

Introduction à l'utilisation du

KRIGEAGE

pour la substitution d'un simulateur physico-numérique

Matthias De Lozzo

Document mis à jour le lundi 30 janvier 2017

- → De la nécessité d'un modèle de substitution
 - ➤ Modèle physico-numérique
 - Diverses applications d'un simulateur physico-numérique
 - Le coût calculatoire de ces simulateurs
 - Remplacer un simulateur par un métamodèle
- ♣ Présentation du métamodèle "krigeage"
 - Cadre bayésien : un métamodèle par processus gaussien
 - Cadre géostatistique : le meilleur prédicteur linéaire sans biais
 - ➤ Cadre fonctionnel : problème de régularisation dans un RKHS
- ♣ Pour aller plus loin...
 - ➤ Simuler une instance de processus gaussien
 - Analogies entre le krigeage et le filtre de Kalman





Exposé du problème

Tout commence par un code de calcul modélisant un phénomène φ .

Modèle

Soit un **simulateur physico-numérique** \mathcal{M} associant une grandeur d'intérêt réelle y à un vecteur $\mathbf{x}=(x_1,\ldots,x_d)$ composé de d paramètres d'entrée réels :

$$\begin{array}{cccc} \mathcal{M} & : & \mathcal{D}_{\mathbf{x}} \subset \mathbb{R}^d & \to & \mathbb{R} \\ & \mathbf{x} & \mapsto & \mathcal{M}(\mathbf{x}) =: y \end{array}$$

Paramètres d'entrée

- x conditions aux limites;
- x conditions initiales;
- paramètres du modèle physique;
- paramètres du simulateur numérique;
- **X** ...

Exposé du problème

Notre fil rouge : un code de calcul "jouet".

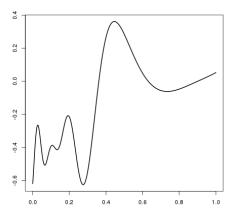


FIGURE: $\mathcal{M} =$ fonction de Xiong et al. (2007)

Exposé du problème Que faire avec un code de calcul?

Quelques exemples d'utilisation de ${\mathcal M}$

- Modélisation du phénomène physique pour une valeur de x
- X Optimisation du type :

```
 \begin{array}{l} \checkmark \  \, \mathbf{x}^* \in \underset{\mathbf{x} \in \mathcal{D}_{\mathbf{x}}}{\operatorname{argmax}} \  \, \mathcal{M}(\mathbf{x}) \\ \checkmark \  \, \mathbf{x}^* \in \underset{\mathbf{x} \in \mathcal{D}_{\mathbf{x}}}{\operatorname{argmin}} \  \, \mathbb{P}\left(\mathcal{M}(\mathbf{X}) > \mathbf{x}\right) \leq \alpha \end{array}
```

- X Gestion des incertitudes du type quantification, propagation et analyse de sensibilité
- X Problème inverse, du type assimilation de données
- X ...

Exposé du problème

Code de calcul : un point \oplus versus un point \ominus

Un code de calcul à la fois précis et lent

Le simulateur physico-numérique \mathcal{M} est :

- X précis par rapport aux observations du phénomène physique et aux théories sous-jacentes;
- X lent en temps d'exécution, du fait principalement de la résolution de systèmes d'équations non linéaires.

Des applications sollicitant fortement \mathcal{M}

- X Optimisation : non linéarités, grande dimension, algorithmes coûteux, ...
- X Incertitude : échantillonnage pour l'estimation des moments, quantiles, extrema...
- X Problème inverse : échantillonnage pour l'estimation des matrices de corrélation de l'EnKE...

Exposé du problème

Mener des applications gourmandes en calcul malgré le coût de ${\mathcal M}$

Le problème : le coût

Comment mener des applications requérant un grand nombre d'appels au simulateur physico-numérique ${\cal M}$ sachant que celui est coûteux en temps exécution?

Une réponse : le modèle de substitution

On remplace ${\mathcal M}$ par un **simulateur numérique** $\widehat{{\mathcal M}}:={\mathcal M}_{\hat{ heta}_{\hat{ heta}}}$ où

 $m{\mathsf{X}}$ le vecteur des **hyperparamètres** $\hat{ heta}_{\hat{p}}$ et

 \mathbf{X} l'espace de définition $\Theta_{\hat{p}} \in \{\Theta_p, \ p \in \mathcal{P}\}$ de $\hat{\theta}_{\hat{p}}$ sont calibrés via une base d'**apprentissage** $(\mathbf{x}^{(i)}, \mathcal{M}(\mathbf{x}^{(i)}))_{1 \leq i \leq N}$ de sorte à approcher : $\underset{\theta \in \Theta_p, \ p \in \mathcal{P}}{\operatorname{argmin}} \int_{\mathcal{D}_{\mathbf{x}}} (\mathcal{M}(\mathbf{x}) - \mathcal{M}_{\theta_p}(\mathbf{x}))^2 d\mathbf{x}$.

On résoud un **problème de régression non linéaire**.

Exposé du problème Plusieurs familles de métamodèles possibles

On veut remplacer le simulateur physico-numérique \mathcal{M} par le simulateur numérique $\widehat{\mathcal{M}}:=\mathcal{M}_{\widehat{\theta}_{\widehat{p}}}$ appelé métamodèle, **modèle de substitution**, modèle réduit ou encore modèle de régression.

Quelle famille de métamodèles choisir?

- x modèle linéaire généralisé;
- X réseau de neurones artificiels;
- X développement en polynômes de chaos;
- modèle de krigeage;
- X régressions par machines à vecteurs de support;
- X ...

Choix retenu dans cette présentation

Modèle de krigeage, a.k.a. modèle par processus gaussien.



- → De la nécessité d'un modèle de substitution
 - → Modèle physico-numérique
 - Diverses applications d'un simulateur physico-numérique
 - Le coût calculatoire de ces simulateurs
 - Remplacer un simulateur par un métamodèle
- + Présentation du métamodèle "krigeage"
 - ➤ Cadre bayésien : un métamodèle par processus gaussien
 - Cadre géostatistique : le meilleur prédicteur linéaire sans biais
 - Cadre fonctionnel : problème de régularisation dans un RKHS
- ♣ Pour aller plus loin...
 - ⇒ Simuler une instance de processus gaussien
 - Analogies entre le krigeage et le filtre de Kalman



Un métamodèle par processus gaussien

Voir $\mathcal M$ comme la réalisation d'un processus aléatoire.

HYPOTHÈSE : \mathcal{M} est l'instance d'un processus gaussien. QUESTION : Lequel ? INDICES : $\mathcal{M}\left(\mathbf{x}^{(1)}\right),\ldots,\mathcal{M}\left(\mathbf{x}^{(N)}\right)$.

