

Los Cuadernos del Centro de Morfología Matemática
de Fontainebleau

Fascículo 2

CURSO DE GEOESTADÍSTICA

Por

Georges MATHERON

1969

Traducido al español por

Marco ALFARO

2005

Propiedad del Centro de Geoestadística de la Escuela de Minas de París.

Prohibida su venta.

CURSO DE GEOESTADISTICA

Tabla de Materias

PREFACIO 3

0 - INTRODUCCION. 4

- 0-1 Notaciones 4
- 0-2 Producto de convolución 5
- 0-3 Geoestadística y Teoría de las variables regionalizadas 6
- 0-4 Métodos transitivos y Teoría intrínseca 10

1 - LOS METODOS TRANSITIVOS. 9

- 1-1 Ejemplo introductorio 9
- 1-2 El covariograma transitivo 11
- 1-3 Regularización y subida 12
- 1-4 La estimación de las variables regionalizadas 13
- 1-5 Aplicación a la estimación de una superficie S 17

2 - TEORIA DE LAS F. A. INTRINSECAS. 20

- 2-1 Definiciones generales 20
- 2-2 Propiedades de la covarianza y del variograma 23
- 2-3 Regularización de una F.A.I. 26
- 2-4 Varianzas de extensión y de estimación 29
- 2-5 Efecto de pepita y fenómeno de transición 34
- 2-6 Calculo de varianzas de estimación 37

3 - EL ESQUEMA ESFERICO. 45

4 - EL KRIGEADO 47

- 4-1 El problema del krigeado y su interés 47
- 4-2 Las ecuaciones generales del krigeado 49
- 4-3 Krigeado completo y balance mineral metal 51

5 - INVESTIGACION DEL OPTIMO EN EL RECONOCIMIENTO Y LA PUESTA EN EXPLOTACION DE DEPOSITOS MINEROS: 53

- 5-0 Introducción 53
- 5-1 Elección de una ley de corte y de un ritmo anual de explotación 56
- 5-2 Reconocimiento óptimo para el dimensionamiento de la explotación 60
- 5-3 Optimización secuencial y término del reconocimiento 63

6. EJERCICIOS 70

BIBLIOGRAFIA 77

PREFACIO DEL TRADUCTOR

Una vez finalizada la traducción del Fascículo 5, "La Théorie des Variables Régionalisées, et ses Applications" del Maestro Georges Matheron, publicada en 1970, un amigo geoestadístico me solicitó traducir "Cours de Géostatistique", el cual a su juicio era un subconjunto del Fascículo 5.

Al realizar la traducción me di cuenta de varias diferencias, por ejemplo, en el tratamiento del krigeado, en el cual se llega a un sistema de ecuaciones de $n-1$ incógnitas, siendo n el número de datos (esto se conoció, en el pasado, como sistema de Matheron. Debido a que el tiempo computacional para resolver un sistema es proporcional a n^2 , yo he utilizado este sistema en mis programas computacionales, construyendo un programa más rápido que los otros).

La diferencia más notable es que en el Fascículo 5 no existe un capítulo acerca de los problemas de economía minera. En este curso de geoestadística se estudian los problemas económicos para determinar el ritmo de producción y la ley de corte óptimos. Matheron llega a la conclusión que para determinar la ley de corte y el ritmo de producción que maximizan los beneficios futuros de la explotación de un depósito, se debe utilizar una tasa de actualización $i=0$, en abierta contradicción con la tendencia actual, la cual conduce a un "floreo técnico" de nuestros depósitos.

Además de rendirle un homenaje de reconocimiento y de afecto al creador de la Geoestadística, con la satisfacción de haber sido su Secretario en la efímera Asociación Internacional de Geoestadística, el propósito de esta publicación es doble: por una parte servirá de ayuda y complemento para los estudiantes de minas, de geología y los profesionales interesados en la estimación de recursos mineros, por otra parte, como una referencia para mostrar que ciertos conceptos y el uso de algunas herramientas han sido desvirtuados con el correr de los años. La varianza de estimación, noción capital en todo el libro, constituye un excelente ejemplo.

La traducción es casi textual pero he omitido - en lo cual G. Matheron estaba de acuerdo - el modelo de De Wijs. La palabra "semivariograma" es reemplazada siempre por "variograma". En vez del anglicismo "kriging" he preferido traducir "krigeage" por la palabra correcta en español "krigeado" o "krigeaje" (se prefiere krigeado porque krigeaje no suena bien en castellano).

Existen otros manuscritos que he realizado, que iré poco a poco pasando a formato electrónico, poniéndolos a disposición de la comunidad minera.

M. ALFARO

CURSO DE GEOESTADISTICA

0 - INTRODUCCION

Designaremos por x, y, \dots puntos del espacio de n dimensiones ($n = 1, 2$ o 3), por dx, dy, \dots elementos de longitud ($n = 1$), de superficie ($n = 2$) o de volúmenes ($n = 3$), centrados en los mismos puntos, por $f(x), g(y)$ etc., ... funciones de estos puntos.

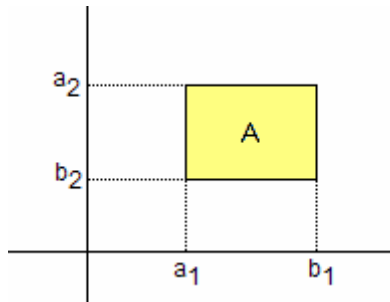
La integral (extendida a todo el espacio) de una función $f(x)$ se escribe:

$$\int f(x)dx$$

Por ejemplo, para $n = 3$ dimensiones, y, designando por (x_1, x_2, x_3) las tres coordenadas del punto x , esta notación se explicita de la manera siguiente:

$$\int f(x)dx = \int_{-\infty}^{+\infty} dx_1 \int_{-\infty}^{+\infty} dx_2 \int_{-\infty}^{+\infty} f(x_1, x_2, x_3) dx_1 dx_2 dx_3$$

De la misma manera, escribiremos $\int_A f(x)dx$ para la integral de la función $f(x)$ sobre un dominio A del espacio de n dimensiones.



Para $n = 2$, si A es el rectángulo de la figura, se tiene, de manera explícita:

$$\int_A f(x)dx = \int_{a_1}^{b_1} dx_1 \int_{a_2}^{b_2} f(x_1, x_2) dx_1 dx_2$$

Estas notaciones son a la vez muy condensadas y muy parlantes: en una primera lectura se las puede interpretar en el espacio de una dimensión, donde su significación es, en general, muy clara. La generalización a los espacios de dos o tres dimensiones se hace así más fácil.

Ejemplo: Sea V un volumen, y $g(h) = g(h_1, h_2, h_3)$ una función del vector h de coordenadas h_1, h_2, h_3 . Sean $x = (x_1, x_2, x_3)$ e $y = (y_1, y_2, y_3)$ el origen y la extremidad del vector h , es decir $h=y-x$. El valor medio de la

función $g(h)$ cuando los dos puntos x e y recorren (cada uno a su manera) el volumen V , se escribe:

$$\frac{1}{v^2} \int_v dx \int_v g(y-x) dy$$

En notación explícita, lo anterior se escribiría como:

$$\frac{1}{v^2} \int_v \int_v \int_v dx_1 dx_2 dx_3 \int_v \int_v \int_v g(y_1 - x_1; y_2 - x_2; y_3 - x_3) dy_1 dy_2 dy_3$$

La primera notación es un símbolo, que describe directamente el concepto de valor medio; la segunda es un algoritmo que indica el camino a seguir para calcular esta cantidad, pero cuya significación no aparece a primera vista.

0-2 PRODUCTO DE CONVOLUCION (media móvil)

El producto de convolución de dos funciones $f_1(x)$ y $f_2(x)$ se define por:

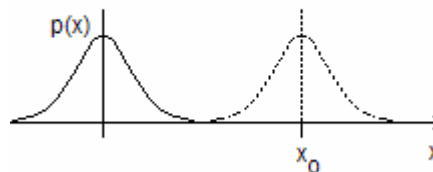
$$g(x) = \int f_1(y) f_2(x-y) dy = \int f_2(y) f_1(x-y) dy$$

Se escribe, en notación simbólica:

$$g = f_1 * f_2$$

Esta operación de convolución juega un papel fundamental en Geoestadística (como también en probabilidades y en toda la física teórica). La convolución se puede asociar con la noción intuitiva de "media móvil":

Regularizada de una función f (media móvil ponderada de la función)



Sea $p(y)$ una función de ponderación. El valor en el punto x_0 de la media móvil de la función f ponderada por p es:

$$f_p(x_0) = \int p(y) f(x_0 + y) dy = \int p(-y) f(x_0 - y) dy$$

(se atribuye el peso $p(y)dy$ al valor tomado por f en el punto $x_0 + y$. La segunda expresión se obtiene al cambiar y por $-y$).

Sea \tilde{p} la función transpuesta de la función p (por definición: $\tilde{p}(x) = p(-x)$). $f_p(x_0)$ es el valor en x_0 de $f * \tilde{p}$:

$$f_p = f * \tilde{p}$$

Esta media móvil f_p de la función f (ponderada por p) se llama regularizada (de f por p).

Ejemplo 1 (toma de una muestra v en el punto x)

Sea v la muestra implantada en el origen de coordenadas, y sea $k(x)$ su función indicatriz (por definición $k(x) = 1$ si $x \in v$ y $k(x) = 0$ si $x \notin v$). Tomemos como función de ponderación $p(x) = (1/v)k(x)$. La media móvil correspondiente:

$$f_v = \frac{1}{v} f * \tilde{k}$$

representa la ley media de la muestra v tomada en el punto x . De manera explícita, se tiene:

$$f_v(x) = \frac{1}{v} \int_v f(x+y) dy$$

Ejemplo 2 (radiactividades)

Si en el origen 0 de coordenadas se pone una masa unitaria de una sustancia radiactiva, se observa a la distancia d una radiactividad igual a $Ae^{-\lambda d}/d$ (A es una constante y λ es el coeficiente de absorción del medio). Sea ahora $f(x)$ la ley en el punto x de esta sustancia radiactiva. Designemos por $d(x-x_0)$ la distancia entre dos puntos x y x_0 . La masa $f(x)dx$ localizada en x genera en x_0 una radiactividad:

$$Ae^{-\lambda d(x-x_0)} / d(x-x_0) f(x) dx$$

En total, se observa en x la radiactividad:

$$A \int f(x) \frac{e^{-\lambda d(x-x_0)}}{d(x-x_0)} dx$$

Esto no es otra cosa que el producto de convolución $Af * \tilde{p}$ (con una función p de ponderación $p(x) = e^{-\lambda d}/d$, la cual es además simétrica, es decir $p = \tilde{p}$)

0-3 GEOSTADISTICA Y TEORIA DE LAS VARIABLES REGIONALIZADAS.

La Geoestadística es la aplicación de la teoría de las variables regionalizadas a la estimación de los depósitos mineros (con todas las aproximaciones que esto implica). De manera general, diremos que un

fenómeno es regionalizado cuando se desplaza en el espacio, manifestando una cierta estructura. Las ciencias de la tierra, entre otras, nos proporcionan numerosos ejemplos. Si $f(x)$ designa el valor en el punto x de una característica f de este fenómeno, diremos que $f(x)$ es una variable regionalizada, abreviado, una V.R. Se trata de un término neutro, descriptivo, anterior, en particular a toda interpretación probabilística.

Entonces, del punto de vista matemático, una V.R. es simplemente una función $f(x)$ del punto x , pero es, en general, una función muy irregular: ejemplo: una ley en un depósito minero. Una variable regionalizada se presenta bajo dos aspectos contradictorios (o complementarios):

- un aspecto aleatorio (alta irregularidad, y variaciones imprevisibles de un punto a otro)
- un aspecto estructurado (la V.R. debe sin embargo reflejar a su manera las características estructurales de un fenómeno regionalizado)

La teoría de las V.R. se propone entonces dos objetivos principales:

- en el plano teórico, expresar estas características estructurales en una forma matemática adecuada
- en el plano práctico, resolver el problema de la estimación de una V.R. a partir de un muestreo fragmentario.

Estos dos objetivos están relacionados: para un mismo conjunto de muestras, el error de estimación depende de las características estructurales; este error, por ejemplo, es más grande cuando la V.R. es más irregular y más discontinua en su variación espacial.

Campo y Soporte de una V.R.

El campo V de una V.R. es el dominio donde ésta es diferente de 0. Un panel es un subconjunto V' de V .

Soporte - A menudo, no se conoce $f(x)$, sino más bien su valor medio $f_v(x)$ dentro de la muestra v tomada en el punto x . Esta regularizada $f_v(x)$ es efectivamente más regular que la V.R. $f(x)$. El volumen v se llama soporte de la V.R. f_v , regularizada de f . Otra tarea importante de la teoría de las variables regionalizadas consistirá entonces en determinar las características de f_v conociendo las de f . Ejemplo: en un depósito, prever las características de los paneles V' (variable f_v), conociendo las características de f o de f_v (muestras).

Más generalmente, se buscará relacionar las características de f con las de una regularizada f_p dada (ejemplo de las radiactividades).

0-4 METODOS TRANSITIVOS Y TEORIA INTRINSECA.

Para alcanzar los objetivos disponemos de dos grupos de métodos:

- métodos transitivos: absolutamente generales, en particular no necesitan ninguna hipótesis de naturaleza probabilística, y, a fortiori, ninguna hipótesis de estacionaridad.

- teoría intrínseca: se trata de una aplicación de la teoría de las funciones aleatorias; entonces se introducen interpretaciones probabilísticas, y además, una cierta hipótesis de estacionaridad (hipótesis intrínseca)

Desde el punto de vista teórico, estos dos grupos de métodos conducen a resultados equivalentes, lo cual muestra que los resultados de la teoría intrínseca no están, en realidad, ligados a la hipótesis de estacionaridad (por otra parte, se puede construir una teoría probabilística sin utilizar esta hipótesis, la cual permite encontrar los resultados principales de la Geoestadística).

En la práctica la teoría intrínseca es más fácil de construir, y es casi siempre la teoría que se utiliza, salvo el caso particular, muy importante, de la estimación de una superficie o de un volumen (problema geométrico)

Respecto de la bibliografía (ver referencias al término de este Fascículo), se encontrará en [4] la exposición completa de la teoría de las variables regionalizadas, y en [5], una exposición abreviada de esta misma teoría, además de un tratado completo de Geoestadística aplicada, conteniendo, en particular, el estudio de los problemas de optimización económica (el texto original en francés puede ser consultado en la Biblioteca de la Escuela de Minas de París). La tesis de J. Serra proporciona, por una parte una exposición general muy concreta (sin dificultades matemáticas), y por otra parte, un estudio detallado del esquema esférico. El antiguo tratado [3] está obsoleto en gran parte, salvo en lo relacionado con el esquema de De Wijs. La tesis de A. Carlier [1] estudia (solamente para el esquema de De Wijs) los problemas especiales propuestos por los depósitos de substancias radiactivas. Los problemas de optimización económica se tratan en dos artículos de los Annales des Mines [8], lo más esencial figura en [5].

1 - LOS METODOS TRANSITIVOS

1-1 EJEMPLO INTRODUCTORIO

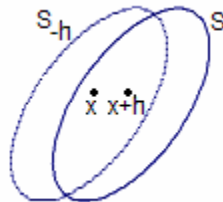
Consideremos el fenómeno de "transición" más simple que se pueda imaginar: presencia o ausencia de una característica. Sea por ejemplo una formación geológica de extensión limitada. Un sondaje perforado en el punto x la intersecta o no la intersecta. Designemos por $k(x)$ la indicatriz del conjunto S , es decir la función definida por:

$$k(x) = \begin{cases} 1 & \text{si } x \in S \\ 0 & \text{si } x \notin S \end{cases}$$

Se trata de un fenómeno único, para el cual no es posible realizar una formulación probabilística: hablar de la probabilidad de que un punto dado x pertenezca a S no tendría gran sentido. Se puede observar también que todo el interés se concentra aquí en la frontera de S . En efecto, $k(x)$ es constante tanto al interior como al exterior de S , solamente al pasar esta frontera $k(x)$ varía, pasando de 1 a 0 o de 0 a 1. Esta es la justificación del nombre de fenómeno de transición, y del nombre métodos transitivos.

El área S de nuestra formación está dada, evidentemente, por:

$$S = \int k(x) dx$$



El valor S constituye una información muy interesante desde el punto de vista práctico: a menudo, se busca estimar este valor a partir de una red de sondajes. Sin embargo este parámetro no nos aporta ninguna información de naturaleza realmente estructural. En efecto, se puede definir la estructura de un conjunto como el sistema de relaciones existentes entre los elementos o las partes de este conjunto. Tendremos información de naturaleza estructural al hacer intervenir simultáneamente, al menos, dos puntos.

