_j\|$.

```
vario.cloud <- variogram(log(zinc) ~ 1, data = meuse_sf, cloud = TRUE
ggplot(vario.cloud, aes(x = dist, y = gamma)) + geom_point(alpha = 0.
  ggtitle("Nuée variographique")
```

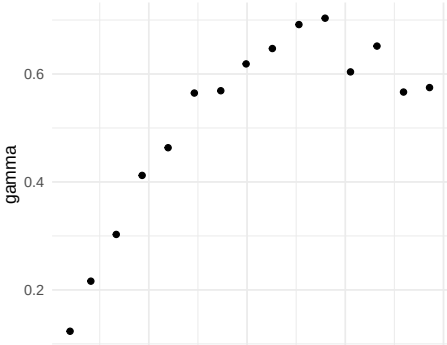


Variogramme empirique

On regroupe les paires par classes de distance d_k (de largeur δ), avec $N(k)$ l'ensemble des n_k paires de la classe k :

$$\tilde{\gamma}(d_k) = \frac{1}{2n_k} \sum_{(i,j) \in N(k)} (Z(s_i) - Z(s_j))^2$$

```
vario.b <- variogram(log(zinc) ~ 1, data = meuse_sf)
ggplot(vario.b, aes(x = dist, y = gamma)) + geom_point()
```



Propriétés du variogramme empirique

- Approximativement sans biais (à la largeur des classes près).
- Ne nécessite pas d'estimer la moyenne (si elle est constante).
- Si Z est gaussien, chaque terme est un $\gamma(h)\chi^2(1)$: estimateur très **sensible aux valeurs extrêmes**.
- Peu fiable aux grandes distances (peu de paires) : on se limite généralement à la moitié de la distance maximale.

Estimateur robuste (Cressie-Hawkins)

$$\tilde{\gamma}(d_k) = \frac{\left\{\frac{1}{n_k}\sum_{(i,j)\in N(k)} |Z(s_i) - Z(s_j)|^{1/2}\right\}^4}{2\left(0.457 + \frac{0.494}{n_k}\right)}$$

Ajustement d'un modèle de variogramme

Le variogramme empirique n'est pas un variogramme valide : on ajuste un modèle paramétrique γ_θ, avec $\theta = (\sigma^2, \rho, \nu, \tau^2)$.

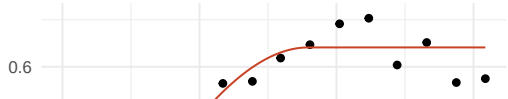
Moindres carrés pondérés

$$\hat{\theta} = \arg\min_\theta \sum_{k=1}^{K} w_k \left(\gamma_\theta(d_k) - \tilde{\gamma}(d_k)\right)^2, \qquad w_k = \frac{n_k}{\gamma_\theta(d_k)^2} \text{ (Cressie)}$$

```
v.fit <- fit.variogram(vario.b, vgm(model = "Sph", psill = 0.6, range
v.fit
```

```
  model      psill    range
1   Nug 0.05066522   0.0000
2   Sph 0.59061054 897.0412
```

```
ggplot(vario.b, aes(x = dist, y = gamma)) + geom_point() +
  geom_line(data = variogramLine(v.fit, maxdist = max(vario.b$dist))
```



Le krigeage en pratique

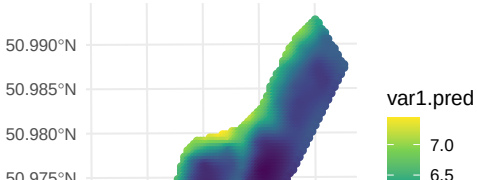
Démarche

1. Explorer les données, modéliser la **tendance** (moyenne) éventuelle.
2. Calculer le variogramme empirique (des résidus), vérifier l'isotropie.
3. Ajuster un modèle de variogramme.
4. Kriger sur une grille : carte de prédiction **et** carte de variance.
5. Valider (validation croisée : `krige.cv`).
6. Si besoin, simuler conditionnellement pour des quantités non linéaires.

```
data(meuse.grid, package = "sp")
grille <- st_as_sf(meuse.grid, coords = c("x", "y"), crs = 28992)
K <- krige(log(zinc) ~ 1, locations = meuse_sf, newdata = grille, mode
```

```
[using ordinary kriging]
```

```
ggplot(K) + geom_sf(aes(col = var1.pred)) + scale_color_viridis_c()
```