Processus stochastique (ou aléatoire)

Soit:

- **X** un espace de probabilité (Ω, \mathcal{F}, P) ;
- **X** un ensemble d'indexation $\mathcal{D}_{\mathbf{x}}$, souvent appelé ensemble des indices et souvent $\mathcal{D}_{\mathbf{x}} \subset \mathbb{R}, \mathbb{R}_+, \mathbb{R}^p, \mathbb{N}, \ldots$;
- **X** un espace métrique (S, Σ) ;

Un processus stochastique à valeurs dans S est une famille de variables aléatoires définies sur l'espace de probabilité (Ω, \mathcal{F}, P) indexée par $\mathcal{D}_{\mathbf{x}}$ et à valeurs dans S. On le note $\{\mathcal{M}_{\mathbf{x}}\}_{\mathbf{x}\in\mathcal{D}_{\mathbf{x}}}$. On note alors $\mathcal{M}_{\mathbf{x}}(\omega)$ la valeur de $\mathcal{M}_{\mathbf{x}}$ en l'événement $\omega\in\Omega$ et S est appelé l'espace d'état.



Un métamodèle par processus gaussien

Voir \mathcal{M} comme la réalisation d'un processus gaussien.

Processus gaussien

Un processus aléatoire à valeurs dans l'espace d'état S et indicé par $\mathcal{D}_{\mathbf{x}}$ est gaussien si et seulement si pour tout ensemble fini d'indices $(\mathbf{x}^{(1)},\ldots,\mathbf{x}^{(k)})\in\mathcal{D}^k_{\mathbf{x}}$, le vecteur $\mathcal{M}_{\mathbf{v}^{(1:k)}}=(\mathcal{M}_{\mathbf{v}^{(1)}},\ldots,\mathcal{M}_{\mathbf{v}^{(k)}})$ suit une loi normale multivariée.

Exemple

On considère le processus gaussien $\{\mathcal{M}_{\mathbf{x}}\}_{\mathbf{x}\in\mathcal{D}_{\mathbf{x}}}$:

X d'espérance : $\forall \mathbf{x} \in \mathcal{D}_{\mathbf{x}}, \ \mathbb{E}[\mathcal{M}_{\mathbf{x}}] = \mu(\mathbf{x})$;

X de covariance : $\forall (\mathbf{x}, \mathbf{x}') \in \mathcal{D}_{\mathbf{x}}^2$, $\operatorname{cov} [\mathcal{M}_{\mathbf{x}}, \mathcal{M}_{\mathbf{x}'}] = \sigma^2 r(\mathbf{x}, \mathbf{x}')$.

On a donc :

$$\forall \mathbf{x}^{(1:k)} \in \mathcal{D}^k_{\mathbf{x}}, \ \mathcal{M}_{\mathbf{x}^{(1:k)}} \sim \mathcal{N}\left(\mu(\mathbf{x}^{(1:k)}); \sigma^2\left(r(\mathbf{x}^{(i)}, \mathbf{x}^{(j)})\right)_{1 \leq i,j \leq k}\right).$$

Prédéfinition d'un processus gaussien a priori

On suppose que le modèle déterministe $\mathcal M$ est l'instance du processus gaussien $\mathcal M_x$ à valeurs dans $\mathbb R$ et indicé par $\mathcal D_x$, dont l'espérance est notée :

$$\mathbb{E}\left[\mathcal{M}_{\mathbf{x}}\right] := \mu(\mathbf{x})$$

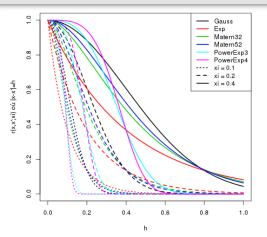
et la covariance :

$$cov[\mathcal{M}_{\mathbf{x}}, \mathcal{M}_{\mathbf{x}'}] = \sigma^2 r(\mathbf{x}, \mathbf{x}'; \boldsymbol{\xi})$$

Choix de la fonction μ

- \mathbf{X} krigeage simple : $\mu(\mathbf{x}) = \beta_0$ (fixé)
- **X** krigeage ordinaire : $\mu(\mathbf{x}) = \beta_0$
- **X** krigeage universel : $\mu(\mathbf{x}) = \sum_{b=0}^{B} \beta_i \phi_i(\mathbf{x})$
- où le plus souvent, B = d et $\phi_i(\mathbf{x}) = x_i$, avec $x_0 = 1$.

Choix du noyau de corrélation stationnaire $r(x, x'; \xi) := r(|x - x|; \xi)$



$$\begin{split} &\text{où } r_{\mathsf{gauss}}(h;\xi) = \exp\left(-\left(\frac{h}{\xi}\right)^2\right), \ r_{\mathsf{exp}}(h;\xi) = \exp\left(-\frac{h}{\xi}\right), \ r_{\mathsf{matern3.2}}(h;\xi) = \left(1 + \frac{\sqrt{3}i}{\xi}\right) \exp\left(-\frac{\sqrt{3}h}{\xi}\right), \\ &r_{\mathsf{matern5.2}}(h;\xi) = \left(1 + \frac{\sqrt{5}j}{\xi} + \frac{5h^2}{3\xi^2}\right) \exp\left(-\frac{\sqrt{5}h}{\xi}\right) \text{ exp} \left(r,\xi,\rho\right) = \exp\left(-\left(\frac{h}{\xi}\right)^\rho\right) \end{split}$$

Choix du noyau de corrélation stationnaire $r(\mathbf{x}, \mathbf{x}'; \boldsymbol{\xi})$

On utilise un noyau tensoriel :

$$r(\mathbf{x}, \mathbf{x}'; \boldsymbol{\xi}) = \prod_{i=1}^{d} r_i(x_i, x_i'; \xi_i) = \prod_{i=1}^{d} r_i(|x_i - x_i'|; \xi_i)$$

Différents noyaux pour différentes dépendances

On peut utiliser une famille de fonctions différente par noyau r_i afin de prendre en compte des profils de variation différents.

Présence de dépendances anisotropes

On peut utiliser un hyperparamètre ξ_i différent par noyau r_i afin de prendre en compte des vitesses de variation différentes.

Aller plus loin

- X noyaux tenant compte d'instationnarités, de symétries, de caractéristiques spatiales, ...
- x noyaux non tensoriels, ...



Intégration d'observations de ${\mathcal M}$

Ce modèle de krigeage *a priori* stipule uniquement que $\mathcal M$ est l'instance d'un processus gaussien $\mathcal M_{\mathbf x}$ prédéfini. Questions

- X Quelles valeurs pour les hyperparamètres de corrélation ξ ?
- **X** Quelles valeurs pour les hyperparamètres de tendance β ?
- X Quel choix d'instance de \mathcal{M}_x ?

Problème d'apprentissage

Pour chaque élément d'un plan d'expériences space-filling $(\mathbf{x}^{(1)},\dots,\mathbf{x}^{(N)})$, évaluer le simulateur \mathcal{M} . Puis à partir de la base d'apprentissage $(\mathbf{x}^{(i)},\mathcal{M}\left(\mathbf{x}^{(i)}\right))_{1\leq i\leq N}$, estimer les hyperparamètres $\boldsymbol{\xi}$ et $\boldsymbol{\beta}$ par maximum de vraisemblance ainsi qu'une instance du processus $\mathcal{M}_{\mathbf{x}}$ conditionné par ces hyperparamètres estimés et par la base d'apprentissage.