Sean entonces dos puntos x y $x+h$ (es decir el conjunto estructurante más pequeño que podamos imaginar. Consideremos entonces la expresión $k(x)k(x+h)$: vale 1 si x y $x+h$ pertenecen los dos a S , y 0 en otro caso. Pero decir que $x+h$ pertenece a S equivale a decir que x pertenece al trasladado S_{-h} de S según la traslación $-h$. Se tiene entonces:

$$k(x)k(x+h) = \begin{cases} 1 & \text{si } x \in S \cap S_{-h} \\ 0 & \text{si } x \notin S \cap S_{-h} \end{cases}$$

Integremos en x esta expresión: obtenemos una función de h:

$$K(h) = \int k(x)k(x+h)dx = \text{medida}(S \cap S_{-h})$$

Que representa la medida (el área) de la intersección de S y de su trasladado por -h. Esta función es simétrica, porque las dos intersecciones $S \cap S_{-h}$ y $S \cap S_h$ se deducen una de la otra por traslación. Esta función K(h) es el covariograma geométrico asociado a S. La función K(h) proporciona una cierta imagen de la forma del conjunto S:

Propiedades del covariograma geométrico K(h)

a/ Simetría: $K(h) = K(-h)$

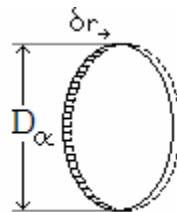
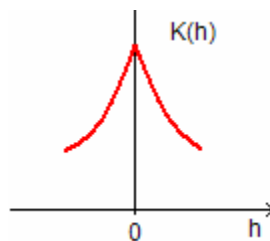
Desigualdades: $0 \leq K(h) \leq K(0)$

Relaciones: $S = K(0)$

$$S^2 = \int K(h)dh$$

b/ Alcances: el alcance $a(\alpha)$ en la dirección α es la distancia a partir de la cual K(h) se anula en esa dirección. Entonces es la dimensión más grande de S en esta dirección.

c/ Pendiente en el origen (derivada a la derecha y a la izquierda)



Si el módulo δr de h es pequeño, se tiene:

$$K(h) = K(0) - \delta r D_\alpha$$

$\delta r D_\alpha$ es la mitad de la superficie pequeña barrida por el vector δr cuyo origen describe el contorno de S. D_α es la variación diametral de S en la dirección α (si S es convexo, D_α es el diámetro aparente en esa dirección).

1-2 EL COVARIOGRAMA TRANSITIVO (caso general)

Sea $f(x)$ una V.R. nula fuera de un campo V acotado. El covariograma transitivo de esta V.R. es la función $g(h)$ definida por:

$$(1) \quad g(h) = \int f(x)f(x+h)dx$$

Propiedades del covariograma transitivo $g(h)$

a/ Simetría:
$$g(h) = g(-h)$$

desigualdad:
$$|g(h)| \leq g(0) = \int [f(x)]^2 dx$$

Sea $Q = \int f(x)dx$ la cantidad de metal, se tiene:

$$(2) \quad Q^2 = \int g(h)dh$$

(para demostrar (2), reemplazar g por su expresión (1), e integrar en h)

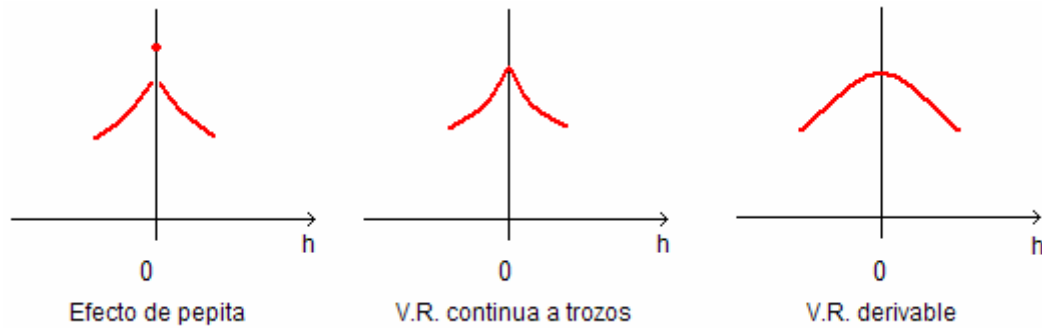
b/ Alcance: $a(\alpha)$ se define por la condición $g(h)=0$ cuando $h>a(\alpha)$ para el vector h de dirección α : es una propiedad del campo de la V.R.

c/ Comportamiento de $g(h)$ en una vecindad del origen. La regularidad de $g(h)$ en una vecindad del origen refleja las propiedades de continuidad de la V.R. en su variación espacial. Lo anterior resulta de:

$$g(0) - g(h) = \frac{1}{2} \int [f(x+h) - f(x)]^2 dx$$

Si $f(x)$ es continua por trozos, $g(h)$ tiene un comportamiento lineal en una vecindad del origen.

Si $f(x)$ es derivable, $g(h)$ tiene un comportamiento parabólico:



A menudo, $g(h)$ no es continuo en $h=0$: dicho de otra manera, $g(0)$ es mayor que el límite de $g(h)$ cuando h tiende a 0. Se dice entonces que hay un efecto de pepita. Más que una discontinuidad verdadera, se trata, frecuentemente, de una zona de transición muy rápida, la cual, experimentalmente, se presenta como una discontinuidad.

Caso isótropo. Si $g(h) = g(r)$ solo depende del módulo $r = |h|$ del vector h , su desarrollo en una vecindad del origen contiene una parte regular (términos de grado par) y una parte irregular (términos en r^λ , λ diferente de un entero par, eventualmente con términos logarítmicos $r^{2k} \log(r)$):

$$g(r) = g(0) + a_2 r^2 + \dots + \underbrace{\sum_{\lambda} a_{\lambda} r^{\lambda}}_{\text{parte regular}} + \underbrace{\sum_k a_{2k} r^{2k} \log(r)}_{\text{parte irregular}}$$

El término de más bajo grado de la parte irregular caracteriza la irregularidad de la V. R. en su variación espacial. Por ejemplo, la V. R. geométrica asociada a una superficie S tiene un covariograma en:

$$K(h) = S - D|h| + \dots$$

el término de más bajo grado de la parte irregular es de grado 1.

1-3 REGULARIZACION DE UNA V.R. Y SUBIDA

1-3-1. Regularización de una V.R.

Sea $f(x)$ una V.R., $p(x)$ una función de ponderación, $f_p = f * \tilde{p}$ la regularizada de f por p . La relación (1) puede escribirse:

$$g = f * \tilde{f}$$

lo cual muestra que el covariograma transitivo es la autoregularizada de la V.R. f . El covariograma

$$g_p = f * \tilde{p} * \tilde{f} * p = f * \tilde{f} * p * \tilde{p} = f_p * \tilde{f}_p$$

se puede poner en la forma:

$$g_p = g * P$$

con $P = p * \bar{p}$: entonces el covariograma de la regularizada se obtiene regularizando g con el covariograma transitivo P de la función de ponderación p: g_p es una función más regular que g, así como f_p es más regular que la V.R. inicial f.

La Subida (caso límite de la regularización). Si $f_3(x) = f_3(x_1, x_2, x_3)$ es una V.R. en el espacio de tres dimensiones, se dice que la V.R.:

$$f_2(x_1, x_2) = \int f_3(x_1, x_2, x_3) dx_3$$

definida en el espacio de dos dimensiones se deduce de f_3 por subida. Por ejemplo, si $f_3(x)$ es una ley puntual, $f_2(x_1, x_2)$ es la acumulación (cantidad de metal al metro cuadrado) del sondaje implantado en el punto (x_1, x_2) de la superficie topográfica.

No hay dificultad para definir subidas de orden superior. Por ejemplo, la subida de orden 2 (paralela al plano x_1, x_2) conduce a la V.R.

$$f_1(x_3) = \iint f_3(x_1, x_2, x_3) dx_1 dx_2$$

Se demuestra (por ejemplo, con la ayuda de la transformada de Fourier), que el covariograma $g_2(h_1, h_2)$ de la V.R. f_2 deducida de f_3 por subida de orden 1 se obtiene efectuando directamente esta misma operación de subida sobre el covariograma $g_3(h_1, h_2, h_3)$ de la V.R. f_3 .

Subida en el caso isótropo

Si $g_3(h) = g_3(r)$ solo depende del módulo r del vector h, se puede efectuar la subida término a término sobre la parte irregular de f_3 (este principio de correspondencia término a término determina entonces el covariograma g_2 salvo una serie entera). Luego al término en r^λ le corresponde así un término en $r^{\lambda+1}$. Si λ es un entero impar, o bien en el caso de un término logarítmico, se obtiene la secuencia singular:

$$\begin{aligned} \log(r) &\rightarrow \pi r \\ r &\rightarrow -r^2 \log(r) \end{aligned}$$

Donde se alternan términos impares y términos logarítmicos. Se observará que, en todos los casos, la subida tiene un efecto regularizante.

1-4 LA ESTIMACION DE LAS VARIABLES REGIONALIZADAS

1-4-1 Expresión rigurosa de la varianza de estimación. Razonaremos en el espacio de una dimensión, sin embargo, los resultados se pueden generalizar sin dificultad. Sea $f(x)$ una V. R. Para estimar la cantidad de metal:

$$Q = \int_{-\infty}^{+\infty} f(x) dx$$

Se dispone de una malla de muestreo con una malla regular a . Si x_0 representa la implantación de una (cualquiera) de las muestras, se conocen entonces los valores numéricos de $f(x_0+pa)$ para p entero positivo o negativo). En efecto, si el campo está acotado, solo un número finito de estos valores es $\neq 0$, y basta con que la red sobrepase un poco el campo. Se toma como estimador:

$$Q^*(x_0) = a \sum_{p=-\infty}^{p=+\infty} f(x_0 + pa)$$

El error de estimación es $Q - Q^*(x_0)$: es una función periódica de período a , porque es indiferente tomar uno u otro punto de muestreo como origen de la red de muestreo.

Para hacer que este error sea una variable aleatoria, basta con admitir que "el origen" x_0 está implantado al azar según una densidad de probabilidad uniforme en el intervalo $(0, a)$. Se tiene entonces:

$$E(Q^*) = \int_0^a \left(a \sum_p f(x_0 + pa) \right) \frac{dx_0}{a} = \int_{-\infty}^{+\infty} f(x) dx = Q$$

Y la varianza de estimación (varianza de este error aleatorio) es:

$$\sigma^2(a) = E(Q^{*2}) - Q^2 = \int_0^a [Q^*(x_0)]^2 \frac{dx_0}{a} - Q^2$$

Evaluemos la integral de $[Q^*(x_0)]^2$. Se tiene:

$$[Q^*(x_0)]^2 = a^2 \sum_{p=-\infty}^{+\infty} \sum_{q=-\infty}^{+\infty} f(x_0 + pa) f(x_0 + qa) = a^2 \sum_k \sum_p f(x_0 + pa) f(x_0 + pa + ka)$$

Integremos de 0 a a . Se tiene:

$$\begin{aligned} \int_0^a [Q^*(x_0)]^2 dx_0 &= a^2 \sum_k \int_0^a \sum_p f(x_0 + pa) f(x_0 + pa + ka) dx_0 = \\ &= a^2 \sum_k \int_{-\infty}^{+\infty} f(x_0) f(x_0 + ka) dx_0 = a^2 \sum_k g(ka) \end{aligned}$$

Al tomar en cuenta la relación (2), se obtiene entonces la fórmula:

$$(3) \quad \sigma^2(a) = a \sum_{k=-\infty}^{+\infty} g(ka) - \int_{-\infty}^{+\infty} g(h) dh$$

En el espacio de n dimensiones, se tiene un resultado totalmente análogo. Por ejemplo, para n=3, y para una malla en forma de paralelepipedo $a_1 a_2 a_3$ se tiene:

$$\sigma^2(a_1, a_2, a_3) = a_1 a_2 a_3 \sum_{k_1} \sum_{k_2} \sum_{k_3} g(k_1 a_1, k_2 a_2, k_3 a_3) - \int g(h) dh$$

Observación - La varianza de estimación (3) es la diferencia entre un valor aproximado y el valor exacto de la integral $\int g(h) dh$. Por consiguiente $\sigma^2(a)$ es más pequeña cuando:

- la malla a es más pequeña
- la función g, luego también la función f, es más regular.

Si la malla es pequeña con respecto al alcance de g(h), la fórmula (3) tiene un gran número de términos. Lo anterior nos conduce a buscar fórmulas de aproximación.

1-4-2 Fórmulas de aproximación para el espacio de una dimensión.

En el espacio de una dimensión, y para una malla pequeña respecto del alcance, se puede aplicar un principio de correspondencia término a término entre la parte irregular de g(h) y el desarrollo limitado de la varianza de estimación $\sigma^2(a)$ en una vecindad de $a = 0$. A cada término irregular en r^λ le corresponde un término en $A_\lambda a^{1+\lambda}$ (A_λ es una constante que solo depende de λ):

En particular:

$$-r \rightarrow \frac{1}{6} a^2$$

$$r^2 \log(r) \rightarrow 0.0609 a^3$$

Observación: Si $L = na$ es la longitud mineralizada, y n el número de muestras positivas, al covariograma en $|r|^\lambda$ le corresponde la varianza de estimación siguiente:

$$(4) \quad \sigma^2(a) = A_\lambda a^{1+\lambda} = \left(A_\lambda L^{1+\lambda} \right) \frac{1}{n^{1+\lambda}}$$

Esta varianza está en razón inversa de $n^{1+\lambda}$ (y no en $1/n$ como lo habría sugerido una aplicación errónea de la estadística clásica).

Zitterbewegung, o término fluctuante. EL principio de correspondencia que hemos enunciado, es una aproximación. El valor exacto de $\sigma^2(a)$ calculado a partir de (3) puede diferir notablemente del proporcionado por el principio de correspondencia (ver la figura 1 adjunta): es el término fluctuante, o Zitterbewegung.

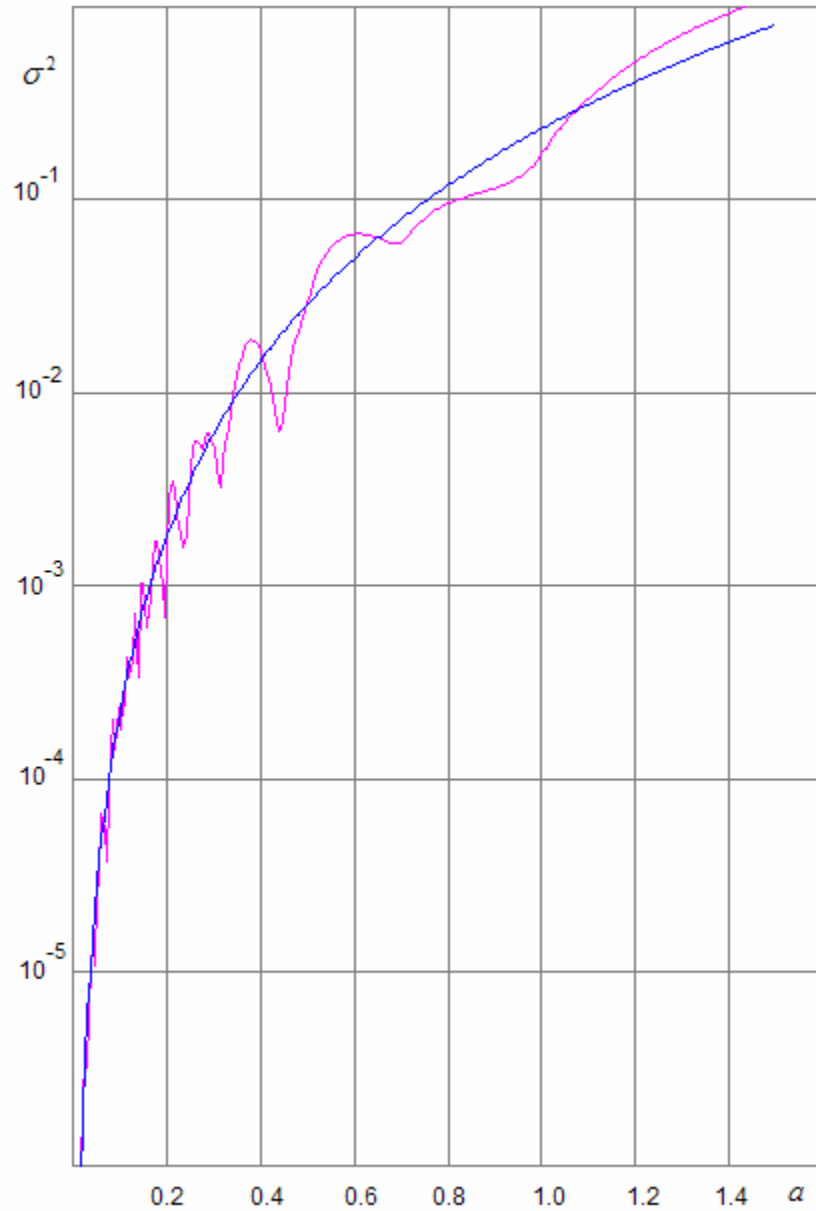
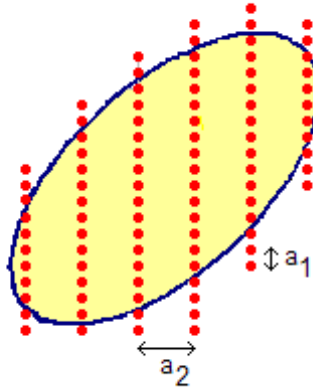


Figura 1: Zitterbewegung en la estimación de la superficie de un círculo de diámetro 1, a partir de una malla cuadrada de lado a . En abscisa, la malla a . En ordenada, la varianza de estimación correspondiente $\sigma^2(a)$. La curva en morado representa el valor exacto, calculado a partir de la fórmula exacta (1-14), la curva en azul representa la fórmula $\sigma^2(a) = 0.2276 a^3 + 0.0047 a^5$ obtenida despreciando el Zitterbewegung y reteniendo los dos primeros términos del desarrollo limitado dado por el principio de correspondencia (Nota del T.: no corresponde a la figura original, la cual es en parte errónea).

