

NOUVEAUX RESULTATS SUR LES VITESSES DE CALCULS GEOSTATISTIQUES

Marco ALFARO
Planificacion Mina
Codelco Dicsion
CHUQUICAMATA
Chili

RESUME. Les gisements de nitrates-iodum dans le Nord du Chili ont une extension énorme, de l'ordre de 600 km² : l'estimation par krigeage est donc très lourde en temps de calcul. Cette circonstance nous a conduit à chercher des algorithmes de calcul de plus en plus rapides.

On présente une série de techniques pour réduire les temps des calculs géostatistiques par ordinateur, avec le but d'apporter une aide à l'utilisateur de la géostatistique sur site minier. Entre ces techniques on distingue une intéressante propriété d'isotropie du variogramme et l'ancien système d'équations de krigeage à N-1 inconnues.

TABLE DES MATIERES

1. INTRODUCTION
2. TECHNIQUES POUR AMELIORER LA VITESSE DES CALCULS GEOSTATISTIQUES
 - 2.1. Algorithme d'inclusion
 - 2.2. Algorithme pour le calcul des distances
 - 2.3. Voisinage d'estimation
 - 2.4. Le modèle de variogramme
 - 2.5. Le système de Matheron à N-1 équations
 - 2.6. Une propriété d'isotropie du variogramme

I. Introduction.

La figure 1 montre le contour d'un gisement de nitrates-iodum dans le Nord du Chili. Ces gisements sont reconnus par une maille irrégulière avec des puits a une profondeur de l'ordre de 2m. On mesure la puissance de découverte, la puissance mineralisée et les teneurs en NaNO_3 , I, Mg, Ca. La caractéristique principale de ces gisements est son extension: si on prend des blocs de 25mx25m, on peut arriver a un million de blocs : le krigeage de la zone prend un temps de calcul enorme dans un micro ordinateur du type PC. Par cette raison nous avons developé des techniques et algorithmes conduisant à des solutions plus économiques en temps de calcul.

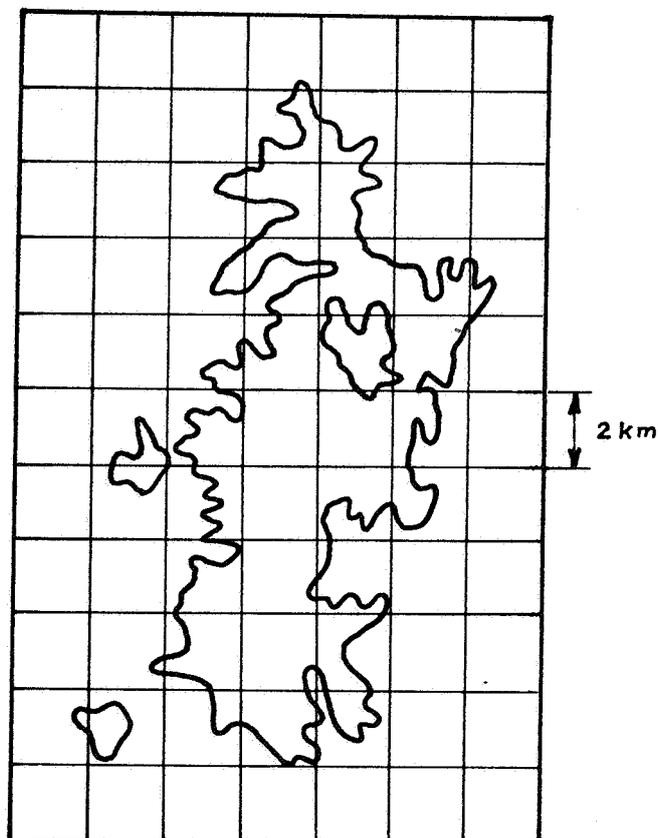


Figure 1: La Oficina de Pampa Yumbes

II. Techniques pour améliorer la vitesse des calculs géostatistiques.

II.1. Algorithme d'inclusion.

Soit S la zone de la figure 1: on veut savoir si un point (x,y) appartient ou non à S. Etant donné la complexité de la frontière de S, les algorithmes classiques (cf. M. David, 1988), sont trop lents. Notre méthode est très simple (on profite du fait que le remplissage des figures est inclus dans le langage de programmation):

i) dans l'écran de l'ordinateur on colore l'aire S:

fill(S) , couleur

ii) Ensuite on s'interroge par la couleur du point (x,y):

z = point(x,y)

Si z correspond au code de la couleur de S alors (x,y) appartient à S.

II.2. Algorithme pour le calcul des distances.

