

9- Géostatistique

9.1 Introduction	1
9.2. Le variogramme.....	1
9.2.1 Hypothèses de base et définition:	2
9.2.2 Estimation du variogramme.....	5
9.2.3 Modélisation	8
9.2.4 Remarques concernant le calcul de variogrammes et l'ajustement de modèles	11
9.3 Variance d'estimation	12
9.4 Krigeage.....	14
9.4.1 Krigeage ordinaire	15
9.4.2 Quelques cas très simples de krigeage ordinaire	17
9.4.4 Propriétés du krigeage	17
9.4.5 Pratique du krigeage	24
9.4.6 Validation croisée	25
9.4.7 Exemple numérique de krigeage	29

9.1 Introduction

Lorsqu'on mesure une caractéristique en un point, on peut considérer la valeur obtenue comme la réalisation d'une variable aléatoire en ce point. Il en est de même pour tous les points d'un site donné. On a donc un grand nombre (ou une infinité) de v.a. représentant conjointement un site. La géostatistique adopte ce point de vue et considère la distribution conjointe de toutes ces v.a.

Il y a deux étapes principales dans une étude géostatistique :

- identification des caractéristiques des v.a.. L'outil principal utilisé est le variogramme (section 9.2)
- utilisation de ces caractéristiques et des valeurs connues pour l'estimation optimale aux points (ou sur des volumes) non mesurés. La méthode utilisée est le krigeage (section 9.3).

9.2. Le variogramme

Idée fondamentale: La nature n'est pas entièrement "imprévisible". Deux observations situées l'une près de l'autre devraient, en moyenne, se ressembler davantage (i.e. être plus corrélées) que deux observations éloignées.

Ex. Soit trois localisations x_0 , x_1 et x_2 , que l'on déplace par translation sur le site. On mesure la valeur en chacun de ces points.

x_1 x_0 x_2

La valeur au point x_1 devrait ressembler plus (en moyenne) à celle observée en x_0 qu'à celle en x_2 .

Supposons maintenant que l'on connaît uniquement la valeur aux points x_1 et x_2 . On a peut-être intérêt à utiliser l'information contenue en x_1 et x_2 pour fournir une bonne estimation de la valeur en x_0 .

Notion de "continuité" de la minéralisation : Implicitement toutes les méthodes d'estimation, d'interpolation et de cartographie reposent sur ce concept. En géostatistique, on cherche à quantifier cette continuité préalablement à l'estimation pour pouvoir l'utiliser lors de l'estimation.

Soit deux points x et $x+h$ séparés d'une distance h .

$$x \longleftarrow \text{-----} \longrightarrow x+h$$

La teneur en x est une variable aléatoire $Z(x)$, la teneur en $x + h$ aussi, $Z(x+h)$.

La différence entre les valeurs prises par ces deux v.a. est $Z(x) - Z(x+h)$. C'est également une v.a. dont on peut calculer la variance. Cette variance devrait être plus petite lorsque les points sont rapprochés (les valeurs se ressemblent plus en moyenne) et plus grande lorsque les points sont éloignés. On appelle variogramme la demi-variance de cette différence, i.e. $\gamma(x,x+h) = 0.5 * \text{Var}(Z(x) - Z(x+h))$

Si l'on considère n localisations différentes x_1, x_2, \dots, x_n , la meilleure description que l'on puisse faire des n variables aléatoires $Z(x_1), Z(x_2), \dots, Z(x_n)$ est d'établir la fonction de distribution conjointe (multivariable). Clairement, ceci n'est pas possible puisqu'on ne peut disposer généralement que d'une seule observation à chacun de ces n points. On pourrait formuler une hypothèse très forte du genre: le vecteur des v.a. suit une loi multinormale de moyennes et variances-covariances spécifiées. Ceci serait beaucoup trop restrictif.

La géostatistique a des visées plus modestes. On veut estimer des paramètres statistiques à partir des données et non imposer un modèle à priori qui aurait toutes les chances de s'avérer inadéquat. Les paramètres que l'on cherchera à estimer ne sont pas la fonction de distribution conjointe, ni même la fonction de distribution bivariable (i.e. les v.a. considérées deux à deux) mais simplement les deux premiers moments (moyenne, variance, covariance) des v.a. prises deux à deux. Même réduit à cela, on ne dispose toujours que d'une seule paire d'observations situées précisément aux points x et $x+h$. On ne peut donc estimer les paramètres statistiques sans formuler certaines hypothèses. Ces hypothèses ont uniquement pour but de permettre l'estimation des paramètres statistiques de notre modèle à partir des données. On les appelle hypothèses de stationnarité du second ordre; elles visent essentiellement à "détacher" les deux premiers moments de localisations précises en permettant des translations des emplacements x et $x+h$. La covariance (et le variogramme) deviennent donc des fonctions dépendant uniquement de la distance séparant les points d'observation et non plus de leur localisation exacte.

9.2.1 Hypothèses de base et définition:

Bref, on suppose que:

- i. L'espérance mathématique ne dépend pas de x ,
i.e. $E[Z(x)] = m$
ou
L'espérance des écarts est zéro
i.e. $E[Z(x) - Z(x+h)] = 0$
 - ii. La covariance entre $Z(x)$ et $Z(x+h)$ ne dépend que de h
i.e. $\text{Cov}(Z(x), Z(x+h)) = C(h)$; stationnarité du second ordre, $C(h)$ est appelé fonction de covariance ou covariogramme
ou
Le variogramme $\gamma(h)$ ne dépend pas de la localisation x , seulement de h (soit en module, soit en module et en direction).
- i.e. $1/2 \text{Var}(Z(x) - Z(x+h)) = \gamma(h)$; hypothèse intrinsèque (cette dernière hypothèse est légèrement moins restrictive que la stationnarité du second ordre)

Évidemment, ces hypothèses supposent une certaine régularité, une certaine homogénéité du gisement étudié. Si on peut reconnaître des zones très différentes géologiquement, on a habituellement intérêt à les traiter séparément.

La fonction la plus utilisée en géostatistique pour décrire la continuité de la minéralisation est le variogramme, et ce surtout parce qu'elle est plus simple à estimer que la covariance (qui demande l'estimation préalable de l'espérance mathématique), mais également parce qu'elle permet d'accommoder les situations où $\text{Var}(Z(x))$ n'est pas définie.

Le variogramme théorique est défini comme:

$$\gamma(h) = \frac{1}{2} \text{Var}[Z(x) - Z(x+h)] = \frac{1}{2} E[(Z(x) - Z(x+h))^2]$$

où x est le vecteur de coordonnées (1, 2 ou 3 coordonnées selon le cas)
 h est le vecteur distance .

Cette fonction, habituellement croissante en fonction de h , synthétise beaucoup d'informations concernant le comportement conjoint des variables aléatoires et concernant "la continuité" de la minéralisation. Ainsi, pour les modèles de variogramme montrant un seuil, on a :

- i. Portée a : Distance où deux observations ne se ressemblent plus du tout en moyenne, elles ne sont plus liées (covariance nulle) linéairement. À cette distance, la valeur du variogramme correspond à la variance de la variable aléatoire.
- ii. Palier $\sigma^2 = C_0 + C$: Variance de la v.a. ($\text{Var}(Z(x))$)
 Écarts les plus grands, en moyenne entre deux v.a.
- iii. Effet de pépite : C_0 : Variation à très courte échelle, erreurs de localisation, erreurs d'analyse et précision analytique.

Ex. Une carotte fendue en deux et dont chaque partie est analysée séparément ne fournira pas exactement les mêmes valeurs pour les deux moitiés. Un même paquet de poudre, séparé en deux parties pour analyse ne donnera pas exactement la même teneur.

Notes : i. Lorsque $h = 0$ on a

$$\gamma(0) = \frac{1}{2} \text{Var}(Z(x) - Z(x)) = 0 \text{ et non } C_0$$

par contre,

$$\lim_{\varepsilon \rightarrow 0^+} \gamma(\varepsilon) = C_0$$

i.e. on a une discontinuité à l'origine du variogramme.

- ii. Parfois les variogrammes ne montrent pas de palier (dans ce cas, la covariance et la variance n'existent pas).
- iii. Lorsque les variogrammes montrent un palier alors on peut facilement établir le lien entre la valeur du variogramme pour la distance h et la covariance pour deux observations séparées de h .

$$\begin{aligned}
 \gamma(h) &= \frac{1}{2} \text{Var} (Z(x) - Z(x+h)) \\
 &= \frac{1}{2} [\text{Var} (Z(x)) + \text{Var} (Z(x+h)) - 2 \text{Cov} (Z(x), Z(x+h))] \\
 &= \sigma^2 - \text{Cov} (Z(x), Z(x+h)) = \sigma^2 - C(h)
 \end{aligned}$$

donc,

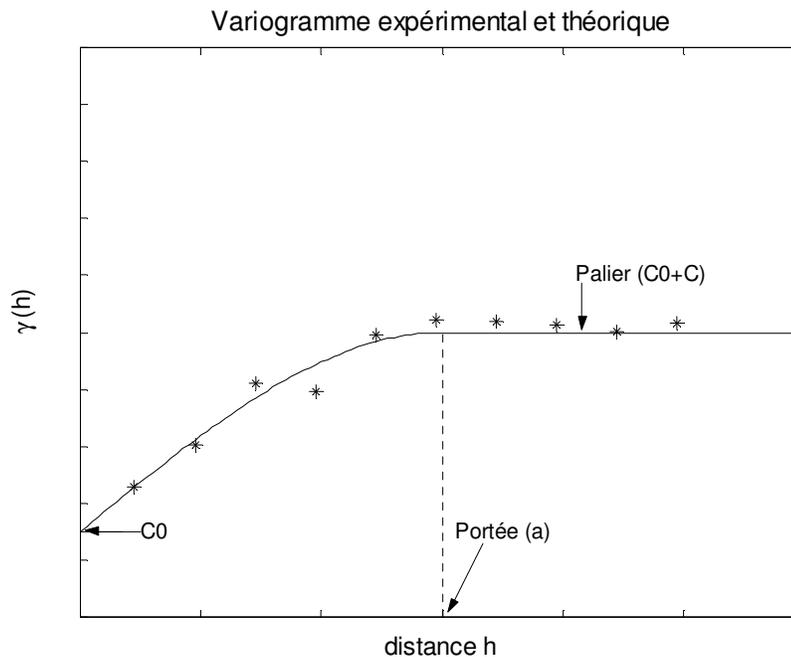
$$\boxed{\gamma(h) = \sigma^2 - C(h)}$$

$C(h)$ est appelé la fonction de covariance de Z . Cette relation est importante et elle est continuellement utilisée en géostatistique.