1-4-3 Fórmulas de aproximación para el espacio de dos o tres dimensiones.

A dos dimensiones (siempre excluyendo el Zitterbewegung) se tienen fórmulas de aproximación del tipo:

$$(5) \quad \sigma^2(a_1, a_2) = B_\lambda a_2^{2+\lambda} + C_\lambda a_1^{1+\lambda} a_2$$



Para un $g(h)$ en $|h|^\lambda$ con $a_1 \leq a_2$. Si $a_1 = 0$, se conocen, de manera perfecta, las acumulaciones de las líneas, entonces el término $B_\lambda a_2^{2+\lambda}$ representa la varianza del error que se comete al estimar Q a partir de los resultados de las líneas (las cuales se suponen perfectamente conocidas): es el término de rebanada (extensión de las líneas en sus rebanadas de influencia). En el espacio de tres dimensiones se tienen resultados análogos.

1-5 APLICACIÓN A LA ESTIMACION DE UNA SUPERFICIE.

Aplicamos la fórmula 5 a la V. R. $f(x) = k(x)$ igual a 1 o a 0 según que x pertenezca o no a la superficie S : se trata entonces de la estimación del área de esta superficie S a partir de una red de sondajes con malla regular $a_1 a_2$ ($a_1 \leq a_2$). El covariograma transitivo $K(h)$ asociado a S es lineal en una vecindad del origen y se tiene:

$$K(h) = S - |h| D_\alpha + \dots$$

En que D_α es la semi-variación diametral en la dirección α del vector h .

a/ Consideremos primero el caso isótropo, es decir en el caso en que la variación diametral $D_\alpha = D$ es aproximadamente independiente de la dirección α . La fórmula (5) es perfectamente aplicable y proporciona:

$$\frac{\sigma_s^2}{S^2} = \frac{1}{n^{\frac{3}{2}}} \frac{D}{\sqrt{S}} \left(\frac{1}{6} \sqrt{\frac{a_1}{a_2}} + 0.0609 \left(\frac{a_2}{a_1} \right)^{\frac{3}{2}} \right)$$

La varianza es en $1/n^{3/3}$, siendo n el número de sondajes positivos. Para utilizar esta fórmula a partir de los datos disponibles, se puede estimar S atribuyendo a cada sondaje su rectángulo de influencia, lo que proporciona:

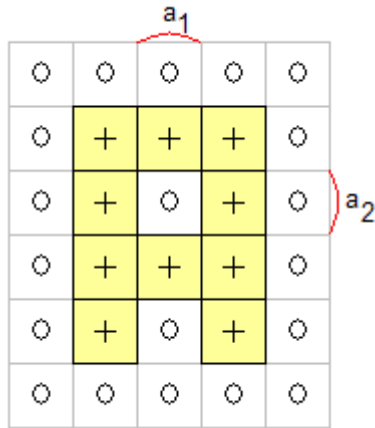
$$S = na_1a_2$$

Se estimará D a partir del contorno de la unión de estas zonas de influencia (ver la figura adjunta). Entonces se contarán los números N_1 y N_2 de elementos paralelos a a_1 y a_2 respectivamente (los cuales constituyen el perímetro) y se tendrá $D = N_1a_1 = N_2a_2$ (debido a la isotropía). Por consiguiente queda la fórmula:

$$(6) \quad \frac{\sigma_s^2}{S^2} = \frac{1}{n^2} \left(\frac{N_2}{6} + 0.061 \frac{N_1^2}{N_2} \right) \quad (N_2 \leq N_1)$$

b/ En general, sin embargo, el contorno no será suficientemente isótropo para que D_α pueda ser considerado como una constante D. Por ejemplo, puede presentar una dirección principal de alongamiento. Si uno de los lados de la malla es paralela a esta dirección principal (lo cual será a menudo el caso), la fórmula (6) anterior sigue siendo válida: en efecto, tomando esta dirección principal como eje de las x, y multiplicando las coordenadas por un módulo conveniente, obtenemos una nueva figura, isótropo esta vez (al menos en primera aproximación) para la cual (6) es válida. Pero esta transformación no ha modificado ni N_1 ni N_2 , ni la varianza relativa σ_s^2/S^2 , de manera que (6) es válida también en el caso de una figura inicial anisótropa.

Ejemplo: en la figura adjunta, se estima el área mineralizada por 10 veces el rectángulo de malla a_1, a_2 . La figura tiene un hoyo (una laguna). Al contar N_1 y N_2 deben figurar también los elementos exteriores y los elementos interiores.



Se lee en la figura:

$$2D_1 = 12 a_1 \quad \text{es decir} \quad N_1 = 6$$

$$2D_2 = 8 a_2 \quad \text{es decir} \quad N_2 = 4$$

Por consiguiente:

$$\frac{\sigma_s^2}{S^2} = \frac{1}{100} \left[\frac{4}{6} + 0.061 \times \frac{36}{4} \right] = \frac{1.21}{100}$$

Es decir una desviación estándar relativa $\sigma_s/S = 11/100$, y un error relativo (con 95% de confianza) de $\pm 22\%$.

No se debe olvidar que este cálculo hace abstracción del término fluctuante o Zitterbewegung, del cual sabemos que puede tener una magnitud enorme.

2-1 DEFINICIONES GENERALES.

Noción de Función Aleatoria. En teoría de probabilidades se define la noción de variable aleatoria (V.A.) vectorial (Y_1, Y_2, \dots, Y_k) de k componentes: es una familia de k V.A. ordinarias Y_1, Y_2, \dots, Y_k (en general no independientes). Cuando el número k de estas componentes se hace infinito, se obtiene una familia infinita de variables aleatorias: esto es una función aleatoria. En particular, si x es un punto del espacio de n dimensiones R^n , se puede definir una familia infinita $(Y(x))$, $x \in R^n$. A todo punto x_0 del espacio corresponde así una V.A. ordinaria $Y(x_0)$. $Y(x)$ es entonces una función del punto x , cuyo "valor" en x_0 no es un número, sino una V.A. (a saber $Y(x_0)$). Se dice que $Y(x)$ es una función aleatoria (abreviado F.A.). Se observará que, en general las V.A. correspondientes a dos puntos de apoyo x_1 y x_2 , es decir $Y(x_1)$ e $Y(x_2)$, no son independientes.

Si Y es una V.A. ordinaria, el resultado particular de un experimento al azar según la ley de probabilidad de Y es un valor numérico particular y . Análogamente, si Y es una V.A. vectorial (Y_1, Y_2, \dots, Y_k) , un sorteo al azar según la ley (de k variables) de Y proporciona un vector $y = (y_1, y_2, \dots, y_k)$, es decir k valores numéricos particulares. Finalmente, si $Y(x)$ es una F.A. - es decir una V.A. vectorial con una infinidad de componentes - un sorteo al azar, efectuado según la ley (de una infinidad de variables) de $Y(x)$ proporciona una función numérica particular $y(x)$, en general, extraordinariamente irregular. Se dice que $y(x)$ es una realización de la F.A. $Y(x)$.

Siempre se puede considerar una realización de una $y(x)$ de una F.A. $Y(x)$ como una variable regionalizada. Inversamente, se puede interpretar una V.R. como una realización de una cierta F.A. $Y(x)$: esta interpretación permite aplicar a las V.R. los resultados de la teoría probabilística de las F.A.

Observaciones

1/ No se puede afirmar que una V.R. $y(x)$ es una F.A. Esto tendría el mismo sentido que decir: el número 98 es una V.A. El enunciado correcto de la hipótesis probabilística de base que deseamos introducir es: $y(x)$ es una realización de una F.A. $Y(x)$.

2/ Para que esta hipótesis probabilística tenga un sentido real, es necesario poder reconstituir, al menos en parte, la ley de la F.A. de la cual suponemos que la V.R. $y(x)$ es una realización, lo cual supone que la inferencia estadística es posible. Por otra parte, la inferencia estadística no es, en general, posible si se dispone de una sola realización $y(x)$ de $Y(x)$ (de la misma manera, no se puede reconstituir la ley de una V.A. Y a partir de un resultado numérico $y=98$ de un único experimento.). Para que la inferencia estadística sea posible, es necesario introducir hipótesis suplementarias acerca de la F.A. $Y(x)$, de manera de reducir en número de "parámetros" de los cuales depende la ley de $Y(x)$. Este es el objetivo de la hipótesis estacionaria que vamos a definir: una F.A. estacionaria se repite, de una cierta manera, en el espacio, y, esta repetición hace posible la inferencia estadística a partir de una realización única. Precisemos ahora esta hipótesis:

F.A. estacionaria. - Se dice que una F.A. $Y(x)$ es estacionaria si su ley es invariante por translación. Dicho de otra manera, si x_1, x_2, \dots, x_k son k puntos de apoyo arbitrarios (k entero cualquiera), las k variables aleatorias $Y(x_1), Y(x_2), \dots, Y(x_k)$ tienen la misma ley de probabilidad (de k variables) que las k V.A. $Y(x_1+h), Y(x_2+h), \dots, Y(x_k+h)$. En adelante, $Y(x)$ designará una F.A. estacionaria.

Esperanza matemática. - Consideremos un punto de apoyo x_0 . Si la V.A. ordinaria $Y(x_0)$ admite una esperanza matemática, ésta es una función $m(x_0)=E[Y(x_0)]$ del punto de apoyo x_0 . Pero $Y(x)$ es estacionaria, por consiguiente, se tiene: $m(x_0+h)=m(x_0)$, para todo vector h , luego $m(x_0)$ es una constante m , independiente de x_0 :

$$m = E[Y(x)]$$

Se puede suponer a menudo que $m=0$, al reemplazar $Y(x)$ por $Y(x)-m$ (suponiendo siempre que esta esperanza existe).

La covarianza $K(h)$. - Consideremos ahora dos puntos de apoyo x_0 y x_0+h . Si las dos V.A. $Y(x_0)$ e $Y(x_0+h)$ admiten varianzas finitas (luego también una esperanza m que supondremos nula), $Y(x_0)$ e $Y(x_0+h)$ admiten también una covarianza $K(x_0;h)$, que depende, en principio, del punto de apoyo x_0 y del vector h . Pero como $Y(x)$ es estacionaria, se tiene: $K(x_0+a;h)=K(x_0;h)$ para todo vector a , luego $K(x_0;a)$ no depende de x_0 , y nosotros escribiremos simplemente $K(h)$:

$$(1) \quad K(h) = E[Y(x)Y(x+h)]$$

Se comparará esta definición a la del covariograma transitivo: $K(h)$ es la transpuesta probabilística de $g(h)$.

Para $h=0$ se tiene $K(0)=E([Y(x)]^2)$: que es la varianza de la V.A. $Y(x_0)$. Para que una F.A. $Y(x)$ admita una función de covarianza $K(h)$, es necesario y suficiente que $Y(x)$ admita una varianza finita $K(0)$.

Hipótesis estacionaria de orden 2. - Diremos que una F.A. $Y(x)$ es estacionaria de orden 2 si la V.A. $Y(x_0)$ admite una esperanza m independiente del punto de apoyo x_0 , y, si para todo vector h la covarianza:

$$K(h) = E[Y(x)Y(x+h)] - m^2$$

existe y no depende de x_0 . Esta hipótesis (la cual no implica la estacionaridad en sentido estricto, tal como la definimos antes) es suficiente para la teoría de las V.R. Sin embargo esta hipótesis supone la existencia de una varianza a priori finita $K(0)$.

Varianza a priori infinita. - Por otra parte, existen numerosos fenómenos que presentan una capacidad de dispersión ilimitada y no pueden ser descritos correctamente si se les atribuye una varianza a priori finita: esta afirmación puede sorprender, sin embargo hay que ver que la naturaleza nos tiende aquí una suerte de trampa. Cuando se toman muestras v en un campo V , se obtiene un histograma a partir del cual se puede calcular numéricamente una varianza, la cual toma un valor perfectamente definido. Pero esta varianza es, en realidad, una función $\sigma^2(v|V)$ del soporte v y del campo V . Esta varianza aumenta, en particular, cuando el

campo V aumenta. Si las muestras de tamaño v poseen una varianza a priori finita, ésta debe ser igual al límite cuando V tiende a infinito, de la varianza experimental $\sigma^2(v|V)$.

Es así como los autores de Africa del Sur (D. G. Krige, etc., ...), a partir de cientos de miles de muestras tomadas en el gran yacimiento de oro de Rand, calcularon la varianza de las muestras en paneles cada vez más grandes, luego en toda una concesión, luego en el yacimiento de Rand en su conjunto: obtuvieron una relación experimental de la forma:

$$\sigma^2(v|V) = \alpha \log\left(\frac{V}{v}\right)$$

El crecimiento de la varianza se produce siempre según esta ley logarítmica (fórmula de De Wijs) hasta el último punto experimental, en el cual V/v es del orden de la decena de billones: Se puede concluir, con plena seguridad, que, en este caso, no existe una varianza a priori finita.

Entonces estamos conducidos a reemplazar la hipótesis estacionaria de orden 2 por una hipótesis más débil pero de significación análoga:

Hipótesis intrínseca. En el caso en que la varianza a priori $K(0)$ no existe (es infinita), puede ocurrir que los incrementos $Y(x_0+h)-Y(x_0)$ tengan una varianza finita. Diremos entonces que la F.A. $Y(x)$ verifica la hipótesis intrínseca si, para todo vector h , el incremento $Y(x_0+h)-Y(x_0)$ admite una esperanza y una varianza independientes del punto de apoyo x_0 (pero dependiendo de h), es decir:

$$E[Y(x+h)-Y(x)] = m(h)$$

$$E[(Y(x+h)-Y(x))^2] = 2\gamma(h)$$

La función $m(h)$ es la derivada lineal. Para demostrar que es lineal en h , se parte de la relación evidente:

$$Y(x+h'+h'')-Y(x) = [Y(x+h'+h'')-Y(x+h')] + [Y(x+h')-Y(x)]$$

y se pasa a las esperanzas, de donde: $m(h'+h'')=m(h')+m(h'')$. Se puede suponer siempre que esta derivada lineal $m(h)$ es nula, al reemplazar $Y(x)$ por $Y(x)-m(x)$

La función $\gamma(h)$:

$$(2) \quad \gamma(h) = \frac{1}{2} E\left[(Y(x+h)-Y(x))^2\right]$$

se llama variograma o función intrínseca. Una F.A. que verifica la hipótesis intrínseca constituye lo que se llama un esquema intrínseco, caracterizado por su variograma.

Observación. Si $Y(x)$ verifica la hipótesis estacionaria de orden 2, verifica también la hipótesis intrínseca y se tiene, en este caso:

$$(3) \quad \gamma(h) = K(0) - K(h)$$

Como $|K(h)| \leq K(0)$, se tiene $\gamma(h) \leq 2K(0)$, de manera que el variograma de una F.A. estacionaria de orden 2 está necesariamente acotado. Existen esquemas intrínsecos de uso muy corriente cuyo variograma $\gamma(h)$ no está acotado, y que no pueden, por consiguiente, verificar la hipótesis estacionaria de orden 2 (tienen una varianza a priori infinita). Ejemplo, esquema de De Wijs ($\gamma(h) = 3\alpha \log|h|$), esquema lineal ($\gamma(h) = A|h|$).