L'un des opérations les plus répétitives en géostatistique est le calcul des distances. La fonction du système sqrt() est trop lente. Pour le calcul des distances nous avons défini deux algorithmes plus rapides que:

$$d = \text{sqrt}(a * a + b * b)$$

méthode 1 (cf. Hall et Knight, 1957):

si $a > b$ alors

$$d = a + \frac{b^2}{2a + \frac{b^2}{2a + \frac{b^2}{2a}}}$$

pour avoir une meilleure précision on peut utiliser plus de termes dans la fraction continue.

méthode 2 (cf. Moler et Morrison, 1983):

cette méthode s'explique par le programme "C" suivant (il faut prendre k petit) :

```

a = abs( a );
b = abs( b );
if ( a < b ) changer( a,b );
if ( a==0. ) return b;
for( i = 1 ; i < = k ; i++ )
{
  r = b / a;
  r = r * r;
  r = r / ( 4. + r );
  a = a * ( 1. + r + r );
  b = b * r;
}
return a;

```

l'idéal est de programmer ces algorithmes en langage machine.

II.3. Voisinage d'estimation.

La méthode classique consiste en un voisinage circulaire avec le test suivant:

```

if ( x * x + y * y > r2 ) ... ( r2 = r2 )

```

Cependant il est plus rapide d'utiliser l'algorithme suivant ($r_1 = 1.4142r$):

```

if ( x > r ) ...
if ( x < -r ) ...
if ( y > r ) ...
if ( y < -r ) ...
if ( abs(x) + abs(y) > r1 ) ...

```

cette méthode n'a pas des multiplications et est un voisinage octogonal (cf. figure 2).

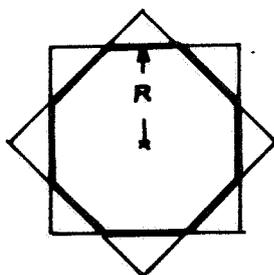


Figure 2

Les algorithmes antérieures peuvent être généralisés à l'espace à trois dimensions.

Maintenant nous étudierons des techniques plus géostatistiques.

II.4. Le modèle de variogramme.

En géostatistique on utilise souvent le modèle sphérique d'équation ($h = h / a$) :

$$\gamma(h) = C (1.5h - 0.5h^3) \quad \text{si } h \leq 1$$

Mais nous préférons le modèle quadratique d'équation (cf. M. Alfaro, 1983) :

$$\gamma(h) = C (2h - h^2) \quad \text{si } h \leq 1$$

Pour voir l'avantage de ce modèle sur le modèle sphérique, étudions les deux programmes suivants :

Programme 1

```
n = 10000;
for(i=1;i<=n;i++)
{
    h = (float)i/n;
    g = h*(2.-h);
}
```

Programme 2

```
n = 10000
for(i=1;i<=n;i++)
{
    h = (float)i/n;
    g = 0.5*h*(3.-h*h);
}
```

Pour le modèle sphérique on a besoin de 3 multiplications contre 1 du modèle quadratique. Le programme 1 est de l'ordre d'un 30% plus rapide que le programme 2.

Dans un gisement avec des millions de blocs le nombre de fois qu'on calcule la fonction $\gamma(h)$ est gigantesque, donc il vaut mieux d'utiliser le modèle le plus simple possible (évidemment de type positif), c'est à dire le modèle quadratique.

II.5. Le système de Matheron a N-1 équations.

Dans les logicielles de géostatistique on utilise toujours le système de $N + 1$ équations suivant (krigeage ordinaire) :

$$\sum_{j=1}^N \lambda_j \sigma_{ij} + \mu = \sigma_{iv} \quad i = 1, 2, \dots, N$$

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

$$\sigma_k^2 = \sigma_{vv} - \sum_{i=1}^N \lambda_i \sigma_{iv} - \mu$$

Si on élimine le paramètre μ et λ_N dans le système antérieur

on a (cf. G. Matheron, 1969):

$$\text{où: } \sum_{j=1}^{N-1} \lambda_j R_{ij} = N_i \quad i = 1, 2, \dots, N-1$$

$$R_{ij} = \sigma_{ij} - \sigma_{Nj} - \sigma_{Ni} + \sigma_{NN}$$

$$N_i = \sigma_{iv} - \sigma_{NV} - \sigma_{Ni} + \sigma_{NN}$$

$$\sigma_k^2 = \sigma_{VV} - \sigma_{NV} - \sigma_{NV} + \sigma_{NN} - \sum_{i=1}^{N-1} \lambda_i N_i$$

On observe que les termes qu'il faut calculer avant de résoudre le système sont les mêmes dans les deux cas, avec l'avantage que le deuxième système a deux équations de moins.

Dans la figure 3 on a les temps d'ordinateur pour les deux systèmes: le gain du dernier système peut être considérable.