Le processus gaussien *a priori*Intégration d'observations de *M*

On recherche les estimateurs du maximum de vraisemblance (EMV) des paramètres (β, σ^2, ξ) en maximisant la log-vraisemblance :

$$-\frac{N}{2}\ln(2\pi)-\frac{N}{2}\ln(\sigma^2)-\frac{1}{2}\ln(\det(\mathbf{R}_{\boldsymbol{\xi}}))-\frac{1}{2\sigma^2}(\mathbf{Y}-\mathbf{\Phi}\boldsymbol{\beta})^T\mathbf{R}_{\boldsymbol{\xi}}^{-1}(\mathbf{Y}-\mathbf{\Phi}\boldsymbol{\beta})$$

avec

X la matrice des caractéristiques :
$$\mathbf{\Phi} = \left(\phi_j(\mathbf{x}^{(i)})\right)_{\substack{1 \leq i \leq N \ 0 < j < B}}$$
;

X la matrice des corrélations :
$$\mathbf{R}_{\xi} = (r(\mathbf{x}^{(i)}, \mathbf{x}^{(j)}; \xi)_{1 < i, i < N};$$

$$\mathbf{Y} = \left(\mathcal{M}\left(\mathbf{x}^{(1)}\right), \ldots, \mathcal{M}\left(\mathbf{x}^{(N)}\right)\right)^T$$
;

et

- X N étant la taille de la base d'apprentissage;
- \times σ^2 la variance du processus gaussien $\mathcal{M}_{\mathbf{x}}$.

Intégration d'observations de ${\mathcal M}$

EMV de β (par calcul analytique)

$$\hat{oldsymbol{eta}} = \left(oldsymbol{\Phi}^{\mathsf{T}} oldsymbol{\mathsf{R}}_{oldsymbol{\xi}} oldsymbol{\Phi}^{\mathsf{T}} oldsymbol{\mathsf{R}}_{oldsymbol{\xi}}^{-1} oldsymbol{\mathsf{Y}} \sim \mathcal{N}\left(oldsymbol{eta},
ight)$$

EMV de σ^2 (par calcul analytique)

$$\widehat{\sigma^2} = rac{1}{N-1} \left(\mathbf{Y} - \mathbf{\Phi} oldsymbol{eta}
ight)^T \mathbf{R}_{oldsymbol{\xi}}^{-1} \left(\mathbf{Y} - \mathbf{\Phi} oldsymbol{eta}
ight)$$

EMV de ξ (par optimisation numérique)

$$\hat{\boldsymbol{\xi}} = \operatorname*{argmin} \det \left(\mathbf{R}_{\boldsymbol{\xi}} \right)^{\frac{1}{N}} \widehat{\sigma^2} \ \ \text{où} \ \ \widehat{\sigma^2} \ \ \text{dépend de } \boldsymbol{\xi}$$



Intégration d'observations de ${\mathcal M}$

On remplace donc les hyperparamètres du processus gaussien $(\mathcal{M}_x)_{x\in\mathcal{D}_x}$ par leurs estimateurs du maximum de vraisemblance.

Définition du processus gaussien a priori

Le modèle déterministe \mathcal{M} est l'instance du processus gaussien $\left(\tilde{\mathcal{M}}_{\mathbf{x}}\right)_{\mathbf{x}\in\mathcal{D}_{\mathbf{x}}}$ à valeurs dans \mathbb{R} et indicé sur $\mathcal{D}_{\mathbf{x}}$, dont l'espérance est notée :

$$\mathbb{E}\left[ilde{\mathcal{M}}_{\mathsf{x}}
ight]=\hat{oldsymbol{eta}}^{\mathsf{T}}\phi(\mathsf{x})$$

et la covariance :

$$\text{cov}[\tilde{\mathcal{M}}_{\mathbf{x}},\tilde{\mathcal{M}}_{\mathbf{x}'}] = \widehat{\sigma^2} r(\mathbf{x},\mathbf{x}';\hat{\boldsymbol{\xi}})$$

Le processus gaussien *a priori*Intégration d'observations de \mathcal{M}

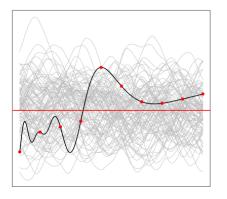


FIGURE: Processus gaussien a priori

Le processus gaussien a priori conditionné par les observations

Étant donné que :

$$\begin{pmatrix} \tilde{\mathcal{M}}_{\mathbf{x}} \\ \tilde{\mathcal{M}}_{\mathbf{x}^{(1:N)}} \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} \left(\boldsymbol{\phi} \left(\mathbf{x} \right) \right)^T \\ \boldsymbol{\Phi} \end{pmatrix} \hat{\boldsymbol{\beta}}, \hat{\sigma}^2 \begin{pmatrix} 1 & \mathbf{r}_{\hat{\xi}}^T(\mathbf{x}) \\ \mathbf{r}_{\hat{\xi}}(\mathbf{x}) & \mathbf{R}_{\hat{\xi}} \end{pmatrix} \right)$$

on obtient :

$$: \widetilde{\mathcal{M}}_{\mathbf{x}} | \widetilde{\mathcal{M}}_{\mathbf{x}^{(1:N)}}, \hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\xi}}, \widehat{\sigma^2} \sim \mathcal{N}\left(\widehat{\mathcal{M}}(\mathbf{x}); \widehat{\mathcal{E}}(\mathbf{x}, \mathbf{x}')\right)$$

dont les paramètres sont appelés :

X estimateur du modèle M:

$$\widehat{\mathcal{M}}(\mathbf{x}) = \phi(\mathbf{x})^T \hat{\boldsymbol{\beta}} + \mathbf{r}_{\hat{\boldsymbol{\xi}}}^T(\mathbf{x}) \mathbf{R}_{\hat{\boldsymbol{\xi}}}^{-1} \left(\mathbf{Y} - \mathbf{\Phi} \hat{\boldsymbol{\beta}} \right)$$

X estimateur de l'erreur du métamodèle $\widehat{\mathcal{M}}$:

$$\widehat{\mathcal{E}}(\mathbf{x}, \mathbf{x}') = \widehat{\sigma^2} \left(1 - \mathbf{r}_{\hat{\xi}}^T(\mathbf{x}) \mathbf{R}_{\hat{\xi}}^{-1} \mathbf{r}_{\hat{\xi}}(\mathbf{x}') \right)$$

Le processus gaussien a priori conditionné par les observations

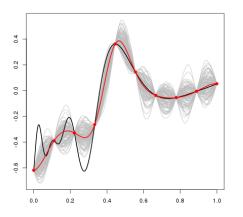


FIGURE: Processus gaussien a posteriori



Le processus gaussien a priori conditionné par les observations

Le métamodèle

$$\widehat{\mathcal{M}}(\mathbf{x}) = \hat{\boldsymbol{\beta}}^T \phi(\mathbf{x}) + \mathbf{r}_{\hat{\boldsymbol{\xi}}}^T(\mathbf{x}) \mathbf{R}_{\hat{\boldsymbol{\xi}}}^{-1} \left(\mathbf{Y} - \mathbf{\Phi} \hat{\boldsymbol{\beta}} \right)$$

se décompose en la somme de deux termes :

X une tendance :

$$\hat{oldsymbol{eta}}^{ au}\phi(\mathbf{x})$$

X une combinaison linéaire de noyaux :

$$\sum_{i=1}^{N} \gamma_{i} r\left(\mathbf{x}, \mathbf{x}^{(i)}; \hat{\boldsymbol{\xi}}\right)$$

où $\gamma = \mathsf{R}_{\hat{\pmb{\varepsilon}}}^{-1} \left(\mathsf{Y} - \mathsf{\Phi} \hat{\pmb{\beta}} \right)$ est le vecteur des pondérations.