2-2 PROPIEDADES DE LA COVARIANZA $K(h)$ Y DEL VARIOGRAMA $\gamma(h)$

a/ Simetría:

$$K(h) = K(-h) \quad , \quad \gamma(h) = \gamma(-h)$$

Desigualdades:

$$K(h) \leq K(0) ; \gamma(h) \geq 0 ; \gamma(0) = 0$$

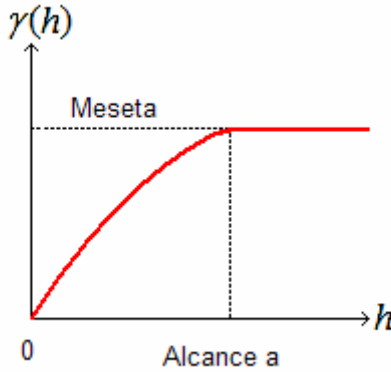
Estas condiciones son necesarias pero no son suficientes. Sea K o γ una covarianza o un variograma que verifican estas relaciones, entonces no necesariamente existe una F.A. estacionaria o intrínseca admitiendo la covarianza K o el variograma γ . En efecto, la condición necesaria y suficiente es que K pertenezca a la clase de funciones "de tipo positivo" y $-\gamma$ a la clase de funciones de "tipo positivo condicional". Por ejemplo, las funciones $\log(r)$ y r^λ con $\lambda < 2$ pueden servir como variogramas, pero no la función r^λ para $\lambda \geq 2$. Estas condiciones expresan, entre otras, que las fórmulas que estableceremos más adelante para las varianzas de extensión o de estimación, conducen, necesariamente a valores positivos (lo que no sería el caso si se toma una función cualquiera como variograma).

b/ Se puede observar que el variograma proporciona un sentido preciso a la noción tradicional de zona de influencia de una muestra: en efecto, su crecimiento más o menos rápido refleja la manera más o menos rápida de cómo se deteriora la influencia de una muestra en zonas cada vez más alejadas del yacimiento.

c/ Las anisotropías se manifiestan por el comportamiento diferente del variograma en las diferentes direcciones del espacio. En ausencia de anisotropía, $\gamma(h) = \gamma(r)$ solo depende del módulo r de h , y no depende de la dirección del vector. Se dice que hay anisotropía geométrica cuando basta con una simple transformación lineal de las coordenadas para restablecer la isotropía.

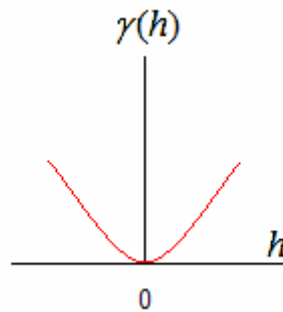
Hay tipos más complejos de anisotropías. Por ejemplo, en el espacio de tres dimensiones, puede suceder que $Y(x)$ solo depende de la tercera coordenada y sea constante en planos paralelos a los dos primeros ejes de coordenadas. El variograma $\gamma(h) = \gamma(h_3)$ solo depende de la tercera coordenada de h . A menudo, $Y(x)$ no será, en realidad, constante sobre planos horizontales, pero variará menos rápido o de manera más regular que en la dirección vertical. Se tomará entonces un variograma de la forma $\gamma(h) = \gamma(h_1, h_2, h_3) + \gamma(h_3)$ (anisotropía zonal).

d/ Alcance. En el caso estacionario de orden 2, el alcance $a(\alpha)$ en una dirección α es el valor de la distancia a partir de la cual, en esta dirección, $Y(x)$ e $Y(x+h)$ no tienen correlación (o su correlación es despreciable): $K(h)=0$ (o ≈ 0) para $|h| \geq a(\alpha)$.

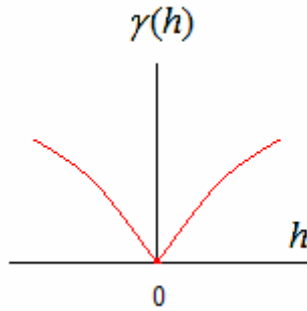


Según (3), $K(h)=0$ equivale a $\gamma(h)=K(0)=\gamma(\infty)$: el alcance es también la distancia a partir de la cual el variograma llega a su valor límite $\gamma(\infty)$ o meseta. Así, una F.A: intrínseca cuyo variograma no está acotado, no puede tener un alcance finito.

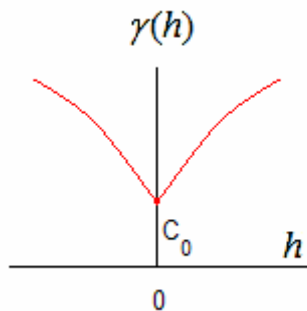
e/ Comportamiento en una vecindad del origen. La continuidad y la regularidad en el espacio de la F.A. $Y(x)$ se manifiestan en el comportamiento de $\gamma(h)$ en una vecindad del origen. Por orden de regularidad decreciente, se pueden distinguir cuatro tipos:



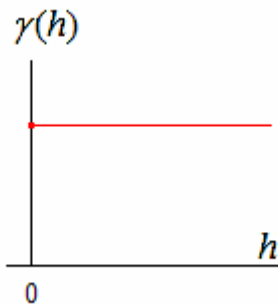
Comportamiento parabólico: $\gamma(h)$ es dos veces derivable en $h=0$. $Y(x)$ es entonces derivable (en media cuadrática), luego presenta un alto grado de regularidad en el espacio.



Comportamiento lineal (tangente oblicua en el origen): $\gamma(h)$ es continuo en $h=0$, pero no es derivable en este punto, $Y(x)$ es continua en media cuadrática, pero no es derivable, luego es menos regular.



Efecto de pepita: $\gamma(h)$ no tiende a 0 cuando h tiende a 0 (discontinuidad en el origen). $Y(x)$ no es continua en media cuadrática, luego es extraordinariamente irregular.



Caso límite completamente aleatorio: $Y(x)$ e $Y(x')$ son independientes para dos puntos cualesquiera distintos, por cercanos que sean (ruido blanco de los físicos o efecto de pepita al estado puro)

Parte irregular. En el caso isótropo ($\gamma(h)=\gamma(r)$) y en ausencia de efecto de pepita, se caracteriza el comportamiento de $\gamma(h)$ por un desarrollo limitado de la forma:

$$\gamma(r) = \sum a_{2n} r^{2n} + \sum c_{\lambda} r^{\lambda} + \sum c_{2n} r^{2n} \log(r)$$

Exactamente como en el caso de los covariogramas transitivos, se distinguirá un parte regular (términos de grado entero par), y una parte irregular (términos en r^{λ} , con λ diferente de un entero par y también

términos logarítmicos del tipo $r^{2n}\log(r)$). Si no hay parte irregular, la F.A. sería indefinidamente derivable, luego perfectamente regular. Por consiguiente solamente la parte irregular representa el grado de irregularidad de la F.A., y, en esta parte irregular, el término de más bajo grado es el que juega el papel principal: se puede definir el grado de regularidad de la F.A. como el grado λ del término irregular principal. Daremos a continuación algunas indicaciones matemáticas complementarias.

Continuidad y derivación en media cuadrática.

Se dice que la F.A. $Y(x)$ es continua en media cuadrática (continua m.c.) si se tiene:

$$E\left([Y(x+h)-Y(x)]^2\right) \rightarrow 0 \quad \text{para } |h| \rightarrow 0$$

Esto ocurre (por definición) si y solo si $\gamma(h)$ es continuo en $h=0$, es decir en ausencia de efecto de pepita.

En el espacio de una dimensión, se dice que la F.A. $Y'(x)$ es la derivada en media cuadrática de $Y(x)$ (derivada m.c.) de la F.A. $Y(x)$ si:

$$E\left(\left[\frac{Y(x+h)-Y(x)}{h}-Y'(x)\right]^2\right) \rightarrow 0 \quad \text{para } |h| \rightarrow 0$$

Se tienen definiciones análogas para el espacio de $n \neq 1$ dimensiones.

Se demuestra que una F.A. intrínseca $Y(x)$ admite una derivada m.c. $Y'(x)$ si y solo si $\gamma(h)$ es derivable dos veces en $h=0$. La derivada segunda $\gamma''(h)$ existe entonces para todo h , $Y'(x)$ es estacionaria de orden 2 (aún en el caso en que $Y(x)$ es solamente intrínseca) y admite como covarianza la función $\gamma''(h)$. Análogamente, $Y(x)$ es derivable m.c. n veces si y solo si la derivada $\gamma^{(2n)}(h)$ existe en $h=0$ (existe entonces para todo h).

Si λ es el grado del término irregular principal de $\gamma(h)$, $Y(x)$ admite entonces una derivada m.c. de orden n si y solo si $\lambda > 2n$ (si el término principal es $r^{2n}\log(r)$, $Y(x)$ es $n-1$ veces derivable m.c., pero no es derivable n veces).

2-3 REGULARIZACION DE UNA F.A. INTRINSECA.

Para simplificar la exposición, haremos los razonamientos en el caso de una F.A. estacionaria de orden 2 con covarianza $K(h)$, sin embargo todos los resultados que expresaremos con la ayuda de la función intrínseca $\gamma(h)=K(0)-K(h)$ seguirán siendo válidos para el caso de una F.A. intrínseca (luego, serán válidos aún en el caso en que $\gamma(h)$ no está acotado, y si, por consiguiente, la covarianza $K(h)$ no existe).

2-3-1 Integral estocástica

$$I = \int_{\nu} Y(x)p(x)dx$$

Se define la integral I como el límite en media cuadrática (si existe) de las sumas discretas:

$$I_n = \sum_{i=1}^n Y(x_i) p(x_i) \Delta x_i$$

(los Δx_i son pequeños elementos de volumen disjuntos cuya unión es el conjunto v , y x_i es un punto de Δx_i). La integral estocástica I es entonces una V.A., como las I_n .

Se demuestra que esta V.A. existe si y solamente si la integral

$$(4) \quad D^2(I) = \int_v p(x) dx \int_v K(x-y) p(y) dy$$

es finita, y $D^2(I)$ es entonces la varianza (finita) de esta V.A. Se encuentra fácilmente este resultado (2-9) mediante el cálculo (no riguroso) siguiente:

$$\begin{aligned} I^2 &= \int_v Y(x) p(x) dx \int_v Y(y) p(y) dy \\ E(I^2) &= \int_v p(x) dx \int_v E[Y(x)Y(y)] p(y) dy = \\ &= \int_v p(x) dx \int_v K(x-y) p(y) dy \end{aligned}$$

(el cálculo no es riguroso porque, al comienzo, no es evidente que se pueden invertir los símbolos E y \int : en efecto, se demuestra que esta inversión es legítima).

2-3-2 Convolution estocástica.

El producto de convolución $Y * f$ de la F.A. $Y(x)$ por una función ordinaria $f(x)$ es la integral estocástica (si existe):

$$\int Y(x-x') f(x') dx'$$

Se puede definir así la regularizada $Y_p = Y * \tilde{p}$ de la F.A. $Y(x)$ por una función de ponderación $p(x)$. Es la "media móvil ponderada":

$$Y_p(x) = \int Y(x+x') p(x') dx$$

Cuidado: la regularizada de $Y(x)$ es también una función aleatoria (más regular que $Y(x)$ pero siempre aleatoria): no basta con alisar o regularizar, por un procedimiento de medias móviles, una F.A. para hacer desaparecer como por arte de magia el carácter aleatorio de esta función.

La varianza de la regularizada Y_p está dada por la fórmula (4) anterior, e $Y_p(x)$ existe en el sentido de la integración m.c. si y solo si $D^2(Y_p) < \infty$. Calculemos la covarianza

$$E[Y_p(x_0)Y_p(x_0+h)] \quad \text{de} \quad Y_p(x_0) \text{ e } Y_p(x_0+h)$$

Se tiene:

$$Y_p(x_0)Y_p(x_0+h) = \iint p(x')p(x'')Y(x_0+x')Y(x_0+h+x'')dx'dx''$$

Pasemos a las esperanzas intercambiando E y \int . Queda:

$$E[Y_p(x_0)Y_p(x_0+h)] = \iint p(x')p(x'')K(h+x''-x')dx'dx''$$

Esta covarianza no depende de x_0 sino solamente de h . Luego, la regularizada Y_p es estacionaria de orden 2, y admite la covarianza:

$$(5) \quad K_p(h) = \int p(x)dx \int K(h+x-y)dy$$

Esta fórmula generaliza (4). Para presentarla en una forma más sintética, hagamos el cambio de variables $x=y+z$. Queda:

$$K_p(h) = \int K(x+z)dx \int p(y+z)dy$$

Sea P el covariograma transitivo ($P=p*\tilde{p}$) de la función de ponderación p . Se tiene, por definición:

$$P(z) = \int p(y+z)p(y)dy$$

y, por consiguiente:

$$K_p(h) = \int K(h+z)P(z)dz$$

es decir, utilizando productos de convolución (porque $P=\tilde{P}$):

$$(5') \quad K_p = K * P$$

Se obtiene la covarianza K_p de la regularizada Y_p al regularizar la covarianza $K(h)$ de Y con el covariograma transitivo de la función de ponderación. Se comparará este resultado con el obtenido en el caso de los métodos transitivos.

El variograma γ_p de la regularizada Y_p es $K_p(0)-K_p(h)$. Al reemplazar $K(h)$ por $K(0)-\gamma(h)$ en (4) y (5), observamos que $K(0)$ se elimina, y queda:

$$(6) \quad \gamma_p(h) = \int \gamma(h+z)P(z)dz - \int \gamma(z)P(z)dz$$

Se puede también escribir:

$$(6') \quad \gamma_p = \gamma * P - A$$

La constante A se determina con la condición $\gamma_p(0)=0$. Estas relaciones (6) y (6') siguen siendo válidas para cualquier F.A. intrínseca (aún en el caso en que la covarianza no existe).

2-3-3 Subida. La subida constituye un caso particular de la regularización. En los métodos transitivos podíamos integrar la V.R. de $-\infty$ a $+\infty$ con respecto a una de las coordenadas, por ejemplo x_3 , debido a que esta V.R. era nula al exterior de un campo acotado. Ahora, solo podemos integrar en un largo finito ℓ . Llamaremos entonces subida bajo potencia constante ℓ a la operación que nos hace pasar de la F.A. $Y_3(x_1, x_2, x_3)$ definida en el espacio de tres dimensiones a la F.A. $Y_2(x_1, x_2)$ definida en el espacio de dos dimensiones por:

$$Y_2(x_1, x_2) = \frac{1}{\ell} \int_0^{\ell} Y(x_1, x_2, x_3) dx_1 dx_2 dx_3$$

Si $Y_3(x)$ es la ley puntual en una formación estratiforme de potencia ℓ , $Y_2(x_1, x_2)$ es la ley media del sondaje implantado en el punto (x_1, x_2) de la superficie topográfica.

La subida tiene un efecto regularizador. En el caso isótropo, se demuestra que el término irregular principal del variograma de Y_2 es en $r^{1+\lambda}$ si el de Y_3 es en r^λ . De manera más precisa, se tienen las reglas:

$$\begin{aligned} r^\lambda &\rightarrow A_\lambda \frac{r^{1+\lambda}}{\ell} \\ \log(r) &\rightarrow \frac{r}{\ell} \left(A_\lambda = \sqrt{\pi} \operatorname{tg} \left(\frac{\lambda\pi}{2} \right) \frac{\Gamma \left(1 + \frac{\lambda}{2} \right)}{\Gamma \left(1 + \frac{1+\lambda}{2} \right)} \right) \\ r &\rightarrow -r^2 \log(r) \end{aligned}$$

2-4 VARIANZA DE EXTENSION Y VARIANZA DE ESTIMACION

2-4-1 Varianza de extensión.

Definamos, en primer lugar, la noción fundamental de varianza de extensión. Sea $Y(x)$ una F.A., la cual supondremos, por el momento, como estacionaria de orden 2, y sea $K(h)$ su covarianza. Designemos por $Z(v)$ y $Z(v')$ las "leyes medias" de dos dominios v y v' del espacio de n dimensiones, es decir las integrales estocásticas:

$$Z(v) = \frac{1}{v} \int_v Y(x) dx \quad ; \quad Z(v') = \frac{1}{v'} \int_{v'} Y(x) dx$$

La fórmula (2-9) muestra que, si v y v' son acotados, $Z(v)$ y $Z(v')$ tienen varianzas finitas. La primera, por ejemplo, tiene la expresión:

$$\sigma^2(v) = \frac{1}{v^2} \int_v dx \int_v K(x-y) dy$$

Calculemos también la covarianza $\sigma(v, v')$ de $Z(v)$ y $Z(v')$. De:

$$Z(v)Z(v') = \frac{1}{vv'} \int_v Y(x) dx \int_{v'} Y(y) dy$$

se deduce, al pasar a las esperanzas matemáticas:

$$\sigma(v, v') = \frac{1}{vv'} \int_v dx \int_{v'} E[Y(x)Y(y)] dy = \frac{1}{vv'} \int_v dx \int_{v'} K(x-y) dy$$

Llamaremos varianza de extensión de v a v' (o de v' a v) a la varianza del error $Z(v') - Z(v)$ cometido al atribuir a v' la ley media $Z(v)$ de v . Esta varianza de extensión σ_E^2 es igual a:

$$\sigma_E^2 = \sigma^2(v) + \sigma^2(v') - 2\sigma(v, v')$$

Al considerar los valores calculados antes, se tiene entonces:

$$\sigma_E^2 = \frac{1}{v^2} \int_v dx \int_v K(x-y) dy + \frac{1}{v'^2} \int_{v'} dx \int_{v'} K(x-y) dy - \frac{2}{vv'} \int_v dx \int_{v'} K(x-y) dy$$

Reemplacemos $K(h)$ por $K(0) - \gamma(h)$: observamos que la constante $K(0)$ desaparece de la expresión de σ_E^2 , y queda la fórmula fundamental:

$$(7) \quad \sigma_E^2 = \frac{2}{vv'} \int_v dx \int_v \gamma(x-y) dy - \frac{1}{v^2} \int_v dx \int_v \gamma(x-y) dy - \frac{1}{v'^2} \int_{v'} dx \int_{v'} \gamma(x-y) dy$$

Se puede demostrar que (7) sigue siendo válida para toda F.A. intrínseca, luego también en el caso en que $K(h)$ no existe.

