Distribution multigaussienne et krigeage simple

Supposons que le processus soit multigaussien. Ceci veut dire que tout ensemble de n points possède une distribution multigaussienne. La fonction de densité conjointe de la loi multigaussienne est :

$$f(Z_1, Z_2, \ldots Z_n) = \frac{1}{(2\pi)^{n/2} \mid \Sigma^{-1} \mid^{0.5}} \exp\left[-(Z-m)'\Sigma^{-1}(Z-m)/2\right]$$

où

Σ est la matrice n x n des covariances
m est le vecteur n x 1 de $E[Z_i]$, i=1..n
Z est le vecteur n x 1 des « n » variables aléatoires

Une propriété remarquable de la loi multigaussienne est que toutes les lois conditionnelles sont également multigaussiennes. Ainsi, supposons que l'on a observé « q » parmi « n » valeurs. La distribution conditionnelle des « p=n-q » valeurs restantes est multinormale de moyenne conditionnelle :

$$(1) \qquad E\left[Z_p \mid Z_q\right] = \Sigma_{pq}\Sigma_{qq}^{-1}(Z_q - m_q) + m_p$$

et de variance conditionnelle :

$$(2) \qquad \Sigma_{pp|q} = \Sigma_{pp} - \Sigma_{pq}\Sigma_{qq}^{-1}\Sigma_{qp}$$

où

$$\Sigma = \begin{bmatrix} \Sigma_{pp} & \Sigma_{pq} \\ \Sigma_{qp} & \Sigma_{qq} \end{bmatrix}, \ m = \begin{bmatrix} m_p \\ m_q \end{bmatrix}, \ Z = \begin{bmatrix} Z_p \\ Z_q \end{bmatrix}$$

sont découpés de façon à regrouper les valeurs inconnues en tête.

Pour établir le lien avec le krigeage simple, supposons que l'on veuille déterminer la distribution conditionnelle en un seul point situé en x_0 conditionnellement aux valeurs prises en « n » points. On se place dans le contexte stationnaire, i.e. « m » est constant. On aura alors :

$$\Sigma_{qp} \equiv k_0 \quad \text{et} \quad \Sigma_{qq} = K$$

où k_0 et K sont les membres de droite et de gauche du système de krigeage simple.

Substituant dans (1) et (2), on trouve immédiatement que :

$$E\left[Z_p \mid Z_q\right] = k_0'K^{-1}(Z-m) + m = \lambda'(Z-m) + m = Z_{0,KS}^*$$
$$\Sigma_{pp|q} = \sigma^2 - \lambda' k_0 = \sigma_{KS}^2$$

Ce qui montre que, dans le cas multigaussien, le krigeage simple coïncide avec l'estimation de la distribution conditionnelle.

Connaissant la distribution de Z_0, on peut en déduire toute probabilité désirée, tout quantile, etc.

Généralisation 1 : cas ponctuel, distribution non gaussienne.

Toujours dans le cas ponctuel, que faire si le processus n'est pas multigaussien? L'idée qui vient naturellement est de transformer Z en une variable Y qui, elle, sera multigaussienne.

Problème : il n'est pas possible de définir une transformation sur un ensemble de variables aléatoires assurant que les variables ainsi transformées suivent une loi multigaussienne.

Par contre, il est très facile de transformer la loi marginale vers une loi gaussienne. Il est aussi possible de vérifier si la variable ainsi transformée est compatible avec une loi bigaussienne. Il n'est généralement pas possible d'aller plus loin dans les tests.

Transformation de la loi marginale (ou de l'histogramme)

Algorithme de transformation « graphique »:

- classer les observations par ordre croissant;
- associer à chaque valeur la valeur p=rang/(n+1), « n » est le nombre de valeurs connues;
- associer à chaque « p » la valeur N(0,1) correspondante tirée d'une table N(0,1) (en pratique évaluée numériquement par une formule d'approximation). Ceci définit la transformation Z=> Y, et Y est N(0,1) par construction.

Vérification du caractère plausible de la distribution bi-gaussienne :

- on calcule le variogramme (usuel) de Y $\gamma_2(h) = \frac{1}{2}E\left[(Z(x) - Z(x+h))^2\right]$;
- on calcule le variogramme d'ordre 1 de Y. Le variogramme d'ordre 1 est défini comme : $\gamma_1(h) = \frac{1}{2}E\left[| Y(x) - Y(x+h) |\right]$.
- on trace le graphe $\sqrt{\gamma_2(h)}$ en fonction de $\gamma_1(h)$; les points doivent définir une droite.