Le processus gaussien a priori conditionné par les observations

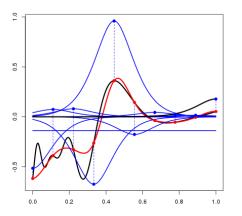


FIGURE: Décomposition du métamodèle en tendance et noyaux pondérés



Approche 100% bayésienne

Il peut être opportun de tenir compte de l'incertitude des estimateurs $\hat{\beta}$, $\hat{\xi}$ et $\widehat{\sigma^2}$, notamment lorsque le nombre N de simulations de $\mathcal M$ est faible. Pour cela, une approche 100% bayésienne est couramment utilisée :

- 1. On définit les lois a priori des hyperparamètres β , ξ et σ^2 .
- 2. On construit les lois a posteriori des hyperparamètres β , ξ et σ^2 .
- 3. On conditionne $\mathcal{M}_{\mathbf{x}}$ par rapport aux observations.

Ainsi, au lieu d'avoir :

$$\tilde{\mathcal{M}}_{\mathbf{x}}|\tilde{\mathcal{M}}_{\mathbf{x}^{(1:N)}},\hat{\boldsymbol{\beta}},\hat{\boldsymbol{\xi}},\widehat{\sigma^2}\sim\mathcal{N}\left(\widehat{\mathcal{M}}(\mathbf{x});\widehat{\mathcal{E}}(\mathbf{x},\mathbf{x}')\right) \hspace{0.5cm} [\mathsf{Bayesien partiel}]$$

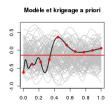
on a:

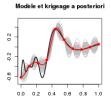
$$\mathcal{M}_{\mathbf{x}}|\mathcal{M}_{\mathbf{x}^{(1:N)}} \sim \mathcal{N}\left(\widehat{\mathcal{M}}(\mathbf{x}); \widehat{\mathcal{E}}(\mathbf{x}, \mathbf{x}')\right) \qquad [100\% \text{ bayesien}]$$

où cette fois-ci, les formulations de $\widehat{\mathcal{M}}(\mathbf{x})$ et $\widehat{\mathcal{E}}(\mathbf{x}, \mathbf{x}')$ sont plus complexes qu'au cas [Bayésien partiel], bien qu'analogues.

Résumé graphiques des étapes

1) Gaussian prior; 2) Observations; 3) Gaussian posterior; 4) Mean







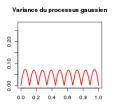


FIGURE: Du modèle ${\mathcal M}$ au métamodèle $\widehat{{\mathcal M}}$

- - → De la nécessité d'un modèle de substitution
 - → Modèle physico-numérique
 - Diverses applications d'un simulateur physico-numérique
 - Le coût calculatoire de ces simulateurs
 - Remplacer un simulateur par un métamodèle
 - ♣ Présentation du métamodèle "krigeage"
 - Cadre bayésien : un métamodèle par processus gaussien
 - Cadre géostatistique : le meilleur prédicteur linéaire sans biais
 - Cadre fonctionnel : problème de régularisation dans un RKHS
 - ♣ Pour aller plus loin...
 - → Simuler une instance de processus gaussien
 - Analogies entre le krigeage et le filtre de Kalman



Aux origines du "krigeage"

Le mot "krigeage" vient de l'ingénieur minier sud-africain Danie G. Krige (1919-2013), spécialiste de la géostatistique et professeur à l'université du Witwatersrand, en Afrique du Sud. Ses travaux empiriques pour évaluer les ressources minières ont été formalisés par le mathématicien et géologue français Georges Matheron (1930-2000).



D. G. Krige, A statistical approach to some basic mine valuation problems on the witwaters and, Technometrics, Vol. **52**, p. 119–139, 1951.



🦠 G. Matheron, Traité de géostatistique appliquée, Paris : Technip, Coll. Mémoires du Bureau de recherches géologiques et minières, 2 vol. (333, 171 p.), 1962-1963.

De la pratique géostatistique à sa théorie

"Pour tout mineur et géologue minier, la connaissance des teneurs et du tonage et l'appréciation des erreurs sur ces grandeurs est fondamentale. Or, jusqu'à présent, il n'existait pas d'approche scientifique correcte de ces problèmes.

La géostatistique, dont les principes sont résumés dans cet article, est la nouvelle science qui permet cette approche. L'auteur indique les méthodes statistiques antérieures et encore courantes et donne quelquesuns des résultats principaux de la géostatistique.

Toute évaluation de gisement et tout décision de mise en exploitation devrait être précédée d'une étude géostatistique permettant de limiter le risque d'une déconvenue ultérieure."



G. Matheron, Principles of geostatistics, *Economic Geology*, **Vol. 58**, p. 1246–1266, 1963.

Un formalisme probabiliste pour un phénomène déterministe

Au milieu du 20ème siècle, "[..] geostatistics started elaborating their own methods and their own mathematical formalism, which is nothing else than an abstract formulation and a systematization of secular mining experience. This formalisme has inherited from its statistical origin a language in which one still speaks of variance and covariance. including however in those notions new content. This similarity in vocabulary must not deceive. At the end of a protracted evolution, the geostatistical theory had to admit that it was facing, instead of random occurences, natural phenomena distributed in space. And, therefore, its methods are approximately these of mathematical physics and more specially those of harmonic analysis."

G. Matheron, Principles of geostatistics, *Economic Geology*, **Vol. 58**, p. 1246–1266, 1963.

La notion de variable régionalisée

"Thus a grade cannot in any way be assimilated to an aleatory variable. We speak of regionalized variables precisely in order to stress the spatial aspect of the phenomena. A regionalized variable is, sensu stricto, an actual function, taking a definite value in each point of the space."

Trois caractéristiques principales pour le physicien ou le géologue :

- X la localisation de la variation spatiale : les variations d'une variable régionalisée définie sur un support géométrique se situent dans un champ géographique;
- X la continuité de la variation spatiale : elle dépend de la variation plus ou moins importante de la quantité en deux points voisins; on parle parfois d'une continuité "en moyenne", voire d'un "effet de pépite" (nugget effect) si la continuité en moyenne n'est pas confirmée.
- X l'anisotropie : la variation spatiale peut différer selon les directions.