On voit que lorsque la portée est atteinte, il n'y a plus de covariance entre les v.a., i.e. $C(h) = 0$ si $h \geq a$. Lorsqu'il y a un palier, les deux fonctions sont équivalentes en ce sens qu'elles fournissent la même information sur le processus.

Le variogramme possède toutefois deux avantages sur le covariogramme.

- i. Le variogramme est défini même s'il n'y a pas de palier.
- ii. Dans l'expression du variogramme, la constante "m" n'apparaît pas et l'on n'a donc pas besoin de l'estimer comme c'est le cas lorsqu'on veut calculer directement le covariogramme.



9.2.2 Estimation du variogramme

On estime le variogramme à l'aide de

$$\gamma_e(h) = \frac{1}{2 N(h)} \sum_{i=1}^{N(h)} [Z(x_i) - Z(x_i + h)]^2$$

où $N(h)$ nombre de paires dont les points sont espacés de h .

Pour un champ donné, rien n'assure que la continuité soit identique dans toutes les directions. Par exemple, il se pourrait que des teneurs montrent une meilleure continuité parallèlement à la stratigraphie que perpendiculairement à celle-ci. De même, pour la contamination par des hydrocarbures, on pourrait observer une meilleure continuité horizontalement que verticalement en raison de la gravité. Si le nombre d'observations le permet (typiquement au moins 50, préférablement 100), on peut chercher à vérifier ce point en calculant le variogramme expérimental dans différentes directions.

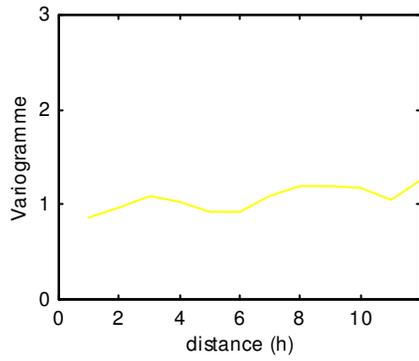
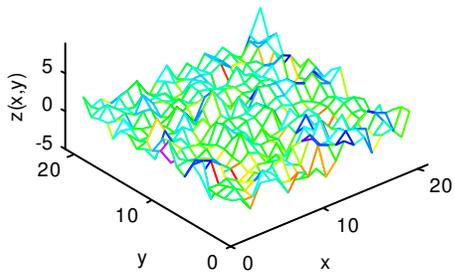
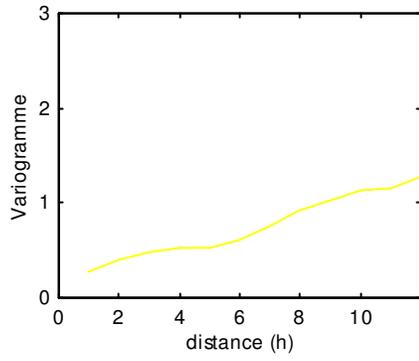
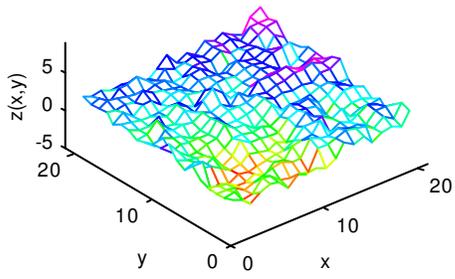
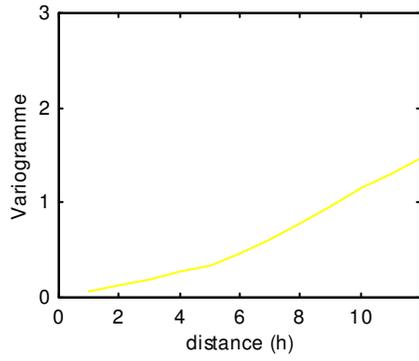
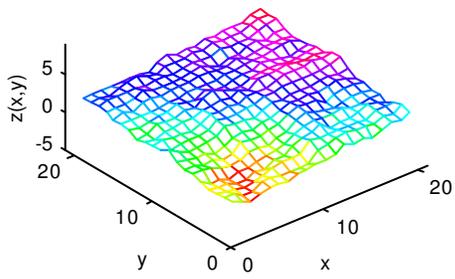
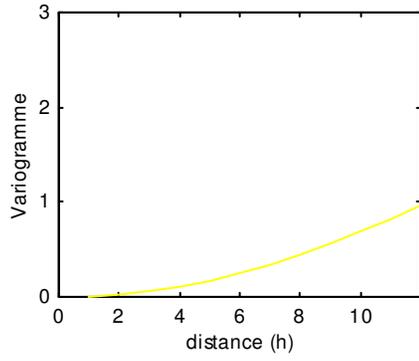
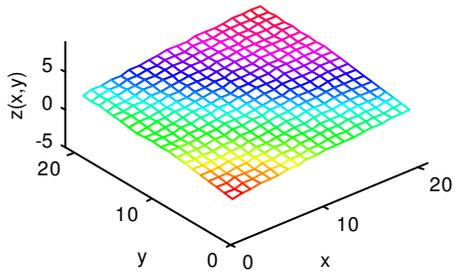
On peut calculer le variogramme selon certaines directions spécifiques:

$$\gamma_e(h, \theta) = \frac{1}{2 N(h, \theta)} \sum_{i=1}^{N(h, \theta)} [Z(x_i) - Z(x_i + h)]^2$$

où $N(h, \theta)$ = nombre de paires séparées de h dans la direction θ .

En pratique on s'accorde une tolérance sur h et sur θ afin d'avoir suffisamment de paires pour chaque h et chaque θ . Pour chacune des classes ainsi formées, on calcule la distance moyenne séparant les extrémités des paires (abscisse) et on évalue le variogramme expérimental pour chaque classe. On obtient donc une série de points expérimentaux auxquels on cherche à ajuster un modèle (i.e. expression analytique) permettant de déduire la covariance entre deux points quelconque en fonction de leur espacement géographique (et, éventuellement, de la direction qu'ils définissent). Une fois le modèle adopté, toute la suite des calculs se fait avec les **valeurs obtenues du modèle** et non avec les valeurs expérimentales.

La figure suivante illustre quelques exemples de surface et le variogramme expérimental correspondant. Les simulations ont été réalisées avec GSLIB-SGSIM, en imposant les valeurs 0, 2, 2 et 4 aux 4 coins. De haut en bas, on a simulé un gaussien de portée 25, un sphérique de portée 25, un sphérique avec 20% d'effet de pépité et portée 25, un sphérique avec 80% d'effet de pépité et portée 25. Comme on le voit, le variogramme expérimental décrit bien le degré d'irrégularité des surfaces.



Exemple numérique

Soit une matrice de données 3 x 3 ayant les valeurs suivantes (la distance horizontale et verticale entre 2 éléments consécutifs est de 1 m et NaN indique une donnée manquante).

3	6	5
7	2	2
4	NaN	0

Le calcul du variogramme selon la direction horizontale donne:

h	$g(h)$	$N(h)$
1	4.375	4
2	7.5	3

Note: $g(1)=0.5/4*[(3-6)^2+(6-5)^2+(7-2)^2+(2-2)^2]$

Dans la direction verticale, on calcule:

h	$g(h)$	$N(h)$
1	5.4	5
2	6.5	2

Dans la direction 45, on calcule:

h	$g(h)$	$N(h)$
1.41	2.33	3
2.82	0.5	1

2e Exemple numérique (1D)

Soit les séquences 1D suivantes :

0 1 2 3 2 1 0

3 1 0 2 1 2 0

Ces deux séries ont même moyenne et même variance, toutefois clairement elles n'ont pas le même degré de continuité spatiale, la 1ère série étant nettement plus continue que la seconde. Voyons leur variogramme:

h	$g(h)$ - 1ère série	$g(h)$ - 2e série
1	0.5	1.25
2	1.6	1.2
3	2.5	1.13

Le variogramme de la 1ère série montre une croissance soutenue alors que la seconde série montre un variogramme à peu près constant à un niveau près de la variance expérimentale (1.06).

9.2.3 Modélisation

Les modèles sont des expressions analytiques que l'on tente d'ajuster le mieux possible aux points des variogrammes expérimentaux.

Condition d'admissibilité des modèles:

Toute fonction ne peut être utilisée comme modèle. Soit une somme quelconque de variables aléatoires (plus généralement, une combinaison linéaire de telles v.a.), la variance de cette combinaison est nécessairement positive (une variance est, par définition, toujours positive). Or cette variance peut s'exprimer en fonction du covariogramme (modèles avec palier) ou du variogramme (modèles avec palier ou sans palier pourvu que la somme des poids de la combinaison linéaire donne 0). Il faut donc que la fonction de covariance ou le variogramme assure des variances positives quelle que soit la combinaison des v.a. considérée.

Bref, soit une combinaison linéaire $\sum_i \lambda_i Z_i$. Dans le cas stationnaire (variogramme avec palier),

$$\text{Var}(\sum_i \lambda_i Z_i) = \sum_i \sum_j \lambda_i \lambda_j \text{Cov}(Z_i, Z_j) = \sum_i \sum_j \lambda_i \lambda_j C(h_{i,j}) \geq 0$$

Dans le cas intrinsèque (variogramme sans palier)

$$\text{Sous la condition } \sum_i \lambda_i = 0, \text{ on a } \text{Var}(\sum_i \lambda_i Z_i) = -\sum_i \sum_j \lambda_i \lambda_j \gamma(h_{i,j}) \geq 0$$

La vérification de l'admissibilité d'un modèle donné est relativement complexe et dépasse le cadre de ce cours. Dans la pratique on se limite à des modèles éprouvés et à des modèles construits à partir de modèles éprouvés en utilisant des propriétés comme :

- une combinaison linéaire (avec coefficients positifs) de variogrammes admissibles donne un modèle admissible;
- un produit de modèles de covariance admissibles donne un modèle de covariance admissible;
- un modèle admissible en \mathbb{R}^p est admissible en \mathbb{R}^{p-1} (l'inverse n'est pas nécessairement vrai).