Varianza de estimación.

Supongamos ahora que en vez de conocer la "ley media" $Z(v')$ de $Y(x)$ en un volumen v' , conocemos la ley media:

$$Z' = \frac{1}{N} \sum_{i=1}^N Y(x_i)$$

de N muestras tomadas en los puntos x_i . Z' es una variable aleatoria, cuyas características se deducen sin dificultad a partir de $K(h)$ o de $\gamma(h)$. Llamaremos varianza de estimación σ_N^2 (de v por las N muestras tomadas en los N puntos x_i) a la varianza de la diferencia $Z(v) - Z(v')$. Para obtener la expresión de σ_N^2 , se debe, en cada una de las etapas del razonamiento que nos condujo a (7), reemplazar las integrales sobre el

conjunto v' por sumas discretas sobre los N puntos x_i . Se encuentra así la segunda fórmula fundamental:

$$(8) \quad \sigma_N^2 = \frac{2}{Nv} \sum_i \int_v \gamma(x_i - x) dx - \frac{1}{v^2} \int_v dx \int_v \gamma(x - y) dy - \frac{1}{N^2} \sum_i \sum_j \gamma(x_i - x_j)$$

Esta fórmula, en la cual se alternan expresiones exactas y aproximadas de las mismas integrales, presenta una estructura sobresaliente; análoga, un poco más compleja, a la fórmula (1-14) de los métodos transitivos. En particular, se observa que la varianza de estimación es más débil:

- cuando la red de muestras x_i es más densa y más representativa de la geometría del volumen v a estimar
- cuando la función $\gamma(h)$ es más regular, luego cuando la F.A. es más continua en su variación espacial.

Sin embargo, en la práctica, si N es grande, la fórmula (8) conduce a cálculos bastante largos. Más adelante proporcionaremos fórmulas de aproximación más simples.

Observación: No hay ninguna diferencia conceptual entre las nociones de varianza de extensión y de estimación: la fórmula (8) es un caso particular de (7), en que v' se reemplaza por la unión de los N puntos x_i . La práctica ha reservado el término varianza de extensión a la extensión de una muestra única en su "zona de influencia", y el de varianza de estimación a la extensión de un número más grande de muestras, en el depósito entero, o en un gran panel.

2-4-2 Varianza de v en V .

La noción de varianza $\sigma^2(v|V)$ de una muestra v en un dominio V parece, a primera vista, evidente desde el punto de vista experimental. Sin embargo, esta noción, solamente tiene sentido en los dos casos a/ y b/ siguientes:

a/ v se reduce a un punto: La varianza $\sigma^2(0|V)$ de la variable puntual $Y(x)$ es igual a la esperanza de $[Y(x) - Z(V)]^2$ cuando se sortea al azar el punto x , con una densidad de probabilidad uniforme en V ($Z(V) = (1/V) \int_v Y(x) dx$ es la ley media de V)

Es entonces el valor medio en $x \in V$ de la varianza de extensión del punto x al volumen V , lo cual es igual a, según (7) o (8):

$$\frac{2}{V} \int_v \gamma(x - y) dy - \frac{1}{V^2} \int_v dy \int_v \gamma(y - y') dy'$$

Al integrar, con x dentro de V , se observa que el primer término es el doble del segundo, y queda:

$$\sigma^2(0|V) = \frac{1}{V^2} \int_V dx \int_V \gamma(x-y) dy$$

(Que es el valor medio en V de $\gamma(h)$ cuando las dos extremidades x e y del vector h describen, cada una por su propia cuenta, el volumen V).

b/ Cuando el volumen V es la unión de N volúmenes v_i disjuntos, iguales a un mismo volumen v de referencia (más precisamente, cada uno de los v_i se deduce de v por traslación). Si Z_i es la "ley media" de $Y(x)$ en v_i , se tiene, esta vez:

$$\sigma^2(v|V) = \frac{1}{N} \sum_{i=1}^N E \left[(Z_i - Z(V))^2 \right]$$

Es entonces la media aritmética de las varianzas de extensión de los v_i en V: Después de un cálculo fácil, la fórmula (7) conduce a:

$$(9) \quad \sigma^2(v|V) = \frac{1}{V^2} \int_V dx \int_V \gamma(x-y) dx dy - \frac{1}{v^2} \int_v dx \int_v \gamma(x-y) dx dy$$

En particular, se tiene:

$$(9') \quad \sigma^2(v|V) = \sigma^2(0|V) - \sigma^2(0|v)$$

c/ En el caso general en el cual v y V son volúmenes cualesquiera, no necesariamente compatibles geoméricamente, se define la varianza por la misma fórmula (9): Esta cantidad represente ahora un simple artificio de cálculo. Esta varianza puede tomar valores negativos; en particular, se tiene siempre:

$$\sigma^2(v|V) = -\sigma^2(V|v)$$

d/ Relación de aditividad. Sean v, V y V' tres volúmenes, con, por ejemplo: $v \subset V \subset V'$. De la relación (9') se deduce:

$$\begin{aligned} \sigma^2(v|V) &= \sigma^2(0|V) - \sigma^2(0|v) \\ \sigma^2(v|V') &= \sigma^2(0|V') - \sigma^2(0|v) \end{aligned}$$

y por diferencia:

$$\sigma^2(v|V') - \sigma^2(v|V) = \sigma^2(0|V') - \sigma^2(0|V) = \sigma^2(V|V')$$

es decir:

$$(10) \quad \sigma^2(v|V') = \sigma^2(v|V) + \sigma^2(V|V')$$

La varianza de la muestra v en el campo V' es igual a la suma de las varianzas de v en el panel V y del panel V en el campo V' .

e/ Covarianza de v y v' en V . De manera análoga, se define la covarianza $\sigma(v, v'|V)$ de dos muestras v y v' (cuya distancia y disposición mutua es fija) en el campo V . Se encuentra:

$$\sigma(v, v'|V) = \frac{1}{V^2} \int_V dx \int_V \gamma(x-y) dy - \frac{1}{Vv'} \int_V dx \int_{v'} \gamma(x-y) dy$$

f/ En particular, a menudo es cómodo expresar la varianza de estimación (8) o la varianza de extensión (7) en función de las varianzas y covarianzas de las muestras en un campo V arbitrario. Se encuentra, por ejemplo, que:

$$\sigma_E^2 = \sigma^2(v|V) + \sigma^2(v'|V) - 2\sigma(v, v'|V)$$

como puede verse al sustituir las expresiones de estas varianzas en V : el término:

$$\frac{1}{V^2} \iint_V$$

desaparece, y los tres términos que quedan proporcionan el segundo miembro de (7). Se obtiene así una expresión, quizás más intuitiva, de la varianza de extensión.

2-4-3 Aplicación: mallas aleatoria y aleatoria estratificada.

El caso de las mallas regulares, que es el más difícil, lo trataremos más adelante, examinaremos entonces los dos tipos usuales de mallas aleatorias.

a/ Malla aleatoria pura. Para estimar la ley $Z(V)$ de un volumen V , se dispone de los valores $Y(x_i)$ de la F.A. en N puntos x_i implantados "no importa donde" en V . Admitiremos que todo ocurre como si cada x_i estuviera implantado al azar en V con una densidad uniforme e independiente de las otras muestras. Se puede obtener la varianza de estimación σ_N^2 al integrar (2-15) en V relativamente a cada uno de los x_i . Este cálculo es simple, pero es aún más simple observar que los errores parciales $Y(x_i) - Z(V)$ son independientes unos de otros (si la realización es fija) y admiten la misma varianza $s^2(0|V)$: el error resultante que es:

$$\frac{1}{N} \sum [Y(x_i) - Z(V)]$$

admite entonces la varianza:

$$\sigma_N^2 = \frac{1}{N} \sigma^2(0|V)$$

En el caso aleatorio puro, la varianza de estimación es igual a la varianza de una muestra en el campo V dividida por el número N de muestras.

b/ Malla aleatoria estratificada. En este caso, el volumen V que se quiere estimar está dividido en N zonas de influencia iguales y disjuntas v_i , y en cada v_i se toma una muestra en un punto x_i elegido al azar en la zona de influencia v_i con una densidad de probabilidad uniforme, e independiente de las otras muestras. Se puede calcular σ_N^2 al integrar (2-15) en x_i dentro de v_i para cada uno de los v_i , pero es más simple observar que el error total es:

$$\frac{1}{N} \sum_i [Y(x_i) - Z(v_i)]$$

y que cada uno de los errores parciales $Y(x_i) - Z(v_i)$ es independiente de los otros y admite la varianza $\sigma^2(0|V)$. Se deduce, entonces:

$$\sigma_N^2 = \frac{1}{N} \sigma^2(0|V)$$

En el caso de la malla aleatoria estratificada, la varianza de estimación es entonces igual a la varianza de una muestra en su zona de influencia dividida por el número N de muestras.

Observación - Es claro que la malla aleatoria estratificada proporciona siempre mejores resultados que la malla aleatoria pura. En efecto, según (9') se tiene:

$$\frac{1}{N} [\sigma^2(0|V) - \sigma^2(0|v)] = \frac{1}{N} \sigma^2(v|V) \geq 0$$

2-5 EFECTO DE PEPITA Y FENOMENOS DE TRANSICION.

Un variograma de alcance a finito caracteriza lo que se llama fenómeno de transición: más lejos de la distancia a, se alcanza la independencia, y el alcance a proporciona la escala de las estructuras elementales del fenómeno regionalizado correspondiente. Por otra parte, a menudo existe superposición de varias estructuras de escalas bien diferentes, anidadas unas en otras. El variograma experimental muestra entonces una sucesión de alcances y de mesetas, cuyo análisis permite reconstituir la jerarquía de estas "estructuras anidadas".

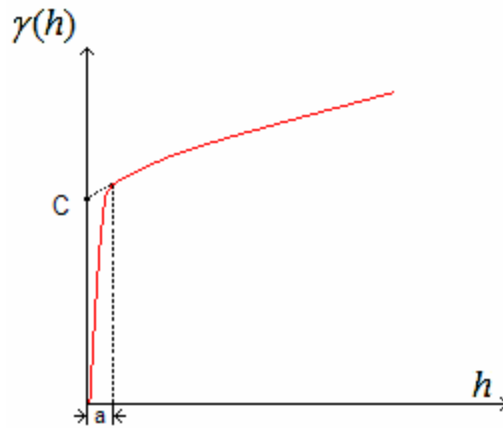
La noción de escala juega aquí un papel primordial. A la escala de la decena o de la centena de metros, un fenómeno de transición cuyo alcance es, por ejemplo, centimétrico, se manifiesta, sobre el $\gamma(h)$ experimental como una discontinuidad en el origen, es decir como un efecto de pepita. De una manera general, un efecto de pepita es una reminiscencia de una estructura de transición cuyas dimensiones fueron sobrepasadas por la escala a la cual se trabaja: los detalles y las características cualitativas de esta estructura anterior han dejado de ser perceptibles hace tiempo, y, la escala superior solo ha conservado un parámetro único

- la constante de pepita - que proporciona una suerte de medida global de la "intensidad" de esta estructura sobrepasada.

Para analizar la génesis de un efecto de pepita, localicémonos al nivel puntual, y supongamos que a una estructura primaria de dimensión a , se superpone una macro-regionalización, es decir una estructura secundaria de dimensiones mucho más grandes. Si solo existiera la estructura primaria, se podría describir la V.R. correspondiente como una realización de una F.A. con covarianza $C(h)$ de alcance a , o con un variograma:

$$\gamma_1(h) = C - C(h)$$

de alcance a y verificando $\gamma(\infty) = C = C(0)$. Para considerar la macro-regionalización, se debe agregar un segundo componente $\gamma_2(h)$, que representa la estructura secundaria, la cual varía con mucha lentitud a la escala de la primera estructura:



$$\gamma(h) = C - C(h) + \gamma_2(h)$$

$\gamma(h)$ presenta entonces, en la vecindad del origen, una zona de crecimiento muy rápido, con dimensiones del orden de a . A la escala de la macro-regionalización, este $\gamma(h)$ presenta entonces un efecto de pepita de amplitud C .

Examinemos las consecuencias de este efecto de pepita. Primero, al nivel macroscópico, las muestras no serán puntuales, sino volúmenes v ya grandes con respecto a a . Determinemos entonces el variograma $\gamma_v(h)$ de estas muestras v (el único variograma accesible experimentalmente). $\gamma_v(h)$ es la suma de una componente muy continua γ_2 (la cual no ha sido sensiblemente alterada en esta regularización) y de lo que se obtiene al aplicar la fórmula (6) a la componente $C - C(h)$: esta componente pepítica es entonces:

$$\gamma_p(h) = \frac{1}{v^2} \int_v dx \int_v C(x-y) dy - \frac{1}{v^2} \int_v dx \int_v C(h+x-y) dy$$

Estudiemos entonces γ_p . Se tiene $\gamma_p(0)=0$, pero cuando h sobrepasa las dimensiones del volumen v , se tiene $|h+x-y| \geq a$ si $x \in v$, $y \in v$ y $C(h+x-y)=0$. El primer término subsiste solo si h no es muy pequeño y se observa experimentalmente una discontinuidad en el origen, cuyo valor σ_p^2 (varianza de pepita) es:

$$\sigma_p^2 = \frac{1}{v^2} \int_v dx \int_v C(x-y) dy$$

Por otra parte, por hipótesis, las dimensiones de v son grandes con respecto a a y se tiene:

$$\int C(x-y) dy = \int C(h) dh$$

(integral extendida a todo el espacio) salvo si x está a una distancia inferior a a de la frontera de v pero estos puntos ocupan un volumen despreciable, se puede entonces despreciar su influencia, que es del orden de a . Queda entonces:

$$\sigma_p^2 = \frac{1}{v^2} \int_v dx \int_v C(h) dh = \frac{1}{v} \int C(h) dh$$

Luego, la constante de pepita σ_p^2 que se observa experimentalmente es inversamente proporcional al volumen de las muestras:

$$\sigma_p^2 = \frac{A}{v}$$

y el coeficiente $A = \int C(h) dh$, que es la covarianza de las micro estructuras, es el único recuerdo de éstas el cual subsiste a la escala de los volúmenes v .

El efecto de pepita aumenta de la misma manera la varianza de v en V , las varianzas de extensión y las varianzas de estimación. En el caso de $\sigma^2(v|V)$, por ejemplo, el aporte del efecto de pepita será:

$$\frac{1}{v^2} \int_v dx \int_v C(x-y) dy - \frac{1}{V^2} \int_V dx \int_V C(x-y) dy$$

es decir, según el cálculo anterior:

$$A \left(\frac{1}{v} - \frac{1}{V} \right)$$

En el caso de la varianza de estimación de V por N muestras de tamaño v , se encuentra una componente pepítica igual a:

$$A\left(\frac{1}{Nv} - \frac{1}{V}\right)$$

En todos los casos, la varianza debida al efecto de pepita es inversamente proporcional al volumen de las muestras. Todo ocurre como si la V.R. admitiera dos componentes independientes, uno muy regular correspondiente al variograma γ_2 , el otro completamente aleatorio y discontinuo el cual considera el efecto de pepita.

2-6 CALCULO DE VARIANZAS DE ESTIMACION.

La fórmula general (8) es bastante pesada para su manejo numérico, entonces vamos a desarrollar algunos principios de aproximación que simplificarán el cálculo de las varianzas de estimación (para el caso de mallas regulares) de manera de presentar los resultados en forma de ábacos fáciles de consultar.