Le variogramme pour caractériser les variations spatiales

"Those different specific aspects of spatial distribution of regionalized variables – far apart from classical probability calculus – must compulsorily be taken into account by geostatistics. This is made possible owing to a simple mathematical tool: the variogram."

Variogramme (ou semi-variogramme)

Courbe représentant le degré de continuité d'une variable régionalisée (ou sa loi de dispersion) :

$$\gamma: \mathcal{V} \subset \mathbb{R}_{+}^{d} \to \mathbb{R}_{+}$$

$$h \mapsto \gamma(\mathbf{h}) = \frac{1}{2 \text{Vol}(V)} \int_{V} \left(\mathcal{M}(\mathbf{x} + \mathbf{h}) - \mathcal{M}(\mathbf{x}) \right)^{2} d\mathbf{x}$$

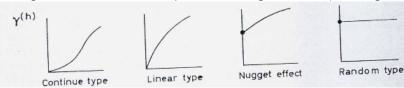
Naturellement, $\gamma(\mathbf{h})$ est d'autant plus grand que $\|\mathbf{h}\|$ est grand.

Le variogramme pour caractériser les variations spatiales

$$\gamma(\mathbf{h}) = \frac{1}{2\mathsf{Vol}(\mathcal{X})} \int_{\mathcal{X}} \left(\mathcal{M}(\mathbf{x} + \mathbf{h}) - \mathcal{M}(\mathbf{x}) \right)^2 d\mathbf{x}$$

Continuité de la variation spatiale

La régularité de f s'étudie à partir de la régularité de γ à l'origine.



De g. à d. : grande continuité, continuité "en moyenne", effet pépite et variable aléatoire. La géostatistique étudie le type de continuité qui est souvent un intermédiaire entre ces quatre.

Le variogramme pour caractériser les variations spatiales

Variogramme d'un processus stochastique stationnaire

$$\begin{split} \gamma(\mathbf{h}) &= \frac{1}{2} \mathbb{V} \left[\mathcal{M}_{\mathbf{x}+\mathbf{h}} - \mathcal{M}_{\mathbf{x}} \right] \\ &= \frac{1}{2} \mathbb{V} \left[\left(\mathcal{M}_{\mathbf{x}+\mathbf{h}} - \mathbb{E} \left[\mathcal{M}_{\mathbf{x}+\mathbf{h}} \right] \right) - \left(\mathcal{M}_{\mathbf{x}} + \mathbb{E} \left[\mathcal{M}_{\mathbf{x}+\mathbf{h}} \right] \right) \right] \\ &= \sigma^2 - \text{cov} \left(\mathcal{M}_{\mathbf{x}+\mathbf{h}}, \mathcal{M}_{\mathbf{x}} \right) \\ &= \sigma^2 - C(h) \end{split}$$



Le meilleur prédicteur linéaire sans biais

The Best Linear Unbiased Predictor (BLUP)

On considère un processus gaussien $\{\mathcal{M}_x\}_{x\in\mathcal{D}_x}$ d'espérance nulle et de structure de covariance $\sigma^2 r(\mathbf{x}, \mathbf{x}')$.

Estimateur linéaire

$$\widehat{\mathcal{M}}(\mathbf{x}) = a_0 + \sum_{i=1}^N a_i \mathcal{M}\left(\mathbf{x}^{(i)}\right) = a_0 + \mathbf{a}(\mathbf{x})^T \mathbf{Y}$$

Estimateur sans biais

$$\mathbb{E}\left[\mathcal{M}_{\boldsymbol{x}}\right] = \mathbb{E}\left[\widehat{\mathcal{M}}(\boldsymbol{x})\right] \Leftrightarrow \textit{a}_0 = 0$$

Estimateur de variance minimale

$$\underset{\widehat{\mathcal{M}}}{\operatorname{argmin}} \ \mathbb{V}\left[\widehat{\mathcal{M}}(\mathbf{x})\right] = \underset{\mathbf{a} \in \{f: \mathcal{D}_{\mathbf{x}} \mapsto \mathbb{R}^{N}\}}{\operatorname{argmin}} \ \mathbb{E}\left[\left(\mathbf{a}(\mathbf{x})^{T}\mathbf{Y} - \mathcal{M}_{\mathbf{x}}\right)^{2}\right]$$



Le meilleur prédicteur linéaire sans biais

The Best Linear Unbiased Predictor (BLUP)

On considère un processus gaussien $\{\mathcal{M}_x\}_{x\in\mathcal{D}_x}$ d'espérance nulle et de structure de covariance $\sigma^2 r(\mathbf{x}, \mathbf{x}')$.

Estimateur linéaire

$$\widehat{\mathcal{M}}(\mathbf{x}) = a_0 + \sum_{i=1}^N a_i \mathcal{M}\left(\mathbf{x}^{(i)}\right) = a_0 + \mathbf{a}(\mathbf{x})^T \mathbf{Y}$$

Estimateur sans biais

$$\mathbb{E}\left[\mathcal{M}_{\boldsymbol{x}}\right] = \mathbb{E}\left[\widehat{\mathcal{M}}(\boldsymbol{x})\right] \Leftrightarrow \textit{a}_0 = 0$$

Estimateur de variance minimale

$$\underset{\widehat{\mathcal{M}}}{\operatorname{argmin}} \, \mathbb{V}\left[\widehat{\mathcal{M}}(\mathbf{x})\right] \, = \underset{\mathbf{a} \in \{f: \mathcal{D}_{\mathbf{x}} \mapsto \mathbb{R}^N\}}{\operatorname{argmin}} \, \sigma^2 \left(1 + \mathbf{a}(\mathbf{x})^T \mathbf{R} \mathbf{a}(\mathbf{x}) - 2\mathbf{a}(\mathbf{x})^T \mathbf{r}(\mathbf{x})\right)$$



Le meilleur prédicteur linéaire sans biais

The Best Linear Unbiased Predictor (BLUP)

On considère un processus gaussien $\{\mathcal{M}_{\mathbf{x}}\}_{\mathbf{x}\in\mathcal{D}_{\mathbf{x}}}$ d'espérance nulle et de structure de covariance $\sigma^2 r(\mathbf{x}, \mathbf{x}')$.