Types de modèles courants

En géologie, les modèles les plus courants sont :

- Effet de pépite.
- Puissance (cas particulier : linéaire).
- Sphérique.
- Gaussien.
- Exponentiel.

$$\text{Effet de pépite: } \gamma(h) = \begin{cases} 0 & \text{si } h = 0 \\ C_0 & \text{si } h > 0 \end{cases}$$

$$\text{Sphérique : } \gamma(h) = \begin{cases} C [1.5 h/a - 0.5 (h/a)^3] & \text{si } 0 < h < a \\ C & \text{si } h \geq a \end{cases}$$

$$\text{Gaussien: } \gamma(h) = C [1 - \exp(-3(h/a)^2)]$$

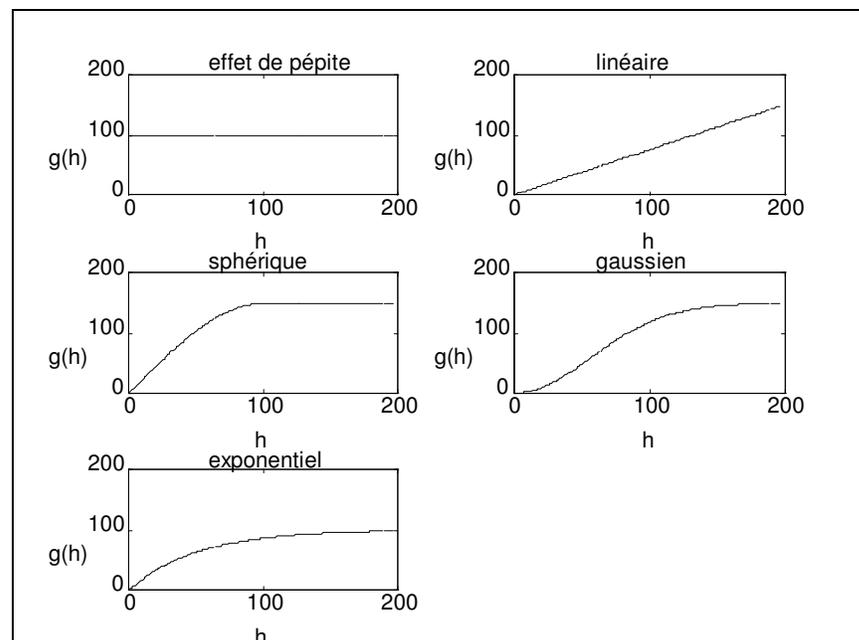
Exponentiel	$\gamma(h) = C [1 - \exp(-3h/a)]$	
Puissance	$\gamma(h) = C h^b$	$0 < b < 2$ (linéaire : $b=1$)

On peut combiner plusieurs modèles en les additionnant. Ainsi, l'effet de pépité est presque toujours présent en association avec un ou plusieurs des autres modèles décrits plus haut. Il est important de noter que ces cinq modèles ne sont pas les seuls que l'on peut utiliser, en réalité, il en existe un très grand nombre.

Il y a un lien étroit entre la nature de la variable étudiée et le type de modèle que l'on est susceptible de rencontrer. Ainsi, le modèle gaussien exprime une très grande continuité typique d'une variable comme la topographie, l'épaisseur d'une formation, le champ gravimétrique, la charge hydraulique. Ce modèle n'est pratiquement jamais rencontré pour des variables comme les teneurs de gisement, les propriétés mécaniques des roches, les analyses géochimiques en général. Pour ces variables, les modèles sphérique et exponentiel sont beaucoup plus courants. Pour des variables discrètes, le modèle gaussien est même à proscrire complètement car incompatible avec ce genre de variables.

Notes :

- i. Lorsque $h = 0$, par définition $\gamma(0)=0$.
Lorsque $h = 0+$, alors $\gamma(h)=C_0$.
L'effet de pépité se présente donc comme une discontinuité à l'origine du variogramme. L'effet de pépité peut représenter des erreurs d'analyse (voir théorie d'échantillonnage de P. Gy), de réelles micro-structures ou/et des structures d'une certaine taille non-détectées par un échantillonnage insuffisant.
- ii. Parfois les variogrammes ne montrent pas de palier (cas du modèle linéaire). D'autre fois ils ne montrent qu'un palier atteint asymptotiquement (cas des modèles exponentiel et gaussien). Dans ce dernier cas, l'on définit la portée effective comme la distance où est atteint 95% du palier. Ainsi, pour les modèles exponentiel et gaussien, la portée effective est "a". Lorsque le variogramme ne montre pas de palier et que sa croissance s'effectue à un taux supérieur à h^2 alors il y a lieu de suspecter une dérive de la moyenne (i.e. l'hypothèse stationnaire ou intrinsèque ne tient pas).



9.2.3.1 Anisotropies

La continuité spatiale n'est pas nécessairement la même dans toutes les directions.

- ex.
- gisement présentant une forme lenticulaire; on peut avoir une meilleure continuité selon l'allongement principal des lentilles;
 - gisement stratiforme; meilleure continuité parallèlement aux strates que perpendiculairement.
 - placer; meilleure continuité le long des paléochenaux que perpendiculairement.
 - etc.

Bien que dans la nature il existe une très grande variété d'anisotropies, en géostatistique, on ne peut modéliser aisément que les anisotropies géométriques. Les autres anisotropies peuvent être approchées en combinant plusieurs modèles isotropes ou avec anisotropie géométrique.

Anisotropie géométrique

Caractéristiques :

- On observe dans diverses directions des **paliers et des composantes pépitiqes identiques mais des portées différentes**.
- Les portées maximales (a_g) et minimales (a_p) s'observent selon deux directions orthogonales.
- On peut rendre les portées identiques (et égales à a_g) suivant toutes les directions en multipliant la composante de la portée parallèle à a_p par le facteur (a_g/a_p). Bref, les portées décrivent une ellipse dont l'axe majeur est orienté parallèlement à a_g .

i.e.

$$\frac{(a_\theta \cos \theta)^2}{a_g^2} + \frac{(a_\theta \sin \theta)^2}{a_p^2} = 1$$

Connaissant a_g et a_p , on peut trouver a_θ , où θ désigne l'angle mesuré par rapport à la direction où est rencontré a_g .

$$a_\theta = \frac{a_g a_p}{\{a_p^2 \cos^2 \theta + a_g^2 \sin^2 \theta\}^{1/2}}$$

Exemple:

Un gisement 2D est modélisé par un modèle avec anisotropie géométrique. Le modèle est sphérique avec $C=17\%^2$ et effet de pépite $C_0=13\%^2$ et les portées sont de 100m dans la direction (convention trigonométrique) de plus grande continuité (30°) et 60m dans la direction de plus petite continuité (120°). Quelle est la valeur du variogramme entre deux observations situées aux coordonnées $(x_1, y_1)=(10, 30)$ et $(x_2, y_2)=(40, 20)$

On calcule la distance séparant les deux points et la direction qu'ils définissent:

$$h = ((20-30)^2 + (40-10)^2)^{1/2} = 31.62\text{m}$$

$$\theta = \arctan\left(\frac{y_2 - y_1}{x_2 - x_1}\right) = \arctan\left(\frac{-10}{30}\right) = -18.4^\circ$$

Cette direction forme un angle de 48.4° avec la direction de plus grande continuité.

On calcule la portée dans cette direction en utilisant la formule plus haut :

$$a_{\theta} = \frac{100 \cdot 60}{\left\{60^2 \cos^2(48.4) + 100^2 \sin^2(48.4)\right\}^{1/2}} = 70.81m$$

On calcule la valeur du variogramme en utilisant l'équation du modèle sphérique pour la distance plus haut et avec la portée 70.81m:

$$\gamma_{\theta}(31.62m) = 13\%^2 + 17\%^2 * \left(1.5 * \frac{31.62m}{70.81m} - 0.5 * \left(\frac{31.62m}{70.81m}\right)^3\right) = 23.63\%^2$$

Remarques importantes concernant la détection d'anisotropies géométriques:

a) Le facteur d'anisotropie géométrique obtenu avec les variogrammes expérimentaux sous estime en général le véritable facteur d'anisotropie en raison de l'utilisation d'une fenêtre angulaire et du fait que les variogrammes expérimentaux ne sont pas nécessairement orientés exactement selon les directions principales de l'ellipse d'anisotropie.

b) L'estimation correcte et à la limite, la détection, d'anisotropie géométrique n'est possible, en pratique, qu'à quatre conditions (fortement liées) devant être remplies simultanément:

- Le nombre de données est suffisant (au moins 50)
- Le facteur d'anisotropie est important (au moins 1.5)
- Une des directions utilisées dans le calcul du variogramme est près de la direction de plus grande portée.
- La fenêtre angulaire utilisée est suffisamment étroite.

9.2.4 Remarques concernant le calcul de variogrammes et l'ajustement de modèles

- On accorde plus de poids aux points du variogramme expérimental calculés avec beaucoup de paires.
- On essaie d'avoir $N(h) \geq 30$ pour chaque point expérimental du variogramme. Si ce n'est pas possible pour certaines classes, on accorde moins d'importance à ces points. Si le nombre de paires est très faible (≤ 10), on ne considère plus du tout le point.
- On accorde plus de poids aux premiers points du variogramme (h petit) car ce sont ces valeurs qui ont le plus d'impact dans les calculs géostatistiques.
- Lorsque h dépasse environ $d_{\max}/2$, on ne tient pas compte des valeurs du variogramme. d_{\max} est la taille du phénomène étudié dans la direction considérée.
- On cherche à obtenir des modèles les plus simples possible qui rendent bien compte des valeurs expérimentales.

Stratégie de modélisation (cas 2D)

- Calculer les variogrammes directionnels selon différentes directions (ex. 0° , 45° , 90° , 135°) ainsi que le variogramme omnidirectionnel (i.e. sans tenir compte de la direction).
La géologie peut apporter une information précieuse dans le choix des directions et la présence ou non d'anisotropies.
- Vérifier les critères ci-dessus : $N(h) \geq 30$, $h < d_{\max}/2$
- Si nécessaire, augmenter la tolérance angulaire ou le pas de calcul de façon à augmenter $N(h)$.