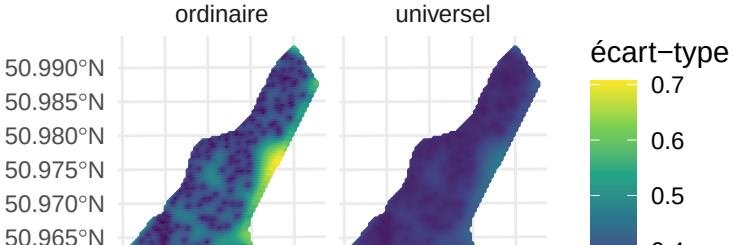
Krigeage universel avec gstat

Le zinc décroît avec la distance à la Meuse : on l'utilise comme covariable (connue sur toute la grille `meuse.grid`).

```
vario.res <- variogram(log(zinc) ~ sqrt(dist), data = meuse_sf)   # va
v.fit.res <- fit.variogram(vario.res, vgm(model = "Sph", psill = 0.2,
KU <- krige(log(zinc) ~ sqrt(dist), locations = meuse_sf, newdata = g
```

```
[using universal kriging]
```

```
bind_rows(K |> mutate(methode = "ordinaire"), KU |> mutate(methode =
  ggplot() + geom_sf(aes(col = sqrt(var1.var)), size = 0.5) + facet_wr
   scale_color_viridis_c(name = "écart-type")
```



Simulation conditionnelle

Pourquoi simuler ?

La carte de krigeage est plus lisse que la réalité :
$\mathbb{V}\mathrm{ar}(\widehat{Z}(s_0)) = \sigma_0^2 - \sigma_K^2(s_0) < \sigma_0^2$.

Elle donne la meilleure prédiction **site par site**, mais elle est mauvaise pour des quantités qui dépendent du champ **dans son ensemble** :

- surface du domaine où la concentration dépasse un seuil,
- maximum du champ sur une région,
- longueur d'un front, connectivité de zones polluées…

Pour une fonctionnelle non linéaire φ : $\varphi(\widehat{Z}) \neq \mathbb{E}(\varphi(Z) \mid \mathbf{Z})$.

Idée : tirer des réalisations du champ **sachant les observations** et calculer φ sur chacune.

Loi conditionnelle dans le cas gaussien

On veut simuler $\mathbf{Z}_G = (Z(g_1), \dots, Z(g_N))$ sur une grille, sachant $\mathbf{Z}$. Comme pour le krigeage simple :

$$\mathbf{Z}_G \mid \mathbf{Z} \sim \mathcal{N}\left(\mu\mathbf{1} + \mathbf{C}_G^T \Sigma^{-1}(\mathbf{Z} - \mu\mathbf{1}),\ \Sigma_{GG} - \mathbf{C}_G^T \Sigma^{-1} \mathbf{C}_G\right)$$

avec $\mathbf{C}_G$ la matrice $n \times N$ des covariances entre sites observés et points de grille.

- Moyenne : la carte de krigeage.
- Covariance : la covariance des erreurs de krigeage (sa diagonale est la variance de krigeage).

Simulation directe
$\mathbf{Z}_G = \widehat{Z}_G + \mathbf{L}\,\mathbf{U}$, avec $\mathbf{L}\mathbf{L}^T$ la covariance conditionnelle (Cholesky) et $\mathbf{U} \sim \mathcal{N}(0, I_N)$. Coût en $O(N^3)$: impossible sur une grande grille.

Conditionnement par krigeage

Propriété clé du cas gaussien : l'erreur de krigeage $Z(s) - \widehat{Z}(s)$ est **indépendante** des données (non corrélée et gaussienne).