Estimateur linéaire

$$\widehat{\mathcal{M}}(\mathbf{x}) = a_0 + \sum_{i=1}^N a_i \mathcal{M}\left(\mathbf{x}^{(i)}\right) = a_0 + \mathbf{a}(\mathbf{x})^T \mathbf{Y}$$

Estimateur sans biais

$$\mathbb{E}\left[\mathcal{M}_{\mathbf{x}}\right] = \mathbb{E}\left[\widehat{\mathcal{M}}(\mathbf{x})\right] \Leftrightarrow a_0 = 0$$

Estimateur de variance minimale

$$\underset{\mathbf{a} \in \{f: \mathcal{D}_{\mathbf{x}} \mapsto \mathbb{R}^N\}}{\operatorname{argmin}} \, \mathbb{V}\left[\widehat{\mathcal{M}}(\mathbf{x})\right] \, = \mathbf{r}(\mathbf{x})^T \mathbf{R}^{-1}$$

The Best Linear Unbiased Predictor (BLUP)

On considère un processus gaussien $\{\mathcal{M}_{\mathbf{x}}\}_{\mathbf{x}\in\mathcal{D}_{\mathbf{x}}}$ d'espérance nulle et de structure de covariance $\sigma^2 r(\mathbf{x}, \mathbf{x}')$.

Formulation du BLUP

$$\widehat{\mathcal{M}}(\mathbf{x}) = \mathbf{r}(\mathbf{x})^T \mathbf{R}^{-1} \mathbf{Y}$$

Erreur quadratique moyenne du BLUP

$$\mathsf{EQM}\left(\widehat{\mathcal{M}}(\mathbf{x})\right) = \sigma^2 \left(1 - \mathbf{r}(\mathbf{x})\mathbf{K}^{-1}\mathbf{r}(\mathbf{x})\right)$$

Lien avec l'approche bayésienne

On retrouve les équations de l'approche [Bayésien partiel]!

The Best Linear Unbiased Predictor (BLUP)

On considère un processus gaussien $\{\mathcal{M}_{\mathbf{x}}\}_{\mathbf{x}\in\mathcal{D}_{\mathbf{x}}}$ d'espérance $\Phi(\mathbf{x})^T \boldsymbol{\beta}$ et de structure de covariance $\sigma^2 r(\mathbf{x}, \mathbf{x}')$.

Estimateur linéaire, sans biais et de variance minimale

$$\left\{ \begin{array}{l} \underset{\boldsymbol{a} \in \{f: \mathcal{D}_{\boldsymbol{x}} \mapsto \mathbb{R}^N\}}{\text{argmin}} \; \mathbb{E}\left[\left(\boldsymbol{a}(\boldsymbol{x})^T\boldsymbol{Y} - \mathcal{M}_{\boldsymbol{x}}\right)^2\right] \\ \text{s.c.} \; \boldsymbol{\Phi}^T\boldsymbol{a}(\boldsymbol{x}) = \boldsymbol{\Phi}(\boldsymbol{x}) \end{array} \right.$$

ou de façon équivalente :

$$\begin{cases} \underset{\mathbf{a} \in \{f: \mathcal{D}_{\mathbf{x}} \mapsto \mathbb{R}^{N}\}}{\operatorname{argmin}} \sigma^{2} \left(1 + \mathbf{a}(\mathbf{x})^{T} \mathbf{R} \mathbf{a}(\mathbf{x}) - 2 \mathbf{a}(\mathbf{x})^{T} \mathbf{r}(\mathbf{x}) \right) \\ \text{s.c. } \mathbf{\Phi}^{T} \mathbf{a}(\mathbf{x}) = \Phi(\mathbf{x}) \end{cases}$$

The Best Linear Unbiased Predictor (BLUP)

On considère un processus gaussien $\{\mathcal{M}_{\mathbf{x}}\}_{\mathbf{x}\in\mathcal{D}_{\mathbf{x}}}$ d'espérance $\Phi(\mathbf{x})^T\beta$ et de structure de covariance $\sigma^2 r(\mathbf{x}, \mathbf{x}')$.

Minimisation via des multiplicateurs de Lagrange

$$\begin{aligned} \underset{(\mathbf{a}, \boldsymbol{\lambda}) \in \{f: \mathcal{D}_{\mathbf{x}} \mapsto \mathbb{R}^{N}\} \times \mathbb{R}^{B}}{\operatorname{argmin}} \ \sigma^{2} \left(1 + \mathbf{a}(\mathbf{x})^{T} \mathbf{R} \mathbf{a}(\mathbf{x}) - 2 \mathbf{a}(\mathbf{x})^{T} \mathbf{r}(\mathbf{x}) \right) \\ + 2 \sigma^{2} \boldsymbol{\lambda}^{T} \left(\boldsymbol{\Phi}^{T} \mathbf{a}(\mathbf{x}) - \boldsymbol{\Phi}(\mathbf{x}) \right) \end{aligned}$$

d'où:

$$\left\{ \begin{array}{l} \nabla_{\pmb{\lambda}} = 0 \\ \nabla_{\pmb{a}} = 0 \end{array} \right. \Leftrightarrow \left\{ \begin{array}{l} \pmb{\Phi}^{\mathsf{T}} \pmb{a}(\pmb{x}) = \Phi(\pmb{x}) \\ \pmb{\Phi} \pmb{\lambda} + \pmb{\mathsf{Ra}}(\pmb{x}) = \pmb{r}(\pmb{x}) \end{array} \right.$$

The Best Linear Unbiased Predictor (BLUP)

On considère un processus gaussien $\{\mathcal{M}_{\mathbf{x}}\}_{\mathbf{x}\in\mathcal{D}_{\mathbf{x}}}$ d'espérance $\Phi(\mathbf{x})^T \boldsymbol{\beta}$ et de structure de covariance $\sigma^2 r(\mathbf{x}, \mathbf{x}')$.

Minimisation via des multiplicateurs de Lagrange

$$\begin{aligned} \underset{(\mathbf{a}, \boldsymbol{\lambda}) \in \{f: \mathcal{D}_{\mathbf{x}} \mapsto \mathbb{R}^N\} \times \mathbb{R}^B}{\operatorname{argmin}} & \sigma^2 \left(1 + \mathbf{a}(\mathbf{x})^T \mathbf{R} \mathbf{a}(\mathbf{x}) - 2\mathbf{a}(\mathbf{x})^T \mathbf{r}(\mathbf{x}) \right) \\ & + 2\sigma^2 \boldsymbol{\lambda}^T \left(\boldsymbol{\Phi}^T \mathbf{a}(\mathbf{x}) - \boldsymbol{\Phi}(\mathbf{x}) \right) \\ \operatorname{d'où} : & \begin{pmatrix} \boldsymbol{\lambda} \\ \mathbf{a} \end{pmatrix} = \begin{pmatrix} \mathbf{0} & \boldsymbol{\Phi}^T \\ \boldsymbol{\Phi} & \mathbf{R} \end{pmatrix}^{-1} \begin{pmatrix} \boldsymbol{\Phi}(\mathbf{x}) \\ \mathbf{r}(\mathbf{x}) \end{pmatrix} \\ & = \begin{pmatrix} -\mathbf{Q} & \mathbf{Q} \boldsymbol{\Phi}^T \mathbf{R}^{-1} \\ \mathbf{R}^{-1} \boldsymbol{\Phi} \mathbf{Q} & \mathbf{R}^{-1} - \mathbf{R}^{-1} \boldsymbol{\Phi} \mathbf{Q} \boldsymbol{\Phi}^T \mathbf{R}^{-1} \end{pmatrix} \begin{pmatrix} \boldsymbol{\Phi}(\mathbf{x}) \\ \mathbf{r}(\mathbf{x}) \end{pmatrix} \end{aligned}$$

avec
$$\mathbf{Q} = (\mathbf{\Phi}^T \mathbf{R}^{-1} \mathbf{\Phi})^{-1}$$
.