- Déterminer s'il y a anisotropie (différences de palier ou de portées qui ne peuvent raisonnablement être imputées à des fluctuations aléatoires du variogramme).
- Procéder à l'ajustement d'un modèle anisotrope ou isotrope selon le cas (habituellement par essai et erreur, bien que l'on puisse aussi obtenir ces ajustements de façon automatique par régression (pondérée, et souvent, non-linéaire)).
- Chercher à respecter la règle de la parcimonie: adopter les modèles les plus simples possibles qui permettent un ajustement adéquat. Comparer des modèles concurrents à l'aide de la *technique de validation croisée*.

9.3 Variance d'estimation

Dans cette section, on cherche à établir les résultats permettant de fournir une mesure de la précision des estimés effectués par une méthode d'estimation quelconque (linéaire).

Soit une v.a. Z_v que l'on veut estimer (ici l'indice v indique que l'on s'intéresse à l'estimation d'un bloc, si v est 0 alors on estime un point), d'une façon ou d'une autre, en formant une combinaison linéaire des valeurs observées en différents endroits, i.e.

$$Z_v^* = \sum_{i=1}^n \lambda_i Z_i \quad (1)$$

où:

Z_i : valeur observée au point x_i (v.a.)
 Z_v^* : estimateur de Z_v

On définit l'erreur d'estimation :

$$e = Z_v - Z_v^*$$

La variance de cette erreur est la **variance d'estimation** :

$$Var(e) = Var(Z_v) + Var(Z_v^*) - 2Cov(Z_v, Z_v^*)$$

Substituant Z_v^* par son expression, en fonction des Z_i , donnée en (1), on obtient:

$$\sigma_e^2 = Var(Z_v) + \sum_i \sum_j \lambda_i \lambda_j Cov(Z_i, Z_j) - 2 \sum_i \lambda_i Cov(Z_i, Z_v) \quad (2)$$

Qui peut être réécrit en fonction du variogramme:

$$\sigma_e^2 = (\sigma^2 - \bar{\gamma}(v, v)) + \sum_i \sum_j \lambda_i \lambda_j (\sigma^2 - \gamma(x_i - x_j)) - 2 \sum_i \lambda_i (\sigma^2 - \bar{\gamma}(x_i, v))$$

Puis finalement, puisqu'on a habituellement $\sum \lambda_i = 1$,

$$\sigma_e^2 = 2 \sum_i \lambda_i \bar{\gamma}(x_i, v) - \bar{\gamma}(v, v) - \sum_i \sum_j \lambda_i \lambda_j \gamma(x_i, x_j) \quad (3)$$

On peut calculer la variance d'estimation soit en utilisant le covariogramme (2) ou le variogramme (3). Dans les cas de modèles sans palier, seul le variogramme peut être utilisé pourvu que $\sum \lambda_i = 1$ dans (1).

- i. Dans les formules précédentes (2 et 3), on reconnaît 3 termes : 1 terme lié au bloc à estimer ($\text{Var}(Z_v)$ ou $\bar{\gamma}(v, v)$), 1 terme lié aux points servant à l'estimation ($\text{Cov}(Z_i, Z_j)$ ou $\gamma(x_i, x_j)$) et 1 terme croisé entre les points servant à l'estimation et le bloc à estimer ($\text{Cov}(Z_i, Z_v)$ ou $\bar{\gamma}(x_i, v)$).
- ii. La variance de l'erreur d'estimation est une mesure de la précision de l'estimation. On pourrait vouloir choisir les λ_i de façon à ce que σ_e^2 soit minimale. C'est ce que nous ferons plus tard avec le krigeage.
- iii. La variance d'estimation est une mesure de précision obtenue, en moyenne, sur l'ensemble du gisement pour une même configuration points-bloc. On constate en effet que la variance d'estimation peut être calculée dès que l'on connaît le variogramme (ou le covariogramme), l'emplacement des points de données et le bloc à estimer. La variance d'estimation ne permet donc pas de tenir compte du fait que certaines portions pauvres du gisement sont peut-être plus faciles à estimer que des zones à haute teneur (effet proportionnel rencontré par exemple avec des distributions lognormales). On peut en tenir compte par l'emploi d'un variogramme relatif (normé par la moyenne locale au carré). Toutefois l'estimation du variogramme relatif est souvent délicate. Les méthodes non-linéaires et les simulations permettent de mieux tenir compte de ce facteur.

Ex. Soit x_1, x_2 et x_3 trois points échantillonnés que l'on veut utiliser pour estimer la valeur inconnue au point x_0 .

x_1

$x_2 \quad x_0 \quad x_3$

Soit la matrice des distances séparant ces points:

	x_0	x_1	x_2	x_3
x_0	0	1.4	1	2
x_1		0	1	3.2
x_2			0	3

Supposons que le variogramme est un modèle linéaire avec pente unitaire et effet de pépité égal à 1, i.e.:

$$\gamma(h) = 1 + h$$

Calculons la variance d'estimation obtenue par trois méthodes différentes d'estimation:

a) Estimation polygonale (plus proche voisin)

$$\lambda_1 = \lambda_3 = 0 \quad \lambda_2 = 1$$

$$Z_0^* = Z_2$$

$$\begin{aligned} \sigma_{\varepsilon}^2 &= 2 \gamma(x_0, x_2) - \gamma(x_0, x_0) - \gamma(x_2, x_2) \\ &= 2(1 + 1) - 0 - 0 = 4 \end{aligned}$$

b) Inverse de la distance

$$1/d_1 = 1/1.4 \quad 1/d_2 = 1 \quad 1/d_3 = 1/2$$

$$\sum_i 1/d_i = 2.21$$

$$\lambda_1 = .32 \quad \lambda_2 = .45 \quad \lambda_3 = .23$$

$$\begin{aligned} \sigma_{\varepsilon}^2 &= 2 \sum_i \lambda_i \gamma(x_i, x_0) - \gamma(x_0, x_0) - \sum_i \sum_j \lambda_i \lambda_j \gamma(x_i, x_j) \\ &= 2 (.32(1+1.4) + .45(1+1) + .23(1+2)) - 0 \\ &\quad - 2 (.32 .45(1+1) + .32 .23(1+3.2) + .45 .23(1+3)) \\ &\approx 2.7 \end{aligned}$$

Ici l'inverse de la distance est très supérieur à la méthode polygonale en terme de variance d'estimation.

c) krigage

Par krigage, on aurait obtenu $\lambda_1=.25$, $\lambda_2=.43$, et $\lambda_3=.32$. Le calcul de la variance d'estimation donne alors $\sigma^2 \cong 2.65$, une amélioration négligeable par rapport à l'inverse de la distance.

9.4 Krigage

Puisqu'on peut calculer la variance d'estimation pour tout estimateur linéaire, pourquoi ne pas choisir celui qui assure la variance d'estimation minimale? C'est précisément ce qu'effectue le krigage. Dans le cas stationnaire, on en reconnaît 2 types principaux, selon que la moyenne du processus est connue ou non, soit le krigage simple et le krigage ordinaire. Ce dernier est, de loin, le plus fréquemment utilisé.

9.4.1 Krigeage ordinaire

Supposons que l'on veuille estimer un bloc v centré au point x_0 . Notons Z_v la vraie valeur (inconnue) de ce bloc et Z_v^* l'estimateur que l'on obtient.

L'estimateur est linéaire, i.e.:

$$Z_v^* = \sum_{i=1}^n \lambda_i Z_i$$

où les Z_i désignent les v.a. correspondant aux points échantillons.

On veut minimiser:

$$\sigma_e^2 = \text{Var}[Z_v - Z_v^*] = \text{Var}[Z_v] + \text{Var}[Z_v^*] - 2 \text{Cov}[Z_v, Z_v^*]$$

Substituant l'expression de l'estimateur dans cette équation, on obtient:

$$\sigma_e^2 = \text{Var}[Z_v] + \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \text{Cov}[Z_i, Z_j] - 2 \sum_{i=1}^n \lambda_i \text{Cov}[Z_v, Z_i]$$

Pour que l'estimateur soit sans biais, il faut que:

$$\sum_{i=1}^n \lambda_i = 1$$

En effet, dans ce cas, $E[Z_v^*] = \sum_i \lambda_i E[Z_i] = \sum_i \lambda_i m = m$

On a un problème de minimisation d'une fonction quadratique (donc convexe) sous contrainte d'égalité que l'on résout par la méthode de Lagrange. On forme le lagrangien:

$$\begin{aligned} L(\lambda) &= \sigma_e^2 + 2\mu \left(\sum_{i=1}^n \lambda_i - 1 \right) \\ &= \text{Var}[Z_v] + \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \text{Cov}[Z_i, Z_j] - 2 \sum_{i=1}^n \lambda_i \text{Cov}[Z_v, Z_i] + 2\mu \left(\sum_{i=1}^n \lambda_i - 1 \right) \end{aligned}$$

Où μ est le multiplicateur de Lagrange. Le minimum est atteint lorsque toutes les dérivées partielles par rapport à chacun des λ_i et par rapport à μ s'annulent. Ceci conduit au système de krigeage ordinaire:

Système de krigeage ordinaire

$$\sum_{j=1}^n \lambda_j \text{Cov}[Z_i, Z_j] + \mu = \text{Cov}[Z_v, Z_i] \quad \forall i = 1 \dots n$$

$$\sum_{j=1}^n \lambda_j = 1$$

La variance d'estimation minimale, appelée variance de krigeage, est obtenue en substituant les équations de krigeage dans l'expression générale pour la variance d'estimation:

$$\sigma_k^2 = \text{Var}[Z_v] - \sum_{i=1}^n \lambda_i \text{Cov}[Z_v, Z_i] - \mu$$

Note: Cette variance de krigeage ne dépend pas des valeurs observées, elle ne dépend que du variogramme et de la configuration des points servant à l'estimation par rapport au point (ou bloc) à estimer.