D'autre part, le système classique de krigeage universel peut être écrit de la même façon que le dernier système (cf. A. Maréchal, 1973).

Remarque : les temps antérieures peuvent changer d'ordinateur à ordinateur.

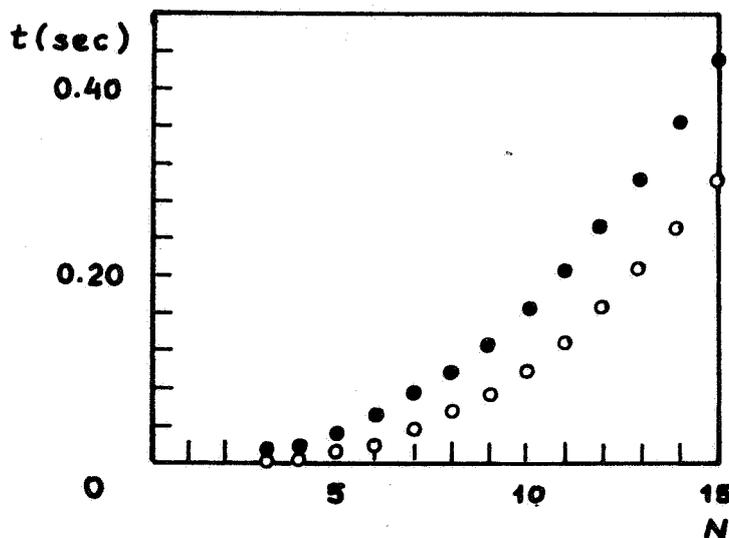


Figure 3

II.6. Une propriété d'isotropie du variogramme.

Si on krige un gisement, l'opération qui consomme le plus de temps d'ordinateur est le calcul des intégrales du second membre:

$$\bar{y} = \frac{1}{V} \int_V y(x_i - x) dx \quad (1)$$

L'intégral est approchée par discretization (avec un minimum de 36 points dans \mathbb{R}^2 et 64 points dans \mathbb{R}^3 , cf Journal et Huij-

bregts 1978).

Une façon d'aller plus vite est de trouver l'expression explicite de l'intégral (cf. M. Alfaro, 1988) et éviter les sommes discrètes. Cependant, en général, l'expression de l'intégral est trop compliquée (avec des termes logarithmiques et des fonctions trigonométriques inverses) et il est presque impossible de trouver l'expression dans l'espace \mathbb{R}^3 .

L'approximation de l'intégral (1) que nous proposons est basée dans le fait suivant:

Soit $Z(x)$ une fonction aléatoire stationnaire isotrope. Soit un bloc carré de côté l et un point x_0 (figure 4) :

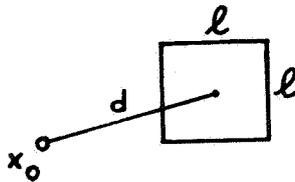


Figure 4

L'approximation consiste à observer que la valeur moyenne (1) dépend, de façon approchée, seulement du modèle de variogramme et de la distance d entre le point et le centre de gravité du bloc.

$$\frac{1}{V} \int_V \gamma(x_0 - x) dx \cong \mathcal{G}(d) \quad (2)$$

Pour voir la qualité de l'approximation (2), soit un modèle sphérique de palier $C = 1$ et portée $a = 2l$ et considérons les 15 points de la figure 5:

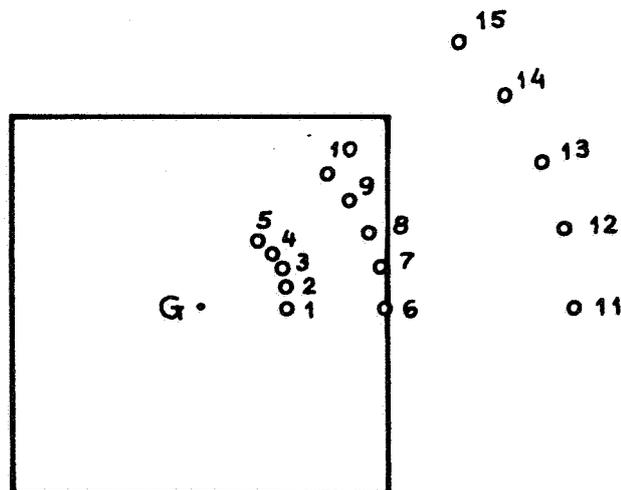


Figure 5

Le tableau suivant résume les valeurs de l'intégral (1) pour les 15 points:

point	distance	\bar{y}
1	0.25	0.3186
2	0.25	0.3183
3	0.25	0.3185
4	0.25	0.3184
5	0.25	0.3184
6	0.50	0.4241
7	0.50	0.4235
8	0.50	0.4223
9	0.50	0.4211
10	0.50	0.4205
11	1.00	0.6962
12	1.00	0.6961
13	1.00	0.6955
14	1.00	0.6947
15	1.00	0.6944