The Best Linear Unbiased Predictor (BLUP)

On considère un processus gaussien $\{\mathcal{M}_{\mathbf{x}}\}_{\mathbf{x}\in\mathcal{D}_{\mathbf{x}}}$ d'espérance $\Phi(\mathbf{x})^T\beta$ et de structure de covariance $\sigma^2 r(\mathbf{x}, \mathbf{x}')$.

Minimisation via des multiplicateurs de Lagrange

$$\begin{aligned} \underset{(\mathbf{a}, \boldsymbol{\lambda}) \in \{f: \mathcal{D}_{\mathbf{x}} \mapsto \mathbb{R}^{N}\} \times \mathbb{R}^{B}}{\operatorname{argmin}} & \sigma^{2} \left(1 + \mathbf{a}(\mathbf{x})^{T} \mathbf{R} \mathbf{a}(\mathbf{x}) - 2 \mathbf{a}(\mathbf{x})^{T} \mathbf{r}(\mathbf{x}) \right) \\ & + 2 \sigma^{2} \boldsymbol{\lambda}^{T} \left(\boldsymbol{\Phi}^{T} \mathbf{a}(\mathbf{x}) - \boldsymbol{\Phi}(\mathbf{x}) \right) \end{aligned}$$

d'où:

$$\mathbf{a}(\mathbf{x}) = \mathbf{R}^{-1} \mathbf{\Phi} \mathbf{Q} \Phi(\mathbf{x}) + \left(\mathbf{R}^{-1} - \mathbf{R}^{-1} \mathbf{\Phi} \mathbf{Q} \mathbf{\Phi}^T \mathbf{R}^{-1}\right) \mathbf{r}(\mathbf{x})$$

avec
$$\mathbf{Q} = (\mathbf{\Phi}^T \mathbf{R}^{-1} \mathbf{\Phi})^{-1}$$
.

Le r

Le meilleur prédicteur linéaire sans biais

The Best Linear Unbiased Predictor (BLUP)

On considère un processus gaussien $\{\mathcal{M}_{\mathbf{x}}\}_{\mathbf{x}\in\mathcal{D}_{\mathbf{x}}}$ d'espérance $\Phi(\mathbf{x})^T\beta$ et de structure de covariance $\sigma^2 r(\mathbf{x}, \mathbf{x}')$.

Formulation du BLUP

$$\widehat{\mathcal{M}}(\mathbf{x}) = \Phi(\mathbf{x})^T \hat{\boldsymbol{\beta}} + \mathbf{r}(\mathbf{x})^T \mathbf{R}^{-1} \left(\mathbf{Y} - \mathbf{\Phi} \hat{\boldsymbol{\beta}} \right)$$

où
$$\hat{\boldsymbol{\beta}} = \left(\mathbf{\Phi}^T \mathbf{R}^{-1} \mathbf{\Phi}\right)^{-1} \mathbf{\Phi}^T \mathbf{R}^{-1} \mathbf{Y}.$$

Erreur quadratique moyenne du BLUP

$$\mathsf{EQM}\left(\widehat{\mathcal{M}}(\mathbf{x})\right) = \sigma^2 \left(1 - \mathbf{r}(\mathbf{x})\mathbf{K}^{-1}\mathbf{r}(\mathbf{x})\right)$$

$$+ \sigma^2 \left(\Phi(\mathbf{x})^T - \mathbf{r}(\mathbf{x})^T \mathbf{R}^{-1} \mathbf{\Phi} \right) \left(\mathbf{\Phi}^T \mathbf{R}^{-1} \mathbf{\Phi} \right)^{-1} \left(\Phi(\mathbf{x})^T - \mathbf{r}(\mathbf{x})^T \mathbf{R}^{-1} \mathbf{\Phi} \right)^T$$

Lien avec l'approche bayésienne

On retrouve les formes d'équations de l'approche [100% bayésien]!

- - → De la nécessité d'un modèle de substitution
 - → Modèle physico-numérique
 - Diverses applications d'un simulateur physico-numérique
 - Le coût calculatoire de ces simulateurs
 - Remplacer un simulateur par un métamodèle

+ Présentation du métamodèle "krigeage"

- Cadre bayésien : un métamodèle par processus gaussien
- Cadre géostatistique : le meilleur prédicteur linéaire sans biais
- Cadre fonctionnel : problème de régularisation dans un RKHS

♣ Pour aller plus loin...

- → Simuler une instance de processus gaussien
- Analogies entre le krigeage et le filtre de Kalman



* Un problème de régularisation dans un RKHSÀ RÉDIGER.....À RÉDIGER.....À RÉDIGER......À RÉDIGER......



- ♣ De la nécessité d'un modèle de substitution
 - → Modèle physico-numérique
 - Diverses applications d'un simulateur physico-numérique
 - Le coût calculatoire de ces simulateurs
 - Remplacer un simulateur par un métamodèle
- + Présentation du métamodèle "krigeage"
 - Cadre bayésien : un métamodèle par processus gaussien
 - Cadre géostatistique : le meilleur prédicteur linéaire sans biais
 - Cadre fonctionnel : problème de régularisation dans un RKHS
- ♣ Pour aller plus loin...
 - ➤ Simuler une instance de processus gaussien
 - Analogies entre le krigeage et le filtre de Kalman



Pour aller plus loin... Simuler une instance de processus gaussien

Selon le théorème de Karhunen–Loève, le processus gaussien $(\mathcal{M}_x)_{x\in\mathcal{D}_x}$ de carré intégrable et de fonction de covariance continue peut s'écrire :

$$\mathcal{M}_{\mathbf{x}} = \mu(\mathbf{x}) + \sum_{i \geq 1} \sqrt{\lambda_i} v_i(\mathbf{x}) \varepsilon_i$$

 v_i et λ_i sont les $i^{\text{ème}}$ fonction propre et vecteur propre de la fonction de covariance de $(\mathcal{M}_x)_{x \in \mathcal{D}_x}$, v_i et v_j étant orthonormales; ils sont solutions de l'équation de Fredholm:

$$\int_{\mathcal{D}_{\mathbf{x}}} \operatorname{cov}\left(\mathcal{M}_{\mathbf{x}}, \mathcal{M}_{\mathbf{x}'}\right) v_{i}(\mathbf{x}') d\mathbf{x}' = \lambda_{i} v_{i}(\mathbf{x})$$

X $(\varepsilon_i)_{i\geq 1}$ sont des variables aléatoires gaussiennes standards et indépendantes.

Pour aller plus loin...