Système de krigeage écrit en terme du variogramme:

Comme la variance d'estimation s'écrit aussi directement en terme du variogramme, on peut aussi écrire le système de krigeage en fonction du variogramme. Ceci tient au fait que $C(h) = \sigma^2 - \gamma(h)$ et que $\sum \lambda_i = 1$.

$$\sum_{j=1}^n \lambda_j \gamma(x_i, x_j) - \mu = \bar{\gamma}(v, x_i) \quad \forall i = 1 \dots n$$

$$\sum_{j=1}^n \lambda_j = 1$$

et, alors

$$\sigma_k^2 = \sum_{i=1}^n \lambda_i \bar{\gamma}(v, x_i) - \bar{\gamma}(v, v) - \mu$$

Il est intéressant de visualiser le système de krigeage ordinaire et la variance de krigeage ordinaire sous forme matricielle:

$$K_o \lambda_o = k_o$$

$$\sigma_{k_o}^2 = \sigma_v^2 - \lambda_o' k_o$$

où

$$K_o = \begin{bmatrix} \sigma^2 & \text{Cov}(Z_1, Z_2) & \bullet & \text{Cov}(Z_1, Z_n) & 1 \\ \text{Cov}(Z_2, Z_1) & \sigma^2 & \bullet & \text{Cov}(Z_2, Z_n) & 1 \\ \bullet & \bullet & \bullet & \bullet & \bullet \\ \text{Cov}(Z_n, Z_1) & \text{Cov}(Z_n, Z_2) & \bullet & \sigma^2 & 1 \\ 1 & 1 & \bullet & 1 & 0 \end{bmatrix}$$

$$k_o = \begin{bmatrix} \text{Cov}(Z_1, Z_v) \\ \text{Cov}(Z_2, Z_v) \\ \bullet \\ \text{Cov}(Z_n, Z_v) \\ 1 \end{bmatrix}, \quad \lambda_o = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \bullet \\ \lambda_n \\ \mu \end{bmatrix}$$

Note : tous les termes $\text{Cov}(Z_i, Z_v)$ et $\text{Var}(Z_v)$ s'obtiennent directement du variogramme par intégration numérique considérant que $Z_v = \frac{1}{v} \int_v Z(x) dx$ où « x » désigne une localisation dans le bloc « v ». Profitant de la linéarité de l'opérateur espérance, on montre que :

$$\text{Cov}(Z_i, Z_v) = \frac{1}{v} \int \text{Cov}(Z_i, Z(x)) dx \quad \text{où } \ll x \gg \text{ est un point dans } \ll v \gg, \ll x \gg \text{ balaie } \ll v \gg.$$

et

$$\text{Var}(Z_v) = \frac{1}{v^2} \iint \text{Cov}(Z(x), Z(y)) dx dy \quad \text{où } \ll x \gg \text{ et } \ll y \gg \text{ sont deux points dans } \ll v \gg, \ll x \gg \text{ et } \ll y \gg \text{ balaient tous deux indépendamment } \ll v \gg.$$

9.4.2 Quelques cas très simples de krigeage ordinaire

Ces quelques cas sont présentés dans le seul but d'acquérir une certaine intuition du comportement du krigeage. On suppose un variogramme sphérique de portée finie "a".

i. Estimation d'un point par un autre point situé à une distance "h"

$$\lambda_1 = 1, \quad \sigma_{k_o}^2 = 2(\sigma^2 - C(h)) = 2\gamma(h) \quad (\text{Note si } h > a, \sigma_{k_o}^2 = 2\sigma^2)$$

Remarque: Il est possible d'avoir une variance de krigeage ordinaire supérieure à la variance théorique de la variable étudiée!

ii. Estimation d'un point situé en x_0 par deux points situés en " x_1 " et " x_2 "

$$\lambda_1 = \frac{\sigma^2 + C(x_0, x_1) - C(x_1, x_2) - C(x_0, x_2)}{2(\sigma^2 - C(x_1, x_2))}, \quad \lambda_2 = \frac{\sigma^2 + C(x_0, x_2) - C(x_1, x_2) - C(x_0, x_1)}{2(\sigma^2 - C(x_1, x_2))}$$

Note: dans les deux cas, les poids peuvent être négatifs dépendant de la position respective des trois points.

iii. Estimation d'un point par "n" points en présence d'un variogramme effet de pépité pur.

$$\lambda_i = \frac{1}{n}, \quad \text{et} \quad \sigma_{k_o}^2 = \frac{(n+1)}{n} \sigma^2$$

9.4.4 Propriétés du krigeage

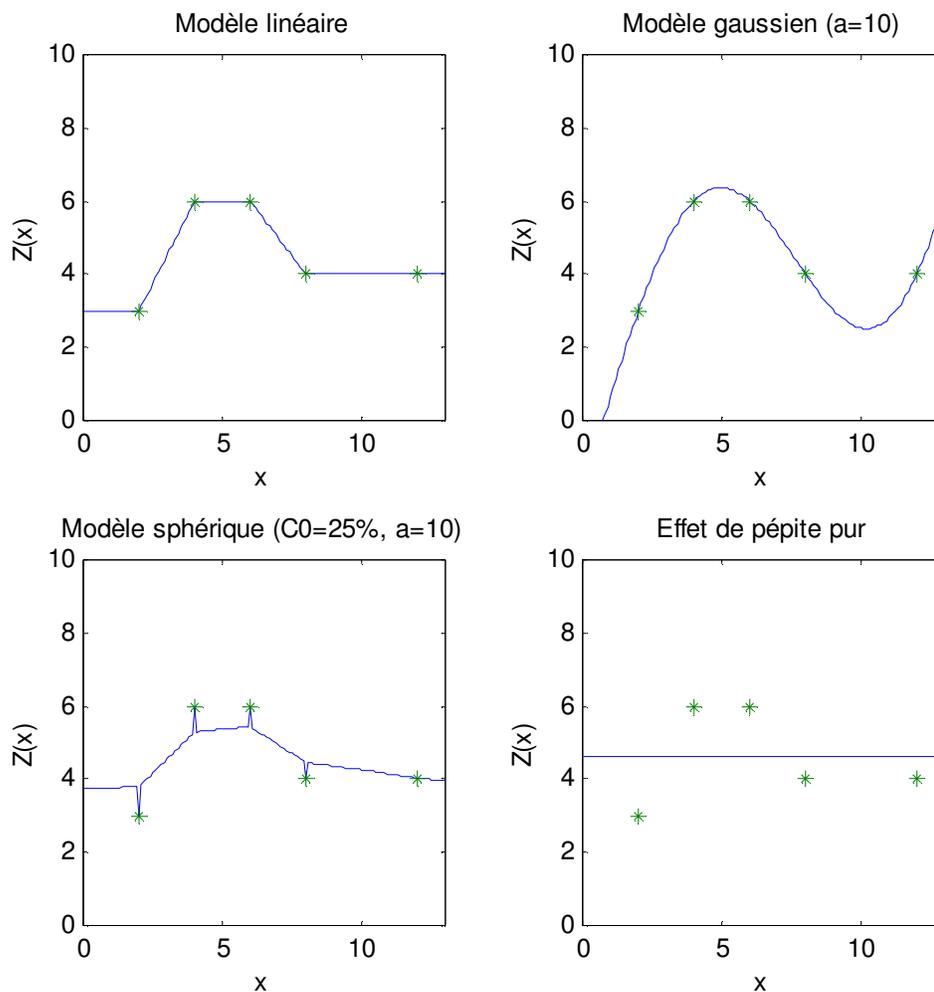
Les principales propriétés et caractéristiques associées au krigeage sont:

- i. Linéaire, sans biais, à variance minimale, par construction.
- ii. Interpolateur exact. : si l'on estime un point connu, on retrouve la valeur connue.
- iii. Présente un effet d'écran: les points les plus près reçoivent les poids les plus importants. Cet effet d'écran varie selon la configuration et selon le modèle de variogramme utilisé pour le krigeage. Plus l'effet de pépité est important, moins il y a d'effet d'écran.
- iv. Tient compte de la taille du champ à estimer et de la position des points entre eux.

- v. Par l'utilisation du variogramme, tient compte de la continuité du phénomène étudié (effet de pépite, anisotropie, etc.).
- vi. Effectue généralement un lissage, i.e. les estimations sont moins variables que les teneurs réelles (point ou bloc) que l'on cherche à estimer.
- vii. Presque sans biais conditionnel. Ceci signifie que lorsqu'on applique une teneur de coupure à des valeurs estimées, on récupérera approximativement la teneur prévue. C'est une propriété très importante pour les mines. Cette propriété implique que l'estimateur utilisé soit plus lisse que la valeur qu'il cherche à estimer, ce qui est le cas pour le krigeage.
- viii. Transitif. Si l'on observe en un point une valeur coïncidant avec la valeur krigée pour ce point, alors les valeurs krigées en d'autres points ne sont pas modifiées par l'inclusion de ce nouveau point dans les krigeages. Par contre les variances de krigeage, elles, sont diminuées. De même, si l'on krige un certain nombre de points et que l'on utilise les valeurs krigées comme si c'étaient de nouvelles données, alors les krigeages subséquents ne s'en trouvent pas modifiés (sauf pour la variance de krigeage).

INTERPOLATEUR EXACT

Exemples d'interpolation par krigeage en 1D, utilisant différents modèles de variogrammes:

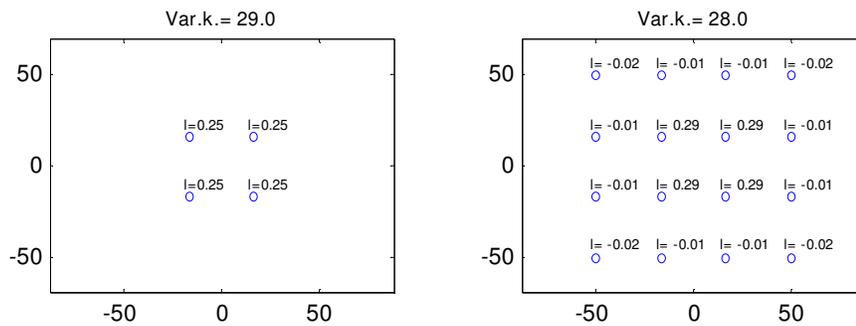


Note: Aux points échantillons, le krigeage retourne la valeur de l'échantillon. Pour éviter les discontinuités dans des cartes il est donc recommandé de ne pas kriger un point échantillon. En somme, on s'assure d'avoir au moins une distance "epsilon" entre le point à kriger et le point échantillon. Comme souvent l'effet de pépite représente une erreur de mesure, il est justifié de s'écartier des valeurs observées.