2-6-1 Caso del espacio de una dimensión

a/ Las funciones intrínsecas auxiliares. En las aplicaciones, además de $\gamma(h)$, se utiliza constantemente las funciones siguientes:

$$\chi(h) = \frac{1}{h} \int_0^h \gamma(x) dx$$

$$F(h) = \frac{2}{h^2} \int_0^h x \chi(x) dx = \frac{2}{h^2} \int_0^h (h-x) \gamma(x) dx$$

$F(h)$ es el valor medio de $\gamma(h')$ cuando las dos extremidades de h' recorren el segmento $(0, h)$.

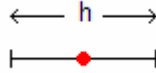
Luego, la varianza del segmento h en el segmento L es:

$$\sigma^2(h|L) = F(L) - F(h)$$

$\chi(h)$ permite también el cálculo de covarianzas: La covarianza en L del segmento h con una de sus extremidades (puntual) es:

$$\sigma(0, h|L) = F(L) - \chi(h)$$

b/ Varianza de extensión elemental. Se llama así a la varianza de extensión al segmento h de la muestra puntual implantada en el centro del segmento.

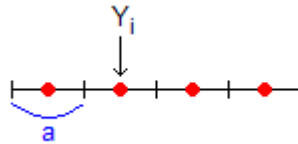


Está dada por:

$$(13) \quad \sigma_E^2 = 2\chi\left(\frac{h}{2}\right) - F(h)$$

Se deduce (13) de (12) y de las relaciones establecidas en a/.

c/ Principio de composición de varianzas de extensión elementales. Se desea estimar la ley media del segmento $L = na$ dividido en n zonas de influencia de longitud a , en el centro de cada una de estas zonas se ha tomado una muestra puntual:



Sea Y_i la ley de la muestra i , Z_i la ley de su zona de influencia, y $Z = (1/n)\sum Z_i$, la ley media de L . El error total es la media:

$$\frac{1}{n} \sum_i (Y_i - Z_i)$$

de los errores parciales $Y_i - Z_i$. El principio de aproximación consiste en admitir que estos errores parciales son independientes entre sí (este principio se verifica con una aproximación muy razonable para los $\gamma(h)$ de tipo usual). Como la varianza de $Y_i - Z_i$ es justamente la varianza de extensión elemental calculada en (13), se encuentra:

$$(14) \quad \sigma_N^2 = \frac{1}{n} \sigma_E^2 = \frac{1}{n} \left[2\chi\left(\frac{a}{2}\right) - F(a) \right]$$

Se obtiene entonces la varianza de estimación al dividir la varianza de estimación elemental de una muestra en su zona de influencia por el número n de muestras.

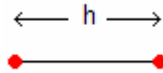
d/ Caso del dispositivo cerrado. A partir de $n + 1$ muestras a malla regular, a , se desea estimar el segmento $L = na$ comprendido entre la primera y la última muestra:



Se demuestra que este dispositivo cerrado es equivalente al dispositivo centrado examinado en c/, y que la varianza de estimación está dada también por (2-23) (siempre que n sea superior a 1): el dispositivo

cerrado con n+1 muestras es equivalente al dispositivo centrado con n muestras.

Observación: para n = 1, el dispositivo cerrado con dos muestras proporciona la varianza de estimación:



$$\sigma_{E'}^2 = 2\chi(h) - F(h) - \frac{1}{2}\gamma(h)$$

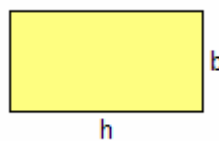
que difiere de la varianza de estimación elemental (13). No se puede dividir $\sigma_{E'}$ por n, porque los errores parciales cometidos no son necesariamente independientes (en particular, estos dos segmentos tienen en común la muestra central).



2-6-2. Caso del espacio de dos dimensiones.

Supondremos que el variograma es isótropo (solo depende de r). En caso de anisotropía geométrica es fácil adaptarse a este caso.

a/ Las funciones auxiliares. Resulta cómodo introducir las funciones siguientes:



$\gamma_b(h)$: media de $\gamma(x-y)$ cuando x e y recorren los dos lados paralelos de longitud b del rectángulo b x h (salvo una constante $\gamma_b(h)$ se deduce de $\gamma(h)$ por subida de orden 1).

$\chi_b(h)$: media de $\gamma(x-y)$, cuando x describe uno de los lados del rectángulo e y recorre el interior del rectángulo. Se tiene:

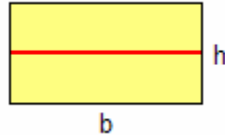
$$\chi_b(h) = \frac{1}{h} \int_0^h \gamma_b(x) dx$$

$F(b, h)$: media de $\gamma(x-y)$ cuando x e y describen el rectángulo. Esta función es simétrica en b y h, y verifica:

$$F(b, h) = \frac{2}{h^2} \int_0^h x \chi_b(x) dx = \frac{2}{h^2} \int_0^h (h-x) \gamma_b(x) dx$$

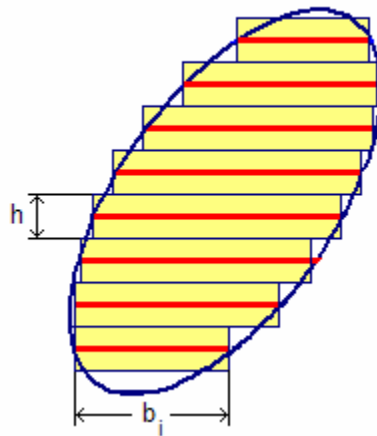
$Q(b, h)$: media de $\gamma(x-y)$, cuando x describe el lado b e y describe el lado h , o también, x describe el rectángulo e y está fijo en uno de los vértices.

b/ Extensión de un segmento medio en su rectángulo de influencia.



$$(15) \quad \sigma_E^2 = 2\chi_b\left(\frac{h}{2}\right) - F(b, h) - \gamma_b(0)$$

c/ Estimación de S por rebanadas paralelas equidistantes.



Sean b_1, b_2, \dots, b_n las longitudes de n rebanadas, h su equidistancia. Se asimila la superficie S que se quiere estimar a la unión de los rectángulos de influencia de las n rebanadas. La fórmula (2-24) proporciona la extensión de b_i a su rectángulo de influencia, es decir, σ_{Ei}^2 . Si Y_i es la ley de b_i , Z_i la de su zona de influencia, admitiendo que los $(Y_i - Z_i)$ son independientes, el error total:

$$\frac{\sum b_i (Y_i - Z_i)}{\sum b_i}$$

admite una varianza que se puede calcular ponderando por el cuadrado de los b_i de las varianzas de extensión σ_{Ei}^2 . Se obtiene entonces la varianza de estimación:

$$(16) \quad \sigma_E^2 = \frac{\sum b_i^2 \sigma_{E_i}^2}{(\sum b_i)^2}$$

Observación. Si los b_i son iguales, queda simplemente:

$$\sigma_E^2 = \frac{1}{n} \sigma_{E_i}^2$$

d/ Composición de término de línea y término de rebanada.

En el caso c/ anterior, sucede que a menudo no se conocen las leyes reales de las líneas, sino solamente una estimación de estas leyes a partir de muestras puntuales en una malla regular a ($a < h$). Se admite entonces que los errores cometidos al estimar las líneas a partir de las muestras y de S a partir de las líneas (supuestamente conocidas) pueden ser mirados como independientes. La varianza de estimación es entonces:

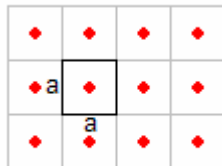
$$(17) \quad \sigma_E^2 = \frac{1}{N} \sigma^2(a) + \frac{\sum b_i^2 \sigma_{E_i}^2}{(\sum b_i)^2}$$

$\sigma^2(a)$ es la varianza de extensión elemental (2-23) de una muestra puntual en su segmento a de influencia, N es el número total de muestras. $(1/N)\sigma^2(a)$ es el término de línea, varianza del error cometido al estimar las líneas a partir de las muestras.

El segundo término, que figuraba ya en (16) es el término de rebanada: varianza del error cometido al extender las líneas a sus rebanadas de influencia.

Este principio (17) de composición es válido siempre que a sea inferior a h. Se aplica, en particular, al caso de un reconocimiento con malla rectangular axh.

e/ Caso de una malla cuadrada. La varianza de extensión de un sondeaje central a su cuadrado de influencia es:



$$(18) \quad \sigma_E^2 = 2Q\left(\frac{a}{2}, \frac{a}{2}\right) - F(a, a)$$

Para una malla cuadrada, se puede admitir que los errores cometidos al estimar cada cuadrado a partir de su sondaje central son independientes. La varianza de estimación con N sondajes es entonces:

$$\sigma_N^2 = \frac{1}{N} \sigma_E^2$$

con un σ_E^2 dado por (18).

f/ Abacos. Para cada esquema isótropo, es decir para cada función $\gamma(r)$, se puede presentar, en forma de ábacos:

1/ La función $F(a, b)$ definida en a/, la cual sirve para el cálculo de la varianza de una muestra en el rectángulo a, b.

2/ La varianza de extensión (15) de una línea media de longitud b en el rectángulo b, h - en particular, si $b = 0$, se encuentra la varianza de extensión elemental (13).

3/ La varianza de extensión (18) de un sondaje en su cuadrado a, a de influencia.

Estos tres documentos permiten un cálculo muy rápido de todas las varianzas de estimación.

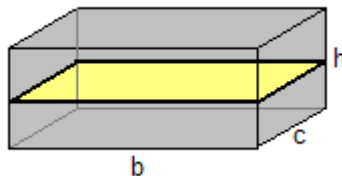
2-6-3 Caso del espacio de tres dimensiones.

Se procederá exactamente como en el caso del espacio de dos dimensiones, componiendo esta vez, tres tipos de términos:

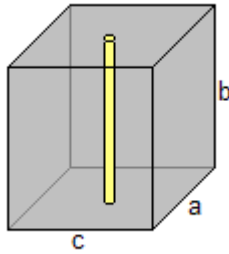
- término de línea: extensión de las muestras puntuales en las líneas
- término de sección: extensión de las líneas en las secciones
- término de rebanada: extensión de las secciones en las rebanadas

Además de los precedentes, se debe disponer de los ábacos siguientes:

- varianza de extensión de un plano mediano b, c en su paralelepípedo rectangular de influencia.



- extensión de un sondaje de longitud b en su prisma recto a, a, b de influencia.



- media de $\gamma(h)$ en el paralelepípedo rectangular a, b, c : esta función $F(a, b, c)$ sirve para el cálculo de la varianza de una muestra en este paralelepípedo.

3 - EL ESQUEMA ESFERICO.

Para representar un fenómeno de transición, se pueden utilizar esquemas de la forma:

$$\gamma(h) = A[K(0) - K(h)]$$

en que $K(h)$ es el covariograma geométrico de un volumen v . Si se toma por v la esfera de diámetro a , se tiene (ver ejercicio 9 de los métodos transitivos, capítulo 1):

$$K(h) = \frac{\pi}{6} a^3 \left(1 - \frac{3h}{2a} + \frac{1}{2} \frac{h^3}{a^3} \right) \quad \text{si } (|h| < a)$$

$$K(h) = 0 \quad \text{si } (|h| \geq a)$$

El esquema esférico estará entonces definido por el variograma:

$$\gamma(r) = C \left(\frac{3r}{2a} - \frac{1}{2} \frac{r^3}{a^3} \right) \quad \text{si } r < a$$

$$\gamma(r) = 0 \quad \text{si } r \geq a$$

El alcance es a , la meseta es $C = \gamma(\infty)$, la pendiente en el origen es $(3/2)(C/a)$.

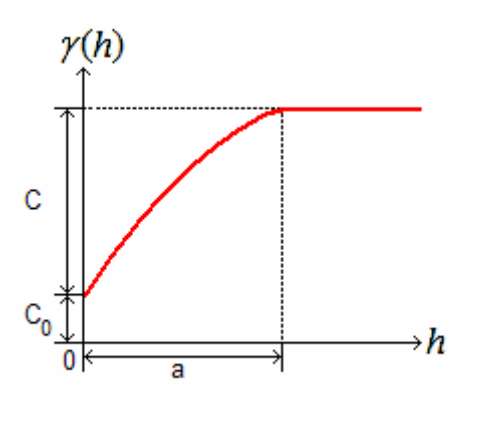
Se tienen los ábacos siguientes:

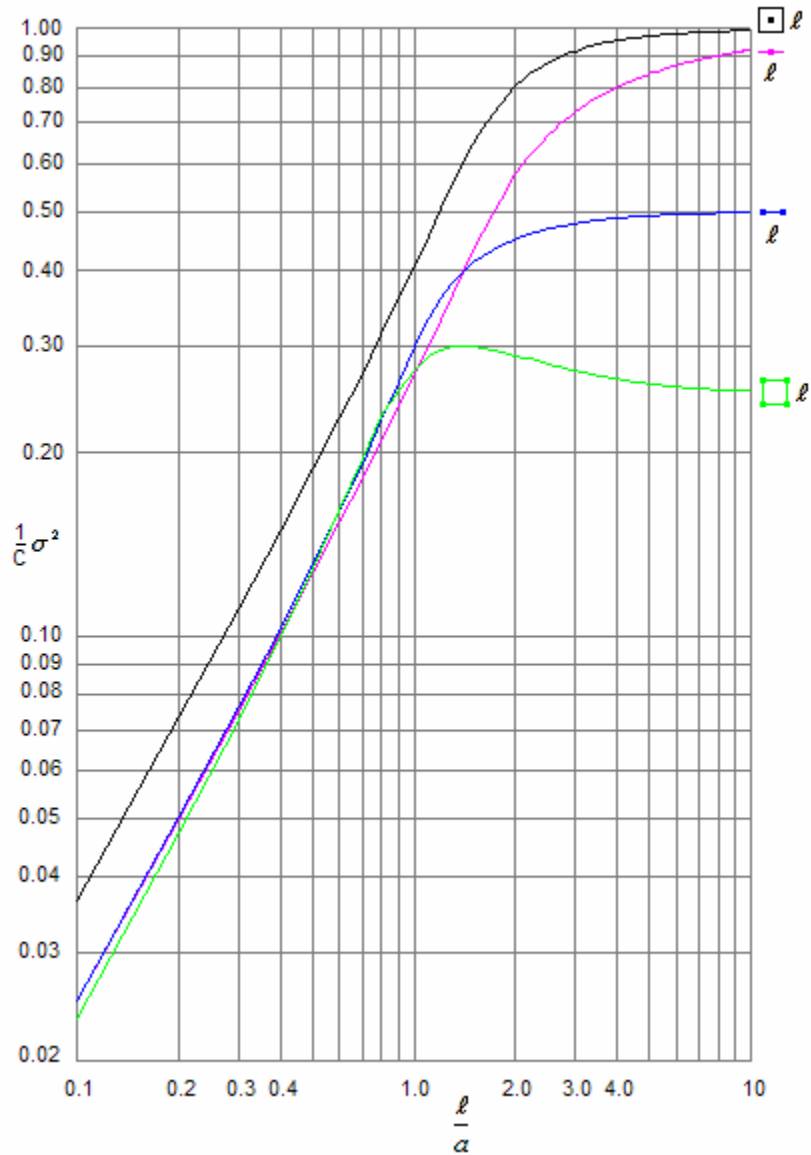
Varianzas de estimación diversas.