Simuler une instance de processus gaussien

Le processus gaussien $(\mathcal{M}_x)_{x\in\mathcal{D}_x}$ peut s'écrire :

$$\mathcal{M}_{\mathbf{x}} = \mu(\mathbf{x}) + \sum_{i \geq 1} \sqrt{\lambda_i} v_i(\mathbf{x}) \varepsilon_i$$

et sa variance totale vaut $\int_{\mathcal{D}_{\mathbf{x}}} \mathbb{V}(\mathcal{M}_{\mathbf{x}}) d\mathbf{x} = \sum_{i>1} \lambda_i$. Il peut être approché par la troncature :

$$\mathcal{M}_{\mathbf{x},P} = \mu(\mathbf{x}) + \sum_{i=1}^{P} \sqrt{\lambda_i} v_i(\mathbf{x}) \varepsilon_i$$
$$= \mu(\mathbf{x}) + \tilde{\mathbf{v}}(\mathbf{x})^T \varepsilon \text{ avec } \varepsilon \sim \mathcal{N}(0, I_P)$$

où P est le plus petit entier vérifiant $100 \frac{\sum_{i=1}^{p} \lambda_i}{\sum_{i\geq 1} \lambda_i} > \alpha$, celui à partir duquel la part de variance totale expliquée dépasse α %.

- - → De la nécessité d'un modèle de substitution
 - → Modèle physico-numérique
 - Diverses applications d'un simulateur physico-numérique
 - Le coût calculatoire de ces simulateurs
 - Remplacer un simulateur par un métamodèle
 - + Présentation du métamodèle "krigeage"
 - Cadre bayésien : un métamodèle par processus gaussien
 - Cadre géostatistique : le meilleur prédicteur linéaire sans biais
 - Cadre fonctionnel : problème de régularisation dans un RKHS
 - ♣ Pour aller plus loin...
 - ⇒ Simuler une instance de processus gaussien
 - → Analogies entre le krigeage et le filtre de Kalman





Pour aller plus loin...

Analogies entre les équations du krigeage et du filtre de Kalman

KRIGEAGE

Modèle d'observation

$$Y_{x} = \Phi(x)^{T} \boldsymbol{\beta} + \boldsymbol{\varepsilon}_{x}$$

$$(\boldsymbol{\varepsilon}_{\mathbf{x}})_{\mathbf{x} \in \mathcal{D}_{\mathbf{x}}} \sim \mathcal{PG}\left[0, \sigma^{2} r(\mathbf{x}, \cdot)\right]$$

.

Vecteur gaussien de la réponse en x et $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(N)}$

$$\begin{pmatrix} \mathbf{Y}(\mathbf{x}) \\ \mathbf{Y} \end{pmatrix} \sim \mathcal{N} \left[\begin{pmatrix} \mathbf{\Phi}(\mathbf{x})^T \boldsymbol{\beta} \\ \mathbf{\Phi} \boldsymbol{\beta} \end{pmatrix}, \sigma^2 \begin{pmatrix} 1 & \mathbf{r}(\mathbf{x})^T \\ \mathbf{r}(\mathbf{x}) & \mathbf{R} \end{pmatrix} \right]$$

Estimation de la réponse accompagnée d'une mesure d'erreur

$$\hat{\mathbf{Y}}(\mathbf{x}) = \mathbb{E}[\mathbf{Y}(\mathbf{x})|\mathbf{Y} = \mathbf{y}] =$$

$$\Phi(\mathbf{x})\boldsymbol{\beta} + \mathbf{r}(\mathbf{x})^{T}\mathbf{R}^{-1}(\mathbf{y} - \mathbf{\Phi}\boldsymbol{\beta})$$

$$\oplus$$

$$\sigma^{2}\hat{\mathbf{r}}(\mathbf{x}, \mathbf{x}') = \text{cov}[\mathbf{Y}(\mathbf{x}), \mathbf{Y}(\mathbf{x}')|\mathbf{Y} = \mathbf{y}] =$$

$$\sigma^{2}\left(1 - \mathbf{r}(\mathbf{x})^{T}\mathbf{R}^{-1}\mathbf{r}(\mathbf{x})\right)$$

FILTRE DE KALMAN

Modèle d'observation

$$\mathbf{Y}^o = \mathbf{G}\mathbf{X}^b + \boldsymbol{\varepsilon}$$

$$\mathbf{X}^b \sim \mathcal{N}\left[\mathbf{X}^t, \mathbf{B}\right], \; \boldsymbol{arepsilon} \perp \mathbf{X}^b \; \mathrm{et} \; \boldsymbol{arepsilon} \sim \mathcal{N}\left[\mathbf{0}, \boldsymbol{R}
ight]$$

↓

Vecteur gaussien de l'observation et de l'analyse

$$\begin{pmatrix} \mathbf{X}^{a} \\ \mathbf{Y}^{o} \end{pmatrix} \sim \mathcal{N} \left[\begin{pmatrix} \mathbf{x}^{b} \\ \mathbf{G} \mathbf{x}^{b} \end{pmatrix}, \begin{pmatrix} \mathbf{B} & \mathbf{B} \mathbf{G}^{T} \\ \mathbf{G} \mathbf{B} & \mathbf{G} \mathbf{B} \mathbf{G}^{T} + \mathbf{R} \end{pmatrix} \right]$$

-

Estimation de l'analyse accompagnée d'une mesure d'erreur

$$\mathbf{x}^a = \mathbb{E}[\mathbf{X}^a | \mathbf{Y}^o = \mathbf{y}^o] =$$

$$\mathbf{x}^b + \mathbf{B}\mathbf{G}^T \left(\mathbf{G}\mathbf{B}\mathbf{G}^T + \mathbf{R}\right)^{-1} \left(\mathbf{y}^o - \mathbf{G}\mathbf{x}^b\right)$$

$$\mathbf{B}^{a} = \operatorname{cov}[\mathbf{X}^{a}|\mathbf{Y}^{o} = \mathbf{y}^{o}] =$$

$$\mathbf{B} - \mathbf{B}\mathbf{G}^T \left(\mathbf{G}\mathbf{B}\mathbf{G}^T + \mathbf{R}\right)^{-1} \mathbf{G}\mathbf{B}$$



Analogies entre le krigeage et le filtre de Kalman

KRIGEAGE

On dispose:

On dispose:

- d'un modèle d'observation incertain

On souhaite obtenir:

- X un modèle d'observation moins incertain
- pour en utiliser :
 - ✓ l'espérance comme estimateur
 - ✓ et la variance comme erreur.

On s'aide pour cela:

de plusieurs observations.

X d'un modèle d'observation incertain

FILTRE DE KALMAN

X avec des paramètres incertaines.

On souhaite obtenir:

- X des paramètres moins incertains
- pour en utiliser :
 - ✓ l'espérance comme estimateur
 - ✓ et la variance comme erreur.

On s'aide pour cela:

X d'une observation.

PROBLÈME DIRECT

PROBLÈME INVERSE

INFÉRENCE BAYÉSIENNE



Pour aller plus loin...

Quels éléments souhaitez-vous voir ajoutés à cette présentation?