EFFET D'ÉCRAN

- Cas extrême : modèle linéaire en 1-D
- Diminue lorsque l'effet de pépite augmente (il n'y a pas d'effet d'écran lorsqu'on a un effet de pépite pur)
- Permet de limiter les systèmes de krigeage aux observations avoisinantes (voisinages glissants)

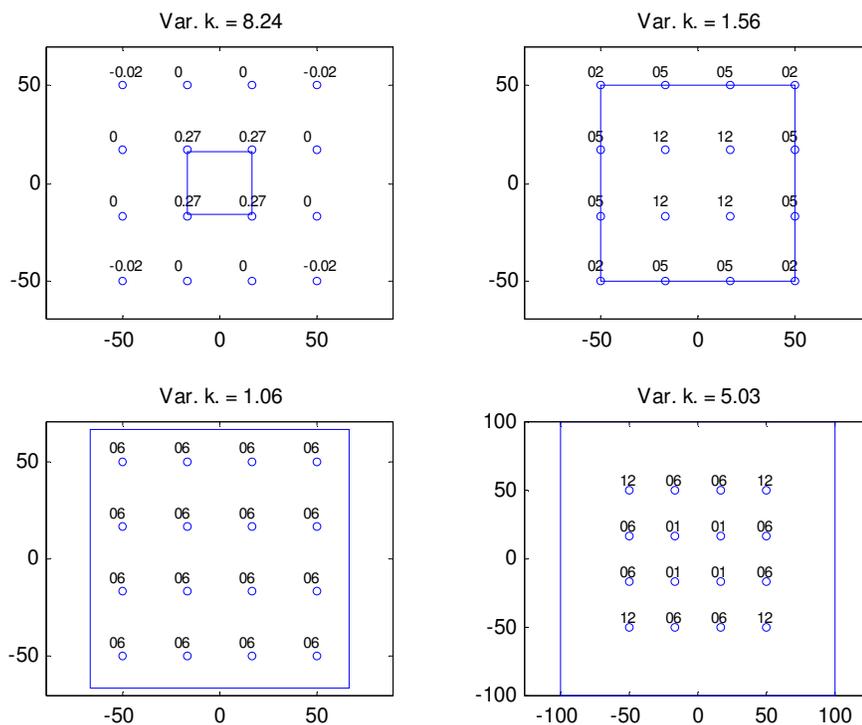
Variogramme sphérique; $C=100$, $a=100$, $C_0=0$



INFLUENCE DE LA TAILLE DU CHAMP

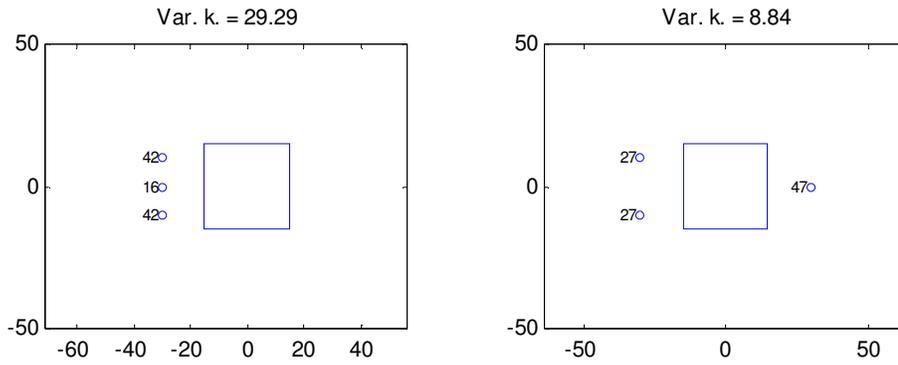
Lorsque la taille du champ estimé augmente,

- Les poids tendent à devenir égaux
- La variance d'estimation diminue puis augmente si on cherche à estimer un champ plus grand que celui renfermant les données (extrapolation)



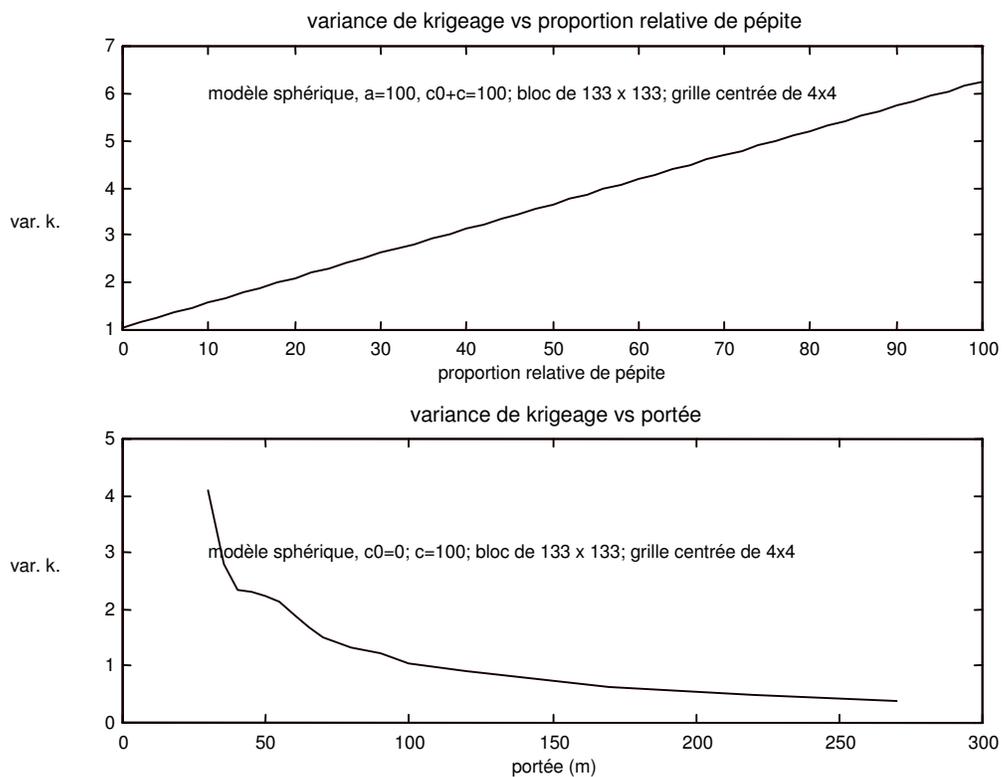
POSITION DES POINTS ENTRE EUX

Contrairement aux méthodes de type "inverse de la distance", la position des points entre eux est très importante. Chaque point est pondéré automatiquement en fonction de sa "zone d'influence". (Les poids par inverse de la distance auraient été 1/3 pour chaque point dans les 2 cas). (Toujours variogramme sphérique avec $a=100$, $C=100$, $C_0=0$).



INFLUENCE DE L'EFFET DE PÉPITE ET DE LA PORTÉE

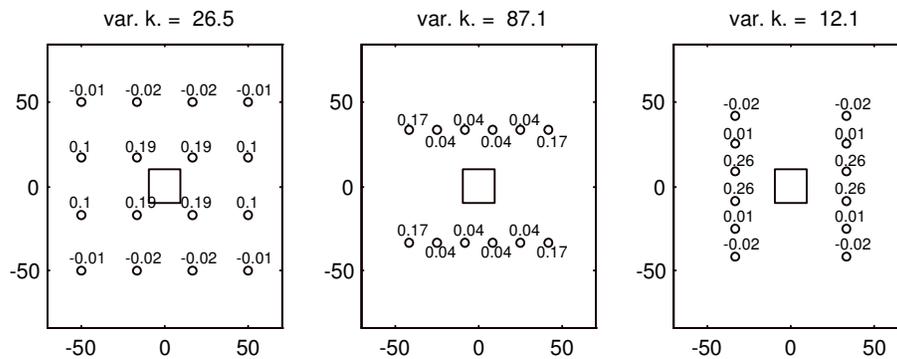
Plus l'effet de pépité est important (relativement à un plateau fixe), plus la variance d'estimation augmente. Inversement, plus la portée augmente, plus la variance d'estimation diminue.



INFLUENCE D'ANISOTROPIES

On doit adapter l'échantillonnage en augmentant la densité d'échantillonnage dans la direction de plus faible portée.

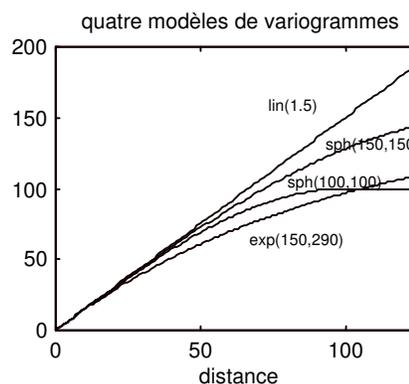
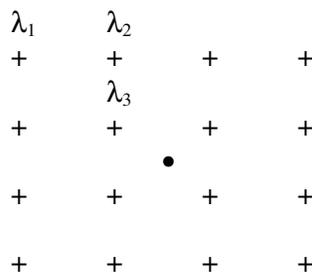
Dans cet exemple, le modèle est sphérique avec $C_0=0$; $C=100$ et $a_x=200$ et $a_y=50$. Les 3 exemples ci-contre correspondent à une même densité d'échantillonnage (1 échantillon par surface de 33×33 unités). Pour le même coût d'échantillonnage on peut donc obtenir des estimations beaucoup plus précises si l'on ajuste la stratégie d'échantillonnage à l'anisotropie.



5

INFLUENCE DU CHOIX DU MODÈLE

Le choix du modèle a peu d'influence sur les résultats du krigeage pour autant que chaque modèle fournisse un ajustement équivalent pour les courtes distances. Ici, le champ fait 100m x 100m et chaque point est espacé de 33.3m. On estime le point au centre de la grille. Les modèles théoriques fournissent à peu près les mêmes valeurs pour les distances de 0 à 25m, or les points centraux, recevant les poids les plus élevés, sont à 24m du point à estimer.