Varianza de un punto en el rectángulo $l \times h$:

$$\frac{1}{C} F \left(\frac{\ell}{a}, \frac{h}{a} \right)$$

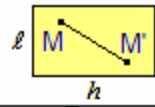
Observación - A menudo, en las aplicaciones, los fenómenos de transición están acompañados de un efecto de pepita: se les representará entonces por un esquema esférico con efecto de pepita C_0 :



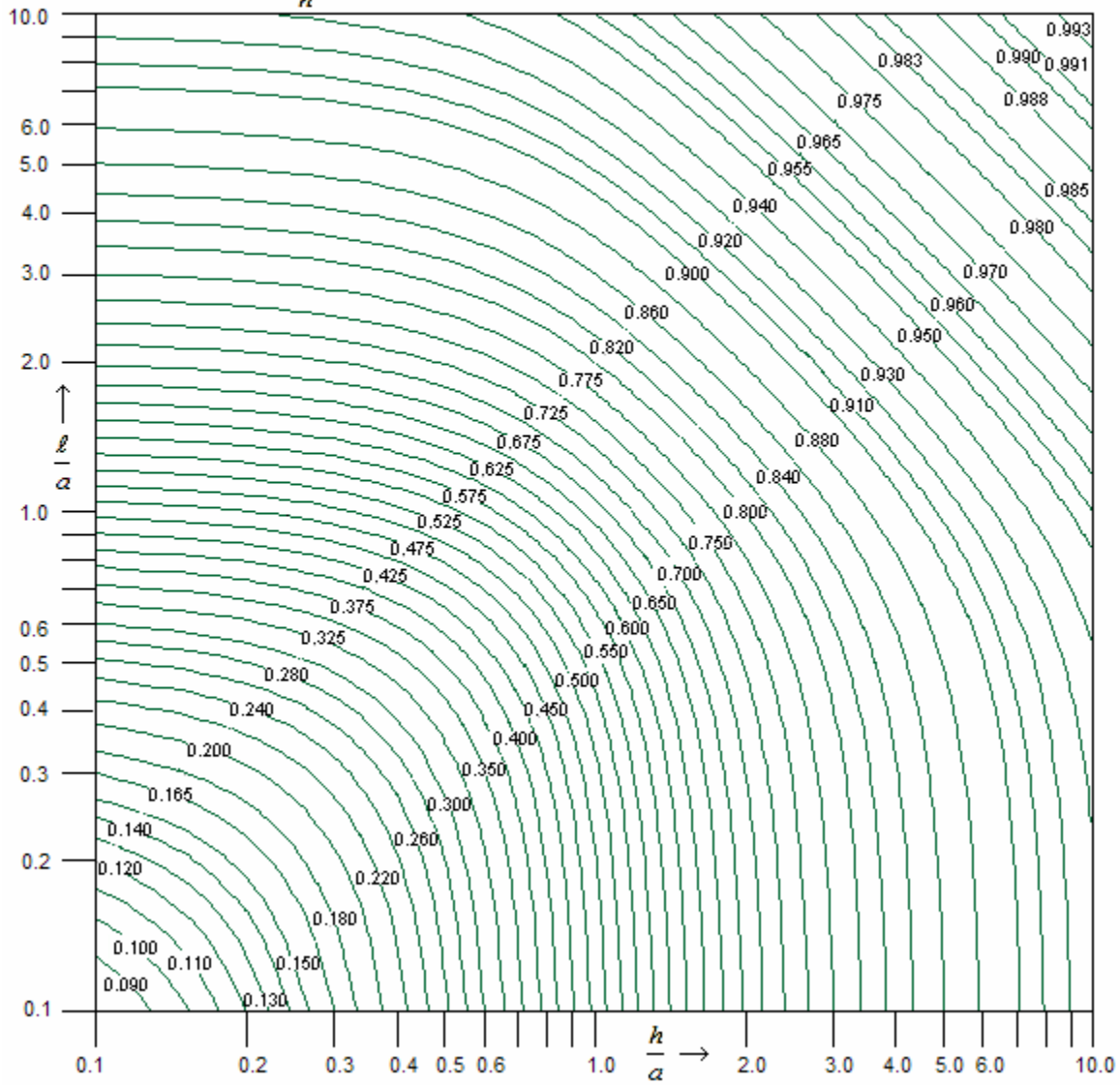


Esquema esférico
 Varianzas de extensión diversas

ESQUEMA ESFERICO



$$\frac{1}{C} F\left(\frac{h}{a}, \frac{l}{a}\right) = \text{varianza de un punto en el rectángulo } l \times h$$



4 - EL KRIGEADO

4-1 EL PROBLEMA DEL KRIGEADO Y SU INTERES.

El kriging consiste en encontrar la mejor estimación lineal posible de la ley de un panel, considerando la información disponible, es decir las leyes de las diferentes muestras que se han tomado, sea al interior, sea al exterior del panel que se quiere estimar. El kriging consiste en efectuar una ponderación, es decir atribuir un peso a la ley de cada muestra, estos pesos se calculan de manera de hacer mínima la varianza de estimación resultante, considerando las características geométricas del problema (formas, dimensiones e implantación relativa del panel y de las muestras). En grueso, como es natural, el kriging atribuirá pesos débiles a las muestras alejadas, e inversamente. Sin embargo esta regla intuitiva puede ser parcialmente puesta en duda cuando aparecen fenómenos más complejos de efecto de pantalla y de transferencia de influencia. No es posible resolver un problema de kriging, es decir calcular efectivamente el peso óptimo que conviene atribuir a cada muestra, sin hacer ciertas hipótesis sobre las características geoestadísticas del depósito que se estudia. En adelante supondremos que el depósito es geoestadísticamente homogéneo, en otras palabras que las leyes, dentro del depósito, consideradas como una variable regionalizada $f(x)$, pueden ser interpretadas como una realización de un esquema intrínseco.

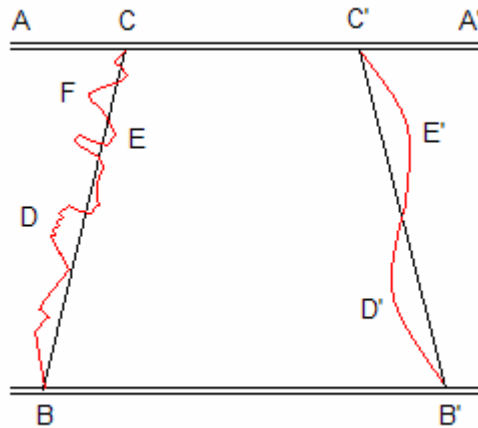
Esta condición de homogeneidad es fundamental. No se puede hacer un kriging entre porciones heterogéneas de un mismo depósito, y más particularmente, en el caso cuando las muestras y el panel a estimar sobrepasan los límites de la mineralización geológicamente homogénea. (Se trata entonces, en sentido estricto, de una dilución y no de un kriging; en general, no es posible, enunciar reglas a priori: los coeficientes de dilución, corrientemente, son determinados por la experiencia).

El primer interés del kriging deriva de su misma definición. Al minimizar la varianza de estimación, estamos seguros de sacar el mejor partido posible de la información disponible, o, si se prefiere, de obtener la estimación más precisa posible del panel en estudio. Esta ventaja es a menudo notable, pero no se justificaría siempre, debido a las complicaciones suplementarias que introduce necesariamente una ponderación. El interés práctico más importante - de lejos - del kriging, proviene, no del hecho que asegura la mejor precisión posible, sino más bien porque permite evitar un error sistemático. En la mayoría de los yacimientos metálicos, se deben seleccionar, para la explotación, un cierto número de paneles, considerados como rentables y se deben abandonar otros paneles considerados no-explotables. D. G. Krige ha mostrado que, si esta selección fuera realizada considerando exclusivamente las muestras interiores a cada panel, resultaría necesariamente - en promedio - una sobre-estimación de los paneles seleccionados. La razón de este problema muy general es que la varianza de las leyes reales de los paneles es siempre más débil que la varianza de los resultados de un muestreo interior. Dicho de otra manera, el histograma de las leyes reales de los paneles comporta menos leyes extremas (ricas o pobres) luego tiene más leyes intermedias que el histograma deducido de las muestras interiores, y, si se calcula el efecto de una selección sobre este último histograma, los paneles

eliminados serán en realidad menos pobres que lo que se había previsto, y los paneles conservados menos ricos.

Nuestra noción de kriging permite interpretar fácilmente este fenómeno, y corregir sus efectos. Cuando se selecciona un panel rico, la aureola de muestras exteriores tiene, en general, una ley más débil que las muestras interiores, sin embargo su influencia sobre el panel a estimar no es despreciable porque el kriging le atribuye un peso no nulo. Si no se toma en cuenta esta aureola exterior, se introduce, necesariamente, una causa de error sistemático por exceso.

Para ilustrar esta noción, imaginemos un yacimiento filoniano reconocido por dos galerías AA' y BB'



Se considera que las leyes de BB' con explotables, mientras que las leyes de AA' son explotables en el segmento CC'. Si nos contentamos con calcular la media de las leyes BB' y CC', estamos seguros de cometer un error por exceso: porque los trozos AC y C'A' - pobres por construcción - tienen una influencia no despreciable sobre el trapecio BB' CC'. Si fuera posible dibujar una frontera precisa entre mineral explotable e inexplorable, en general, la frontera real, no coincidiría con las rectas BC o B'C', sino que sería una curva cualquiera, a menudo bastante irregular, tal como la línea complicada BDEFC. Además en el panel rico subsistirían enclaves pobres y recíprocamente. En la explotación estaremos obligados de abandonar ciertas porciones ricas, y de tomar ciertas partes pobres, y este ensuciamiento traduce de manera concreta la influencia de los trozos pobres AC y C'A' sobre el panel rico (como se trata de un yacimiento homogéneo, y que la frontera precisa de la ley es dudosa, nosotros hablaremos de kriging, y no de ensuciamiento. Emplearemos la palabra ensuciamiento en el caso en el cual los minerales ricos y pobres son geológicamente heterogéneos, y están separados por una frontera continua, representable por una curva relativamente regular, tal como B'D'E'C', que, por otra parte, no tiene ninguna razón para coincidir con la recta B'C'). En este caso el kriging conduce a ponderar las leyes de BB', CC', y de los trozos AC y C'A' por coeficientes convenientes, que serían - por ejemplo - 60% para BB', 27% para CC' y 13% para los dos trozos.

El efecto esencial de esta ponderación es eliminar - en promedio - un error sistemático por exceso, luego un error inquietante. Al considerar este objetivo primordial, la mejora de la precisión propiamente dicha, aparece como secundaria.

4-2. LAS ECUACIONES GENERALES DEL KRIGEADO.

De una manera general, el problema del krigado puede ser formulado de la manera siguiente: sea un depósito homogéneo con un cierto variograma (isótropo o no), en el cual se han tomado n muestras S_1, S_2, \dots, S_n , cuyas leyes (conocidas) son X_1, X_2, \dots, X_n . Se desea encontrar el mejor estimador Z^* , es decir:

$$(21) \quad Z^* = \sum_{i=1}^n a_i X_i$$

De la ley real desconocida Z de un panel P , conociendo la forma, las dimensiones y la implantación relativa de P, S_1, S_2, \dots, S_n . Siempre se puede resolver este problema determinando los coeficientes a_i que cumplen dos condiciones. La primera condición expresa que el error $Z-Z^*$ debe tener una esperanza matemática nula. Esta condición se escribe:

$$(22) \quad \sum_{i=1}^n a_i = 1$$

La segunda condición es una condición de varianza mínima. Esta condición expresa que los coeficientes a_i del krigado tienen valores tales que - teniendo en cuenta la condición imperativa (22) - la varianza $D^2(Z-Z^*)$, la cual mide la magnitud del error posible, tome su valor mínimo. Entonces, teniendo en cuenta la relación (21), esta varianza tiene la expresión:

$$(23) \quad D^2(Z-Z^*) = \sigma_Z^2 - 2 \sum_i a_i \sigma_{zX_i} + \sum_i \sum_j a_i a_j \sigma_{ij}$$

Las notaciones se explican por sí solas. σ_Z^2 es la varianza de la ley real del panel, σ_{zX_i} es la covarianza de Z con la ley X_i de la muestra S_i y σ_{ij} la covarianza entre X_i y X_j (en particular σ_{ii} es la varianza de X_i). Todas estas varianzas y covarianzas se calculan en un gran depósito ficticio, en el cual su dimensión no interviene en realidad: la influencia de este depósito ficticio se manifiesta por una misma constante A que es mayor que todas estas varianzas y covarianzas, y que se elimina, en virtud de (22). Entonces la expresión (23) solo depende del variograma $\gamma(h)$.

Se expresa que la varianza es mínima considerando la condición (22) escribiendo:

$$(24) \quad \begin{cases} \sum_j a_j \sigma_{ij} = \sigma_{zx_i} + \lambda \\ \sum_j a_j = 1 \end{cases}$$

Se obtiene entonces un sistema de $n+1$ ecuaciones lineales con $n+1$ incógnitas (los a_i y el multiplicador de Lagrange λ). Como este parámetro de Lagrange no presenta interés por sí solo, buscaremos eliminarlo. Para ello, se hará jugar un papel particular a una de las muestras, por ejemplo S_n (en la práctica se elegirá una muestra que ocupe una posición privilegiada respecto de las $n-1$ restantes, por ejemplo una posición central) y se escribirá la condición (22) en la forma:

$$(25) \quad a_n = 1 - \sum_{i=1}^{n-1} a_i$$

y se reemplazará en (24) a_n por este valor, es decir:

$$\begin{cases} \sum_{j=1}^{n-1} a_j \sigma_{ij} + a_n \sigma_{in} = \sigma_{zx_i} + \lambda \\ \sum_{j=1}^{n-1} a_j \sigma_{nj} + a_n \sigma_{nn} = \sigma_{zx_n} + \lambda \end{cases}$$

Se elimina λ al restar la última ecuación de las $n-1$ anteriores, y se reemplaza a_n por su valor. Si se pone:

$$(26) \quad \begin{cases} R_{ij} = \sigma_{ij} - \sigma_{nj} - \sigma_{ni} + \sigma_{nn} \\ N_i = \sigma_{zx_i} - \sigma_{zx_n} - \sigma_{in} + \sigma_{nn} \end{cases}$$

el sistema (24) se reduce a:

$$(27) \quad \sum_{j=1}^{n-1} a_j R_{ij} = N_i$$

La solución de este sistema de $n-1$ ecuaciones lineales con $n-1$ incógnitas permite el cálculo efectivo de los n coeficientes de krigeado, a_n se determina según (25).

Los valores así obtenidos se pueden introducir en la expresión (23) de la varianza de estimación, se obtiene entonces el valor mínimo σ_K^2 de la varianza, que llamaremos varianza de krigado. Vemos que esta varianza toma la forma muy simple siguiente:

$$(28) \quad \sigma_K^2 = \sigma_z^2 - 2\sigma_{zx_n} + \sigma_{x_n}^2 - \sum_{i=1}^{n-1} a_i N_i$$

La primera parte de esta fórmula representa la varianza de extensión:

$$\sigma_E^2 = \sigma_z^2 - 2\sigma_{zx_n} + \sigma_{x_n}^2$$

de la muestra S_n en el panel P. Es la varianza de estimación que se tendría si solo se tuviera de la ley X_n de la única muestra S_n . Cada una de las otras muestras S_1, S_2, \dots, S_{n-1} aporta una mejora la cual se traduce por la resta de los términos $a_i N_i$, siempre positivos. Las ecuaciones (27) y (28) constituyen las ecuaciones generales del krigado. Los términos de la matriz R_{ij} y los segundos miembros N_i están dados por (26) y solo dependen de la función intrínseca.

4-3 KRIGEADO COMPLETO y BALANCE MINERAL METAL.

Además de su carácter lineal, las ecuaciones (27) del krigado presentan una particularidad notable, que se observa en las relaciones (26): Los coeficientes de la matriz R_{ij} - es decir los primeros miembros de las ecuaciones (27), solo dependen de las varianzas y covarianzas de las muestras entre ellas, y son independientes del panel a estimar. El panel a estimar interviene solamente en los segundos miembros N_i , por intermedio de las covarianzas σ_{zx_i} de su ley con la de las muestras S_i . Resulta inmediatamente un teorema de superposición de figuras de krigado. Supongamos, en efecto, que se haya krigado, a partir de las mismas muestras S_1, S_2, \dots, S_n dos paneles distintos, de tamaño P y P' y de leyes reales (desconocidas) z y z', es decir que se hayan formado los estimadores:

$$Z^* = \sum a_i X_i, \quad Z'^* = \sum a'_i X_i$$

en que a_i y a'_i son las soluciones del sistema (27), para los valores de los segundos miembros N_i y N'_i correspondientes a los paneles P y P'. Si se desea krigear el panel P+P' constituido por la unión de P con P' (suponiendo que P y P' son disjuntos), cuya ley real es:

$$Z = \frac{P}{P+P'} z + \frac{P'}{P+P'} z'$$

se obtienen los segundos miembros del sistema al ponderar las covarianzas por el tamaño de los paneles:

$$\sigma_{zx} = \frac{P}{P+P'}\sigma_{zx_i} + \frac{P'}{P+P'}\sigma_{z'x_i}$$

es decir efectuando la misma ponderación sobre los segundos miembros. Resulta entonces que el carácter lineal del sistema de ecuaciones implica que los coeficientes A_i del krigado de $P+P'$ son:

$$A_i = \frac{P}{P+P'}a_i + \frac{P'}{P+P'}a'_i$$

y el estimador toma la forma:

$$Z^* = \frac{P}{P+P'}z^* + \frac{P'}{P+P'}z'^*$$

Dicho en otras palabras, existe la superposición de figuras de krigado (con ponderación por los tonelajes). En particular, esta condición se cumple siempre en el caso de un krigado completo - es decir en el caso en que se krigea cada panel con la totalidad de las muestras en el depósito entero. Luego tenemos el teorema: "los krigados completos se pueden superponer".

En particular, si se desea formar el mejor estimador de la ley global de un depósito entero, es lícito ponderar los estimadores obtenidos para cada uno de los paneles, con la restricción que estos hayan sido krigados de manera completa. La estimación panel a panel es entonces equivalente a la estimación global, y el balance mineral y metal queda equilibrado.