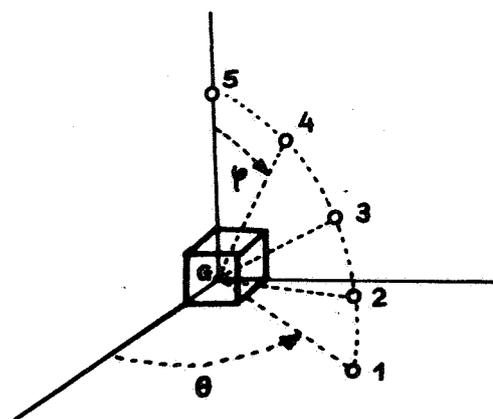
La fonction $\mathcal{C}(d)$ peut être modélisée, par exemple, dans le cas antérieur on a:

$$\mathcal{C}(d) = 1 - 0.7189 \exp(-u^2)$$

$$\text{ou } u = d / (1.08 \ell)$$

Finalement on introduit la fonction $\mathcal{C}(d)$ dans le système de krigeage.

L'approximation est encore meilleure dans l'espace à trois dimensions: soit un bloc $\ell \times \ell \times \ell$, un modèle sphérique de palier 1 et portée 4ℓ (cf. figure 6 et le tableau suivant):



$$\theta = 67.5^\circ$$

$$\varphi = 22.5^\circ$$

Figure 6

point	\bar{y}	
1	0.6952	
2	0.6951	
3	0.6952	(d = 2l)
4	0.6953	
5	0.6953	

Nous avons constaté que la propriété antérieure est valable pour les autres modèles de variogrammes. D'autre part, le krigeage est si robuste que l'approximation (2) n'a presque pas d'influence sur les pondérateurs et les estimateurs de krigeage.

L'approximation est aussi applicable dans certains cas d'anisotropie ou à des blocs rectangulaires (cf. figure 7 où $\bar{y} = (\mathcal{G}(d_1) + \mathcal{G}(d_2)) / 2$)

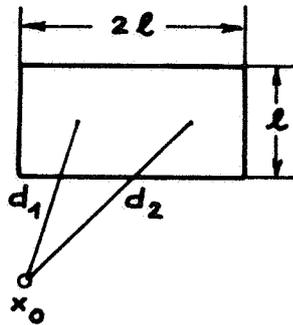


Figure 7

La technique qu'on vient d'étudier peut avoir d'autres applications, par exemple:

- dans d'autres types de krigeage (universel, disjonctif, indicatrices, ...).
- dans la solution approchée du krigeage de la réunion de n blocs et calcul de la variance globale d'estimation.

Soit:

$$s = \bigcup_{i=1}^n v_i \quad , \text{ alors (cf figure 8) :}$$

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n \mathcal{G}(d_i)$$

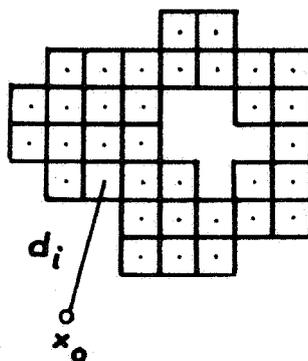


Figure 8.

III. Conclusions

Ces techniques nous ont permis de réduire le temps de calcul du krigeage d'un jour à quelques heures. Le plus important est que les méthodes que nous avons introduit, n'entraînent pas une perte de précision.

IV. Références

- M. Alfaro 1983 Statistical Inference of the semivariogram and the quadratic model. Dans G. Verly et autres: Geostatistics for Natural Resources Characterization. Ed. Reidel.
- M. Alfaro 1988 Estimation non classique d'un gisement d'or. III Congrès International de Géostatistique. Avignon France.
- M. David 1988 Handbook of Applied and Advanced Geostatistics. Ed. Elsevier 1988.
- Hall et Knight 1957 Higher Algebra. Macmillan, Londres.

-
- A. Journel Mining Geostatistics. Ed. Academic Press. New
Ch. Huijbregts York.
1978
- A. Maréchal El sistema de Matheron en krigeage Universal.
1973 Boletín de Geoestadística, Nº4. Ed. Depto. de
Minas, Universidad de Chile.
- Moler et Replacing Square Roots by Pythagorean Sums. IBM
Morrison Journal of Research and Development 27,6.
1983
- G. Matheron Cours de Géostatistique. Les Cahiers du Centre
1969 de Morphologie Mathématique de Fontainebleau,
Fascicule 2. Ed. Ecole des Mines de Paris.