Extension au cas de blocs

Tout comme pour le krigeage d'indicatrices, le changement de support vient compliquer beaucoup les choses. Il existe un certain nombre de modèles pour tenir compte de ce changement de support (dont la correction affine), et tous basés sur un certain nombre d'hypothèses supplémentaires difficiles à vérifier.

Une méthode flexible consiste à représenter le bloc par une série de points le discrétisant (disons « nd ». On écrit alors :

$$P(Z_v(x) < c) = P(\tfrac{1}{v}\int_v Z(x)dx < c) \approx P(\tfrac{1}{nd}\sum_{di=1,nd} Z_{di} < c) = P(\tfrac{1}{nd}\sum_{di=1,nd} f^{-1}(Y_{di}) < c)$$

« di » indique un des points de discrétisation, f^{-1} représente la transformation inverse permettant de passer du domaine gaussien au domaine original. Pour évaluer le dernier terme, il faut simuler des vecteurs Y ayant la matrice de covariance conditionnelle et le vecteur de moyenne conditionnelle que l'on peut calculer à partir des résultats du krigeage simple effectué sur chacun des « m » points ou en appliquant directement les équations (1) et (2). Chacun de ces vecteurs Y est transformé en vecteur Z et l'on estime la probabilité recherchée par simple décompte des vecteurs rencontrant la condition sur le bloc.

En pratique cette approche ne montre pas de réel avantage par rapport à effectuer directement une simulation ponctuelle et elle n'est donc pas utilisée.

Cas particulier : krigeage lognormal

Dans plusieurs situations, la distribution de Y=log(Z) suit approximativement une loi gaussienne, ce qui permet certaines simplifications sous hypothèse que log(Z) est aussi multigaussien. Tout comme pour le cas général, le krigeage de Y spécifie les paramètres conditionnels de la distribution. Cette distribution est gaussienne (pour Y) de moyenne conditionnelle Y*(x) et de variance conditionnelle σ_k^2. Par conséquent Z(x)=exp(Y(x)) aura une distribution conditionnelle log-normale, de moyenne m_Z et variance σ_Z^2 données par les relations générales suivantes :

	Z	Y=ln(Z)
Moyenne	$m_Z = e^{m_Y+\sigma_Y^2/2}$	$m_Y = \ln(m_Z) - \sigma_Y^2/2$
Variance	$\sigma_Z^2 = m_Z^2(e^{\sigma_Y^2} - 1)$	$\sigma_Y^2 = \ln\left(\frac{\sigma_Z^2}{m_Z^2}+1\right)$

Note: Ces relations tiennent strictement entre les paramètres théoriques de la loi log-normale, ils ne tiennent qu'approximativement entre les valeurs estimées pour ces paramètres.

En remplaçant la moyenne et la variance de Y par la valeur krigée et la variance de krigeage, on obtient une estimation de l'espérance conditionnelle et de la variance conditionnelle pour Z. On notera qu'à $\sigma_Y^2 \equiv \sigma_k^2$ constant, la variance conditionnelle de Z croît avec l'espérance conditionnelle illustrant que l'estimation est moins précise aux fortes teneurs qu'aux basses teneurs.

On notera également que l'espérance conditionnelle de Z dépend de la variance de krigeage et donc du modèle de variogramme. Par conséquent, ici il est primordial de valider le modèle de variogramme de façon exhaustive afin d'éviter un fort biais possible.

On a aussi la relation suivante entre les covariances en fonction de la distance séparant deux points :

$$C_Z(h) = m_Z^2\left(\exp(C_Y(h)) - 1\right)$$

Cette relation permet de tester le caractère raisonnable de l'hypothèse de multi-lognormalité en comparant le modèle obtenu par la relation (i.e. $\gamma_Z(h) = \sigma_Z^2 - C_Z(h)$) avec les variogrammes expérimentaux de Z. On peut aussi s'en servir pour déduire le modèle de variogramme à adopter pour Z lorsque les données de Z sont très bruitées. Il est souvent préférable, si l'objectif est de simplement estimer la teneur moyenne, d'effectuer le krigeage sur Z directement quitte à avoir obtenu le variogramme de Z après étude du variogramme de Y et utilisation de la relation précédente.