Algorithme

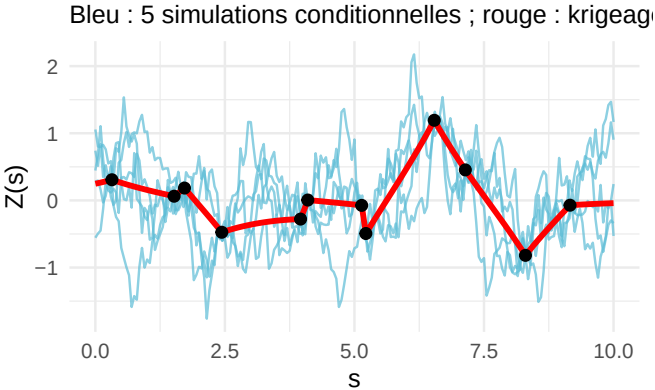
1. Simuler une réalisation **non conditionnelle** Z_s sur la grille et aux sites observés.
2. Kriger à partir des valeurs simulées $Z_s(s_1), \dots, Z_s(s_n)$: on obtient $\widehat{Z}_s$.
3. Poser

$$Z_{cs} = \underbrace{\widehat{Z}}_{\text{krigeage des vraies données}} + \underbrace{(Z_s - \widehat{Z}_s)}_{\text{erreur de krigeage simulée}}$$

Exercice

Montrer que Z_{cs} a la bonne loi conditionnelle et qu'elle interpole les données (sans pépite).

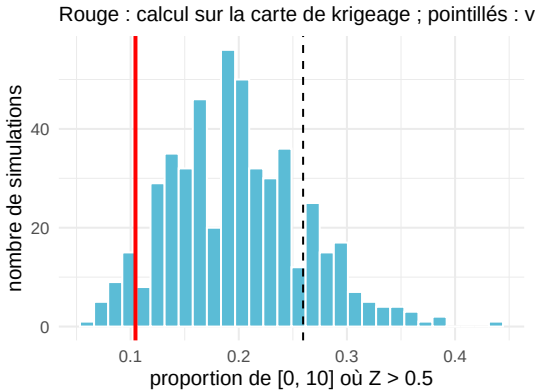
Illustration en 1D



Les simulations passent par les données et ont la **même rugosité** que le vrai champ, contrairement au krigeage.

Exemple : proportion du domaine au-dessus d'un seuil

On cherche la proportion de $[0, 10]$ où $Z(s) > 0.5$.



- Les 500 simulations donnent la loi conditionnelle de la proportion : estimation **et** incertitude.
- Le calcul sur la carte de krigeage, trop lisse, est biaisé.

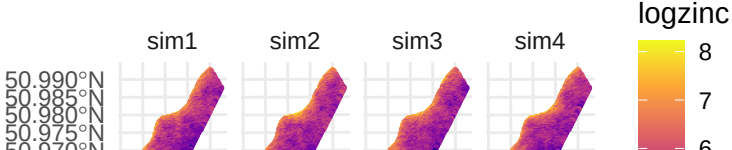
En pratique avec gstat

`krige(..., nsim = )` utilise la simulation séquentielle gaussienne : on parcourt la grille dans un ordre aléatoire et on simule chaque point selon sa loi conditionnelle sachant les données et les points déjà simulés (dans un voisinage `nmax`).

```
set.seed(1)
sim_meuse <- krige(log(zinc) ~ sqrt(dist), locations = meuse_sf, newd
                   model = v.fit.res, nsim = 4, nmax = 30)
```

```
drawing 4 GLS realisations of beta...
[using conditional Gaussian simulation]
```

```
sim_meuse |> pivot_longer(starts_with("sim"), names_to = "sim", valu
  st_as_sf() |> ggplot() + geom_sf(aes(col = logzinc), size = 0.3) + f
   scale_color_viridis_c(option = "plasma")
```



Conclusion

Limites

Le plug-in

On estime la covariance, puis on l'injecte dans les formules comme si elle était connue.

> La variance de krigeage plug-in ne tient pas compte de l'incertitude sur $\widehat{C}$: elle sous-estime l'incertitude réelle. L'approche bayésienne (INLA) permet de propager cette incertitude.

Et si les données ne sont pas gaussiennes ?

- Le BLUP reste valable, mais sans intervalles exacts.
- Souvent préférable : transformer (log, Box-Cox) ou utiliser un modèle hiérarchique (GLM spatial, voir INLA).

En pratique

Voir le TP sur les pluies en Suisse.

Références

Cressie, Noel. 2015. *Statistics for Spatial Data*. John Wiley & Sons.

Gaetan, Carlo, Xavier Guyon, et al. 2010. *Spatial Statistics and Modeling*. Vol. 90. Springer.