Sphérique:	C=100 a=100m $\lambda_1=-.02$ $\lambda_2=-.01$ $\lambda_3=.29$ $\sigma_k^2=28.0$	Sphérique:	C=150 a=150m $\lambda_1=-.01$ $\lambda_2=-.01$ $\lambda_3=.29$ $\sigma_k^2=27.8$
Exponentiel:	C=150 a=290m $\lambda_1=-.01$ $\lambda_2=-.01$ $\lambda_3=.28$ $\sigma_k^2=28.2$	Linéaire:	Pente=1.5 $\lambda_1=-.01$ $\lambda_2=-.01$ $\lambda_3=.28$ $\sigma_k^2=27.6$

EFFET DE LISSAGE

Krigeage ordinaire:

Des équations du krigeage ordinaire, il découle directement que:

$$\text{Var}(Z_v) = \text{Var}(Z_v^*) + \sigma_{ko}^2 + 2\mu$$

Pour "v" fixe, le terme $\text{Var}(Z_v)$ ne dépend pas de la localisation, les termes $\text{Var}(Z_v^*)$ et σ_{ko}^2 et μ , eux, dépendent du bloc considéré et des échantillons disponibles. Normalement, $\sigma_{ko}^2 + 2\mu > 0$, d'où l'effet de lissage annoncé.

Exemple. Considérons un bloc carré de taille 10 x 10 estimé par ses 4 coins. Le variogramme est sphérique avec palier de 1 et portée de 20. L'estimation est faite par krigeage ordinaire (poids égaux à 0.25).

On peut calculer :

$$\text{Var}(Z_v) = 0.6278$$

$$\sigma_{ko}^2 = 0.1311$$

$$\text{De plus, } \text{Var}(Z_v^*) = 1/16 * (4*1 + 8*0.3125 + 4*0.1161) = 0.4353$$

On trouve en substituant dans les équations de krigeage ordinaire $\mu = 0.0307$

On a bien. $0.4353 + 0.1311 + 2*0.0307 = 0.6278$ et $\text{Var}(Z_v^*) < \text{Var}(Z_v)$

BIAIS CONDITIONNEL

Considérons la teneur réelle du bloc Z_v et son estimation Z_v^* . Supposons que l'espérance conditionnelle de Z_v étant donné Z_v^* est linéaire (ce sera assuré si les deux suivent une loi binormale). On aura alors:

$$E[Z_v | Z_v^*] = a + bZ_v^*$$

$$\text{où } b = \frac{\text{Cov}(Z_v, Z_v^*)}{\text{Var}(Z_v^*)} \text{ et } a = (1-b)m$$

Krigeage ordinaire

Par construction on a alors:

$$\text{Var}(Z_v^*) + \mu = \text{Cov}(Z_v, Z_v^*) \Rightarrow b = 1 + \frac{\mu}{\text{Var}(Z_v^*)}, a = \frac{-\mu}{\text{Var}(Z_v^*)}$$

Conséquemment,

$$E[Z_v | Z_v^*] = Z_v^* + \frac{\mu}{\text{Var}(Z_v^*)}(Z_v^* - m)$$

ce qui indique que le krigeage présente un biais conditionnel. Ce biais sera très faible lorsque l'estimation sera précise (faible variance de krigeage et multiplicateur de Lagrange près de zéro, forte $\text{Var}(Z_v^*)$).

Généralement, le multiplicateur de Lagrange est légèrement négatif, ce qui implique que la pente de la régression est inférieure à 1. Donc en utilisant les valeurs krigées directement, on surestime légèrement aux fortes valeurs et on sous-estime aux faibles valeurs.

Note : pour l'estimateur par méthode polygonale (plus proche voisin), l'on a :

$$b = \frac{\text{Cov}(Z_v, Z_v^*)}{\text{Var}(Z_v^*)} = \frac{\text{Cov}(Z_v, Z_i)}{\sigma^2} < 1$$

cet estimateur présente un biais conditionnel qui sera d'autant plus important que le point utiliser sera éloigné du bloc à estimer.

Remarque: lien entre lissage et biais conditionnel

Comme on l'a vu, on a $b = \frac{\text{Cov}(Z_v, Z_v^*)}{\text{Var}(Z_v^*)}$

On peut réécrire cela comme: $b = \frac{\rho\sigma_v\sigma_v^*}{\text{Var}(Z_v^*)} = \frac{\rho\sigma_v}{\sigma_v^*}$

ρ est le coefficient de corrélation entre Z_v et Z_v^* et est nécessairement inférieur (ou égal) à 1 et $\sigma_v^* = \text{Var}(Z_v^*)^{0.5}$. Pour que $b=1$, il faut donc obligatoirement que l'on ait $\sigma_v^* \leq \sigma_v$. On conclut que si un estimateur est plus variable que la quantité qu'il cherche à estimer alors il présente certainement un biais conditionnel (la pente de la régression sera inférieure à 1). C'est, par exemple, le cas pour l'estimateur par méthode polygonale ou la variance des valeurs estimées est égale à la variance des données ponctuelles. Le lissage de l'estimateur (propriété du krigeage) est un préalable essentiel à l'absence de biais conditionnel.

9.4.5 Pratique du krigeage

Grille de krigeage: Souvent, le krigeage est réalisé sur une grille régulière de points ou de blocs.

Dans le cas de points, L'objectif est habituellement de fournir une carte de la variable étudiée. La grille de krigeage doit être alors assez dense pour que la carte corresponde effectivement au krigeage et non à la méthode particulière (souvent inconnue) utilisée pour tracer les isocontours.

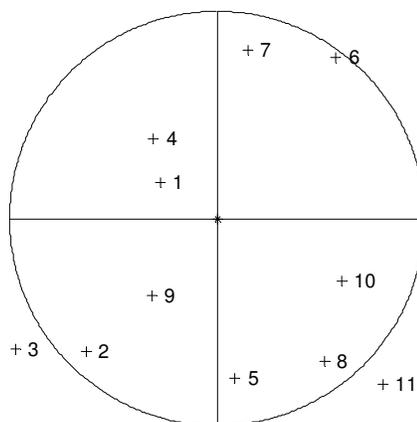
Voisinage utilisé pour le krigeage:

- i. Habituellement en voisinages glissants.
- ii. Nombre de points suffisant (>10 ; peut atteindre jusqu'à 50-100).
- iii. Zone de recherche des points assez grande pour assurer un minimum de points dans le krigeage.
S'il y a anisotropie, on peut adopter une zone de recherche elliptique parallèle à la direction de meilleure continuité. Toutefois une zone de recherche circulaire peut être suffisante si l'on augmente suffisamment le nombre de points dans le krigeage.
- iv. Recherche par quadrants assure une répartition plus uniforme des points (exiger au moins 2 ou 3 points par quadrant)

Exemple: Recherche circulaire avec un maximum de deux points par quadrant.

3 et 11 sont rejetés car en dehors du cercle de recherche.

8 est rejeté car deux autres points sont plus rapprochés du point à estimer dans ce quadrant.



9.4.6 Validation croisée

Une pratique intéressante pour valider le modèle de variogramme et le voisinage utilisé pour le krigeage consiste à effectuer une validation croisée. Le principe est d'éliminer à tour de rôle chaque observation et de l'estimer à l'aide de ses voisins. En chaque point, on obtient donc une valeur vraie et une valeur estimée que l'on peut comparer pour déterminer si le modèle fournit des estimations se comportant comme prévu, si le voisinage utilisé est adéquat, etc.

Plus précisément, soit Z_i^* l'estimation obtenue par krigeage au point "i" (en enlevant la valeur observée Z_i) ainsi que la variance de krigeage σ_{ki}^2 . On peut définir un résidu $e_i = Z_i - Z_i^*$ et un résidu normalisé $n_i = \frac{e_i}{\sigma_{ki}}$. Un modèle et un voisinage adéquats devraient fournir:

- i. $\sum_i e_i \approx 0$ et $\sum_i n_i \approx 0$

ii. $\sum_i |e_i| \text{ min ou } \sum_i e_i^2 \text{ min}$

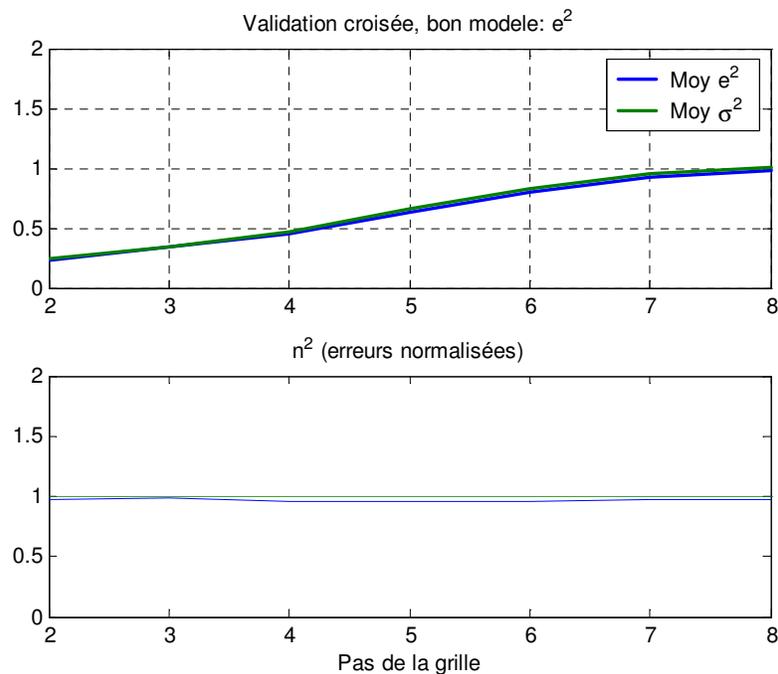
iii. $\left(\frac{1}{n} \sum_i n_i^2 \right)^{0.5} \approx 1$

- iv. Il faut aussi examiner l'histogramme des e_i et des n_i , de même que leur disposition spatiale pour vérifier si les statistiques précédentes pourraient être causées par 1 ou 2 données extrêmes et vérifier si les résidus sont spatialement homogènes.

Illustration de la validation croisée

Les 4 figures suivantes montrent les résultats de simulation effectuées pour 1600 points (40 x 40) à des pas variables (abscisse sur les graphes). En ordonnée, on retrouve dans la figure du haut la moyenne des erreurs de krigeage (par validation) au carré et la moyenne des variances de krigeage. Dans la figure du bas on a la moyenne des erreurs de krigeage normalisées par la variance de krigeage. Tous les krigeages sont effectués avec 50 voisins. Le véritable modèle utilisé pour la simulation est sphérique avec $a=10$ ($C_0=0$) pour les 3 premières figures et un effet de pépite pur pour la dernière. Dans tous les cas, la variance des données simulées est 1.