Extension à des blocs

L'extension à des blocs est possible au prix d'une hypothèse additionnelle. On formule l'hypothèse que les distributions des teneurs de blocs sont également multi-lognormales. Bien que cette hypothèse ne puisse être vraie strictement lorsque les points sont multi-lognormaux, en pratique on observe souvent que les distributions marginales des teneurs de blocs demeurent approximativement log-normales lorsque les points sont log-normaux. Malgré cette hypothèse, le calcul du membre de droite des équations du krigeage devient plus compliqué et ne peut être obtenu par simple krigeage de blocs de Y(x) (voir Journel, 1980, The lognormal approach to predicting local distributions of selective mining unit grades, Mathematical geology, v. 12, # 4, 285-304).

Ressources récupérables globales

Si Z est lognormale de moyenne m et de variance logarithmique β^2, alors, le tonnage au-dessus de la teneur de coupure « c » est donné par:

$$T(c) = T_0 \, F\left(\frac{1}{\beta}\ln\left(\frac{m}{c}\right) - \frac{\beta}{2}\right)$$

où F(x) désigne la loi cumulative d'une N(0,1) évaluée à la valeur x, i.e. P(N(0,1) < x).
T_0 désigne le tonnage total du gisement (i.e. sans teneur de coupure)

De même, la quantité de métal obtenue à la teneur de coupure "c" est donnée par:

$$Q(c) = T_0 \, m\, F\left(\frac{1}{\beta}\ln\left(\frac{m}{c}\right) + \frac{\beta}{2}\right)$$

Exemple : Supposons que l'on connaisse m et σ^2 pour les unités de sélection déterminées, voyons quelle est l'influence du support de sélection.

Soit m = 3%
c = 4.5%
σ = 1.5%
T_o = 1 Mt

Si la distribution est lognormale, on a alors:

$$\beta = \left\{\ln\left(\frac{2.25\%^2}{9\%^2} + 1\right)\right\}^{1/2} = 0.472$$

et

$$T(4.5) = 1Mt\, F\left(\frac{1}{0.472}\ln\left(\frac{3}{4.5}\right) - \frac{0.472}{2}\right) = 1Mt\, F\,(-1.095) = 0.14\,Mt$$

$$Q(4.5) = 1Mt * 3\% * F\left(\frac{1}{0.472}\ln\left(\frac{3}{4.5}\right) + \frac{0.472}{2}\right) = 1Mt * 3\% * F(-0.623) = 0.80\ x\ 10^{-2}\ Mt$$

$$m(4.5) = \frac{0.80x10^{-2}\,Mt}{0.14\,Mt} = 5.71\%$$

Si l'exploitation impose la sélection sur de plus gros blocs ayant un écart-type de 1% (au lieu de 1.5% dans l'exemple précédent), alors, toujours en supposant une loi lognormale,

ß = 0.325

et

T(4.5) = .08 Mt
Q(4.5) = $.42\times10^{-2}$ Mt
M(4.5) = 5.25%

la quantité de métal est réduite de moitié!

Si on définit le profit par T(c) . (m(c) - c), on passerait de 0.169×10^{-2} Mt à $.06\times10^{-2}$ Mt donc une réduction de près des 2/3 du profit escompté.

Note : pour évaluer numériquement F(x), il existe nombre de formules d'approximation plus ou moins précises. Une des plus simples est la suivante:

$$F(x) \approx \frac{1}{2}\left[1+\left(1-e^{-\frac{2x^2}{\pi}}\right)^{\frac{1}{2}}\right] \quad si\ x \geq 0$$

$$F(x) \approx 1 - F(-x) si\ x < 0$$

L'écart maximal entre la vraie valeur et celle approchée par la formule précédente est de 0.003 à x=1.6.

Il existe aussi des formules d'approximation pour le calcul inverse, i.e. déterminer la valeur d'une N(0,1) correspondant à une probabilité donnée.

Exemple : avec la formule précédente, on obtient :

x	F(x) (formule)	F(x) table
0	0.5000	0.5000
1	0.8431	0.8413
2	0.9800	0.9772
3	0.9992	0.9987
4	1.0000	1.0000