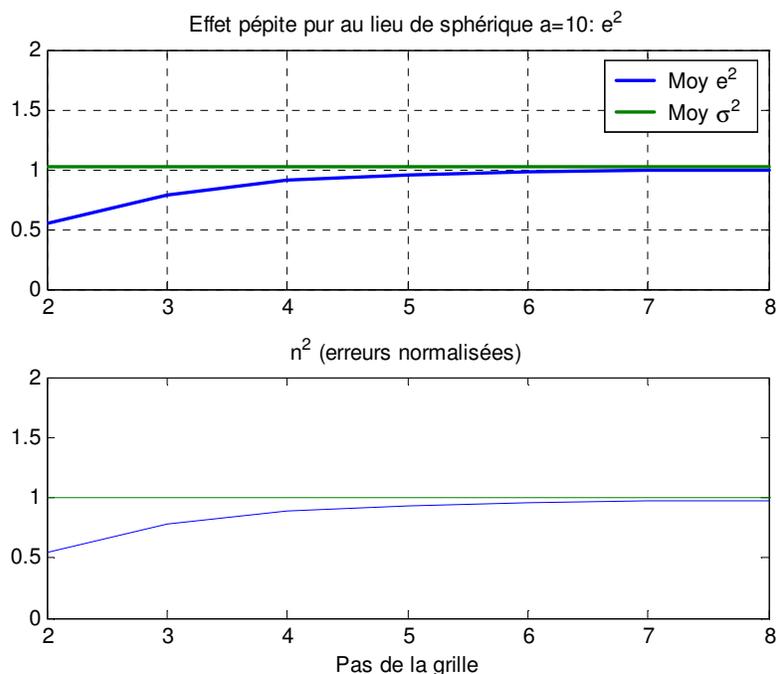
Figure 1 : le krigeage est effectué avec le bon modèle.



On note :

- la variance de krigeage prédit parfaitement la précision accrue due à une grille plus resserrée;
- les erreurs normalisées ont une variance de 1 comme prévu.

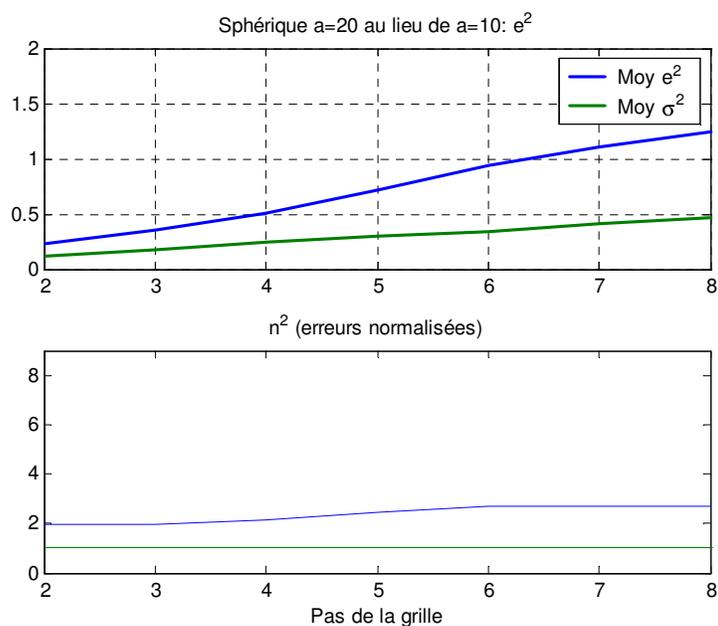
Figure 2 : On a fourni un modèle trop pessimiste (effet de pépité pur) au lieu du vrai modèle :



On note :

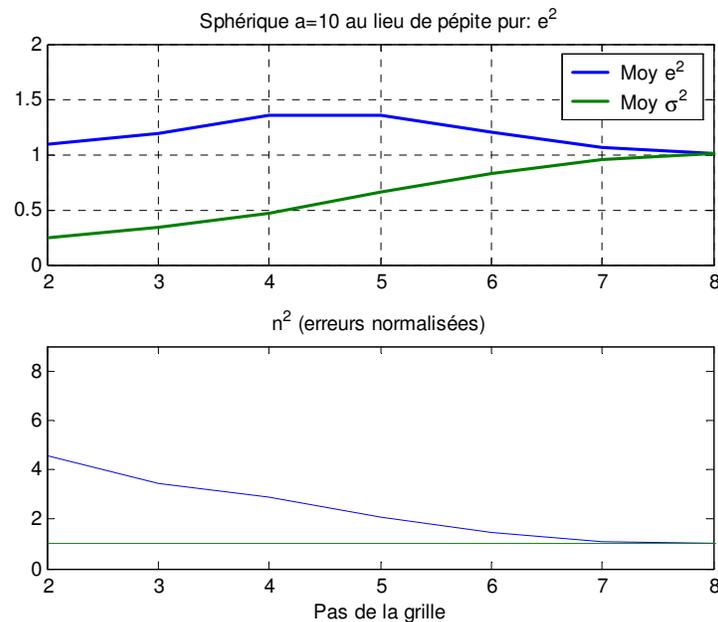
- pour la grille espacée (pas de 6 à 8), la structure est faible et la variance de krigeage prédit assez bien la précision obtenue;
- pour les grilles serrées (2 à 4), la variance de krigeage est supérieure à la variance des erreurs (vue pessimiste) ce qui résulte en une variance des erreurs normalisées inférieure à 1.

Figure 3 : On a fourni un modèle trop optimiste par rapport à la réalité de la simulation ($a=20$ au lieu de $a=10$)



On note que la variance de krigeage sous-estime la variance des erreurs (vue optimiste); la variance des erreurs normalisées est donc supérieure à 1.

Figure 4 : On a fourni un modèle trop optimiste par rapport à la réalité (sphérique avec $a=10$ fourni vs réalité : effet de pépité pur)



On note :

- la variance des erreurs augmente puis décroît en fonction du pas de la grille;
- bien qu'une grille au pas 4 représente 4 fois plus d'échantillons qu'une grille au pas 8, la variance des erreurs est supérieure. En spécifiant le mauvais modèle, on ne profite pas de l'information accrue disponible. Il peut donc être assez dangereux de fournir un modèle exagérément optimiste. L'explication pour la détérioration de la précision est que les poids de krigeage présentent un fort effet d'écran lorsqu'ils sont près du point à estimer (grille serrée) et que l'on se trouve donc à faire une moyenne sur quelques points seulement au lieu des 50 points lorsque la grille est assez espacée pour que les corrélations (et l'effet d'écran) soient faibles.
- la variance des erreurs normalisées est nettement supérieure à 1, indiquant que le modèle est exagérément optimiste

Autres mesures de validation

i. La variance expérimentale des teneurs (i.e. $\hat{\sigma}^2 = \frac{1}{n} \sum (Z_i - \bar{Z})^2$) devrait être égale à la variance de dispersion d'un point dans le gisement $D^2(\bullet|G)$.

ii. La relation de lissage du krigeage (voir section 5.3) fournit naturellement un outil de validation du modèle. Une fois le modèle fixé, on peut calculer les variances de bloc pour différentes tailles de bloc. On peut également réaliser les krigeages pour différentes tailles de bloc et calculer la variance expérimentale des valeurs krigées ($\hat{\sigma}_{Z_v}^*$) les moyennes des multiplicateurs de Lagrange ($\bar{\mu}$) et des variances de krigeage $\bar{\sigma}_{ko}^2$. On devrait alors avoir:

$$\hat{\sigma}_{Z_v}^{2*} \approx D^2(Z_v | G) - \bar{\sigma}_{ko}^2 - 2\bar{\mu}.$$

9.4.7 Exemple numérique de krigeage

Soit les points suivants:

$$\begin{array}{cccc} & x_1 & & \\ x_2 & & x_0 & \\ & & & x_3 \end{array}$$

$$x_1=(0,1)Z_1=9$$

$$x_2=(0,0)Z_2=3$$

$$x_3=(3,0)Z_3=4$$

On veut estimer le point x_0 situé à $(1,0)$. Supposons que l'on a un modèle sphérique, avec effet de pépité 1, palier 11 et portée 3. On calcule d'abord les distances entre toutes les paires de points:

h

	x_0	x_1	x_2	x_3
x_0	0	1.4	1	2
x_1	1.4	0	1	3.2
x_2	1	1	0	3
x_3	2	3.2	3	0

On évalue le variogramme sphérique à chacune de ces distances avec l'équation:

$$\gamma_h = \begin{cases} 0 & \text{si } h = 0 \\ 1 + 10 \left[1.5 \frac{h}{3} - 0.5 \left(\frac{h}{3} \right)^3 \right] & \text{si } 0 < h \leq 3 \\ 11 & \text{si } h > 3 \end{cases}$$

$\gamma(h)$

	x_0	x_1	x_2	x_3
x_0	0	7.55	5.81	9.52
x_1	7.55	0	5.81	11
x_2	5.81	5.81	0	11
x_3	9.52	11	11	0

On calcule la covariance correspondante

$$C(h)=11-\gamma(h)$$

	x_0	x_1	x_2	x_3
x_0	11.0	3.45	5.19	1.48
x_1	3.45	11	5.19	0
x_2	5.19	5.19	11	0
x_3	1.48	0	0	11

Ceci permet de construire le système de krigeage:

$$K \lambda = k_0$$

i.e., dans ce cas:

$$\begin{bmatrix} 11 & 5.19 & 0 & 1 \\ 5.19 & 11 & 0 & 1 \\ 0 & 0 & 11 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \mu \end{bmatrix} = \begin{bmatrix} 3.45 \\ 5.19 \\ 1.48 \\ 1 \end{bmatrix}$$

dont la solution est:

$$\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \mu \end{bmatrix} = \begin{bmatrix} .21 \\ .51 \\ .28 \\ -1.55 \end{bmatrix}$$

L'estimation est alors:

$$\Sigma \lambda_i Z_i = (.21)*9 + (.51)*3 + (.28)*4 = 4.54$$

La variance de krigeage est donnée par:

$$\sigma_{k_0}^2 = 11 - \lambda' k_0 = 8.76$$