En la práctica, casi nunca se hace un krigado completo, debido a la complejidad de los cálculos. Nos limitamos siempre a un número pequeño de aureolas exteriores - las más próximas - que representan la casi totalidad de la influencia de las aureolas exteriores. Como estas aureolas representan una pantalla casi perfecta, el krigado incompleto con el cual nos contentamos, no difiere prácticamente del krigado completo teórico.

El teorema de superposición es sobre todo útil, en el caso de la teoría del krigado continuo, en una forma diferencial. Un panel S a estimar se considera como constituido por la unión de paneles elementales dS infinitamente pequeños. Si $a_i(M)$ es el coeficiente relativo al krigado del elemento dS de centro M , el coeficiente relativo al panel entero se deduce por integración:

$$A_i = \frac{1}{S} \iint_S a_i(M) dS$$

Si se sabe resolver un problema de krigado al nivel puntual, esta fórmula permitirá deducir el krigado de un panel S de forma cualquiera.

De su lado, las ecuaciones del krigado puntual constituyen un caso particular simple de las ecuaciones generales. Sin embargo, escribiremos explícitamente estas ecuaciones, en función del variograma $\gamma(h)$, en el caso en que se desea estimar la ley Z en el punto z a partir de las leyes X_i observadas en los puntos x_i . El estimador es de la forma:

$$Z^* = \sum_{i=1}^n a_i X_i$$

y los coeficientes a_i del krigado verifican el sistema:

$$\begin{cases} \sum_{j=1}^n a_j \gamma(x_i - x_j) = \gamma(z - x_i) + \lambda \\ \sum_{i=1}^n a_i = 1 \end{cases}$$

Al eliminar el coeficiente a_n , se encuentra igualmente:

$$\sum_{j=1}^{n-1} a_j R_{ij} = N_i$$

con:

$$\begin{cases} R_{ij} = \gamma(x_i - x_j) - \gamma(x_n - x_j) - \gamma(x_n - x_i) \\ N_i = \gamma(z - x_i) - \gamma(x_n - z) - \gamma(x_n - x_i) \end{cases}$$

En el caso en que los datos disponibles no son relativos a un número finito de puntos, sino a un conjunto continuo (líneas, superficies o volúmenes), las ecuaciones del krigado se reemplazan por una ecuación integral. Este caso, que constituye lo que se llama el krigado continuo, no lo trataremos aquí (ver [4], [5] o [6]).

5 - INVESTIGACION DEL OPTIMO EN EL RECONOCIMIENTO
Y LA PUESTA EN EXPLOTACION DE LOS DEPOSITOS MINEROS

5-0 INTRODUCCION.

Si se considera la magnitud de las inversiones en la industria minera, y la gravedad del riesgo de ruina, ninguna sociedad, ninguna autoridad responsable, toma la decisión de explotar un nuevo depósito sin haber realizado la estimación más precisa posible de todos los parámetros geológicos, técnicos y económicos que condicionan la rentabilidad de la nueva mina. Estos parámetros son numerosos y variados. Estudios anteriores permiten precisar los factores económicos (precio del mineral). Ensayos de tratamiento en laboratorio, o en planta piloto, y el estudio detallado de diferentes proyectos de explotación, permitirán elegir y cuantificar la mejor solución técnica. Finalmente, los trabajos de reconocimiento por sondajes o labores mineras conducen a una

evaluación de los recursos del depósito, en tonelaje y en ley. Sería deseable conocer perfectamente estos diferentes parámetros. Pero la información cuesta cara. En particular, los trabajos de reconocimiento pueden representar un porcentaje no despreciable del total de las inversiones. Llega un momento en el cual el valor de la información suplementaria ya no está en razón con su costo. Entre estas dos situaciones extremas, caracterizadas, una por gastos de investigación nulos y riesgo de ruina máximo, la otra por una seguridad perfecta y gastos de investigación infinitos, existe necesariamente una situación de compromiso correspondiente a un óptimo. Donde se sitúa este óptimo y como determinarlo, corresponden a las preguntas a las cuales trata de responder esta última sección. En efecto, nos limitaremos al problema de la determinación del volumen óptimo de los trabajos de reconocimiento. Los datos económicos (precios del mineral) se suponen conocidos, igual que las características técnicas del método de explotación y del procedimiento de tratamiento retenidos. Estudiaremos en detalle la influencia de la elección de dos parámetros técnicos muy importantes: el ritmo anual de explotación, y la ley límite de corte.

Queda claro que es difícil plantear el problema. El nivel óptimo de reconocimiento depende de las características del depósito, y no puede ser determinado si no se dispone ahora, de ciertas indicaciones sobre sus características, es decir, en la práctica, si no se ha realizado previamente una primera fase de trabajos de reconocimiento. Se tiene entonces un esquema secuencial, en el cual las decisiones a tomar se presentan en cascada. Al término de la primera fase de reconocimiento, nos encontramos localizados en un punto crucial C (Figura 1):

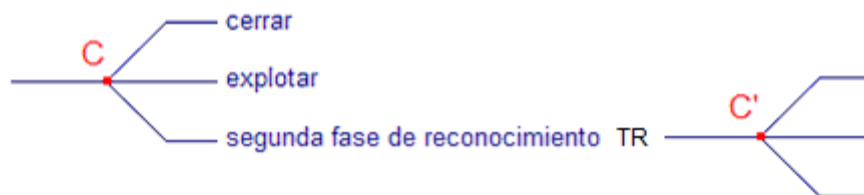


Figura 1: TR=trabajos de reconocimiento

donde se debe elegir una de tres decisiones posibles:

- 1/ Cerrar y abandonar el depósito.
- 2/ Explotar inmediatamente.
- 3/ Hacer una segunda fase de trabajos de reconocimiento.

Si se adopta la tercera decisión, a la salida de esta fase de reconocimiento, nos encontraremos en un nuevo punto crucial C', donde, teóricamente, las tres decisiones son de nuevo posibles, pero pueden ser tomadas basadas en informaciones más completas que en C.

Entre estas tres decisiones, es necesario elegir un criterio: adoptaremos el criterio habitual, el cual consiste en maximizar la esperanza matemática del beneficio futuro. Se elige la decisión que proporciona la mejor esperanza después del punto C. La primera decisión conduce a una esperanza nula. La esperanza matemática de la segunda se calcula haciendo

el balance provisorio del proyecto de explotación construido a partir de la información disponible a la salida de la primera fase. La de la tercera se obtiene anticipando, en probabilidad, los resultados de la segunda fase de trabajos y la decisión a tomar en C' : la esperanza matemática pone en balance la disminución del riesgo de ruina con el costo de los trabajos suplementarios. La Geoestadística proporciona, en cada etapa, una medida objetiva de la información disponible en forma de varianzas de estimación sobre el tonelaje y la ley, razonamientos simples permitirán atribuirles una significación económica para luego calcular numéricamente las diversas esperanzas matemáticas.

Pero esto no es todo. Además del riesgo de ruina, existen otros factores que influyen sobre el volumen óptimo de los trabajos de reconocimiento. En efecto, si se ha tomado la decisión de explotar y el riesgo de ruina ha sido considerado como suficientemente débil, falta preparar un proyecto de explotación, el mejor posible, el cual considera, en particular, la elección óptima del ritmo anual de explotación y de la ley límite. El ritmo óptimo y la ley límite óptima dependen naturalmente de los recursos del depósito. Un error en la evaluación de los tonelajes y de las leyes conduce a un dimensionamiento de la explotación que se aleja del óptimo. Resulta una pérdida, cuyo valor probable puede ser calculado, y debe ser puesto en balance con el costo de los trabajos suplementarios. Aún en el caso de estar seguros de la explotabilidad del depósito, puede suceder que sea interesante hacer más trabajos suplementarios de reconocimiento, no para evitar un riesgo de ruina despreciable, sino más bien para elegir el mejor programa de explotación.

En primer lugar se presenta el esquema secuencial de la figura 1, los problemas de dimensionamiento intervienen después que se ha demostrado la explotabilidad. Sin embargo la exposición teórica debe, necesariamente, seguir el orden inverso, porque una vez que se hayan definido las eventualidades posibles después del punto C , y cifrado sus valores, se pueden calcular las diversas esperanzas matemáticas: es necesario, en primer lugar, profundizar el concepto de explotabilidad, y esto conduce obligatoriamente a estudiar el problema del dimensionamiento óptimo. Luego tenemos el plan de esta sección 6: en una primera parte (6-1), examinaremos la rentabilidad de un depósito cuyos recursos son conocidos, y los problemas asociados a la elección de una ley límite y de un ritmo anual de explotación. En una segunda parte (6-2), se determinará el nivel de reconocimiento óptimo de un depósito en el cual la rentabilidad está asegurada, poniendo en balance el costo de los trabajos y la pérdida debida a un mal dimensionamiento de la explotación. La tercera parte (6-3), finalmente, tomará en cuenta el riesgo de ruina y formulará, de manera precisa, el esquema secuencial de la figura 1. Las dos primeras partes, en realidad, tratarán el problema de la elección de los parámetros técnicos: ritmo de explotación, ley límite y volumen de los trabajos de reconocimiento, estos últimos intervienen en la medida de que sirven para definir el óptimo de los dos primeros. Correlativamente, veremos que los razonamientos deben hacerse con una tasa de actualización nula, lo que conduce al óptimo absoluto. Por el contrario, la tercera parte analiza los criterios de decisión: tiene por objetivo una optimización secuencial, o relativa (relativa a las informaciones disponibles en el punto crucial C), y hace intervenir necesariamente una tasa de actualización diferente de cero. Se encontrará una exposición mucho más detallada en [5] o en [8].

5-1 ELECCION DE UNA LEY DE CORTE Y DE UN RITMO ANUAL DE EXPLOTACION.

5-1-1 Definición de los parámetros (la relación tonelaje/ley).

En esta primera parte, supondremos que las características del depósito se conocen de manera perfecta. En particular, se ha determinado sin ambigüedad el mejor método de explotación y el mejor procedimiento de tratamiento posible. Falta por elegir dos parámetros técnicos esenciales:

- el ritmo anual de producción "t"
- la ley de corte "x"

La significación del segundo parámetro debe ser precisada. Decir que se toma x como ley de corte significa que se decide abandonar los paneles cuya ley media es inferior a x, y explotar aquellos cuya ley es superior a x. Esta noción solo tiene sentido en el caso en que se ha definido el tamaño de los paneles en los cuales se aplicará este corte, es decir el nivel en el cual operará la selección. Este tamaño debe ser considerado como un dato, porque resulta de las características tecnológicas del método de explotación que se ha retenido.

El tonelaje de mineral "T(x)", y la ley media "m(x)" aparecen como dos funciones, una decreciente y la otra creciente, de la ley de corte x. Si

$T(0) = T_0$ designa el tonelaje del depósito entero, la diferencial $-\frac{dT(x)}{T_0}$

representa, en porcentaje del tonelaje total, el tonelaje de mineral cuya ley está comprendida entre x y x + dx. A veces se le interpreta como la densidad de probabilidad f(x)dx de la distribución estadística de las leyes en función de los tonelajes. Esta interpretación no es nada de clara ni esencial. Nosotros no la utilizaremos. La función T(x) representa simplemente el resultado de un método dado de explotación cuando se aplica, con una ley límite x, a un depósito, es decir a un fenómeno natural, determinado.

Se debe tener cuidado, en particular, de no asimilar $T(x)/T_0$ a la distribución acumulada de las leyes de las muestras. La distribución de las muestras no es nunca idéntica a la distribución T(x) de las leyes de los paneles al nivel de los cuales opera la selección. Esta distribución es siempre más dispersa. La geoestadística permite, en general, calcular la varianza de la distribución T(x) de los paneles, a partir de la varianza experimental de las muestras. Si las muestras obedecen, al menos aproximadamente, a una ley de distribución simple (normal, lognormal, etc...), es fácil de reconstituir la función T(x). En ausencia de una ley simple, se podrá construir el histograma de los paneles comprimiendo el histograma de las muestras respecto de su ley media, por una afinidad cuyo módulo es igual a la razón de las desviaciones estándares.

En la práctica, se puede recomendar el método siguiente para la determinación de las funciones m(x) y T(x): respecto del plan de muestreo, adoptar la actitud natural del minero, es decir, elegir y dibujar, para un límite x dado, el modelo de bloques técnicamente posible que podría conducir al mejor resultado previsible, y evaluar la ley media m(x) de los paneles seleccionados, con la ayuda del krigeado (si se verifican las condiciones de homogeneidad), o, en su defecto, con una dilución empírica. Repitiendo la operación para dos o tres valores de x,

se obtendrá una imagen de la variación de $T(x)$ en la zona de leyes de corte útiles, lo que será suficiente para las aplicaciones prácticas.

Cualquiera que sean las dificultades que se encuentran en la práctica para su determinación, las funciones $m(x)$ y $T(x)$ existen, y representan el resultado de la aplicación a un fenómeno natural determinado (el depósito minero) de un procedimiento tecnológico definido (método de explotación). La primera $T(x)$ es decreciente y la segunda $m(x)$ es creciente. A toda ley media m (correspondiente a un corte x) le corresponde un tonelaje bien determinado, cuya ley media es m . Se puede entonces definir una función $T(m)$, la cual proporciona el tonelaje en función de la ley media, y no en función de la ley de corte.

Análogamente, se puede definir una función $m(T)$, la cual proporciona la ley media en función del tonelaje retenido. Se observa que se tiene, necesariamente:

$$(1) \quad x = \frac{d(mT)}{dT} = m + \frac{dm}{d(\log(T))}$$

En efecto, $d(mT)$ representa el tonelaje de metal dQ contenido en la franja marginal dT , cuya ley es x . (En ciertos casos esta relación (1) puede servir para definir el parámetro de corte x . En efecto, puede suceder, que las condiciones de explotabilidad no sean las mismas en las diferentes porciones del depósito. Un panel de igual ley puede ser explotable si está situado en el corazón de una zona rica, o ser inexplorable si está rodeado de zonas pobres y no paga sus costos. En este tipo de casos, no hay una ley de corte universal, pero la función $T(m)$ está siempre definida y, por derivación, la relación (1) permite introducir el parámetro de corte).

La relación $m(T)$ podrá, a menudo, ser representada por una ecuación muy simple:

$$(2) \quad m = \alpha - \beta \log(T)$$

o ley de LASKY. Esta relación empírica no puede ser utilizada en 0 ni en infinito, en este caso se obtienen serias contradicciones. Desde el punto de vista pragmático, esta relación constituye una excelente fórmula de interpolación y permite, en particular, de representar el comportamiento de la curva $m(T)$ entre dos puntos experimentales (conocidos, por ejemplo, a partir de dos proyectos de explotación diferentes). Considerando (2), la ley de corte x se pone en la forma muy simple:

$$(3) \quad x = m - \beta$$

En resumen hemos definido los parámetros técnicos del problema:

- ritmo anual de explotación t ,
- ley de corte x ,
- tonelaje $T(x)$,
- ley media $m(x)$

Falta precisar los parámetros económicos útiles. Entre ellos, distinguiremos principalmente tres: el precio de mercado, los gastos de explotación anuales y el monto de las inversiones.

El precio de mercado, no depende de la voluntad del explotador y puede ser considerado como un dato. Para simplificar lo supondremos conocido y constante. Por otra parte es posible modificar las ecuaciones que propondremos de manera de tener en cuenta la incertidumbre del futuro. Al utilizar la fórmula de venta del concentrado y de la recuperación del tratamiento, el precio de mercado permite definir el valor $V(m)$ del metal recuperable contenido en una tonelada de mineral de ley m .

A menudo se podrá utilizar una fórmula lineal:

$$(4) \quad V(m) = bm - b_0$$

en la cual la constante de castigo b_0 se puede incorporar con el término constante de la expresión $p(t)$ que vamos a definir.

Los gastos anuales por tonelada de mineral representan el conjunto de gastos anuales (explotación, transporte y gastos generales) divididos por la producción anual t . Estos gastos dependen de la producción t . Se les representa por una función $p(t)$, la cual se supone, en general, decreciente. A menudo se puede admitir una relación de la forma:

$$(5) \quad p(t) = a_0 + \frac{a_1}{t}$$

Las inversiones agrupan el conjunto de gastos que es necesario realizar antes de la apertura de la explotación (trabajos mineros preparatorios, planta de tratamiento, eventualmente: campamento, caminos, vías férreas, etc...) Se representan por una función creciente del ritmo t , es decir $I(t)$. Por el contrario, la razón $I(t)/t$ de las inversiones a la producción anual, es una función decreciente de t . En las aplicaciones se puede utilizar una función lineal

$$(6) \quad I(t) = C_0 + C_1 t$$

o bien, más generalmente, una expresión de la forma:

$$(7) \quad I(t) = C_0 + C_1 t^\gamma$$

el exponente γ debe ser positivo pero inferior a 1. Para $C_0=0$ y $\gamma=2/3$, se encuentra la ley empírica, utilizada en varias ramas de la industria, la cual expresa que las inversiones y la capacidad de producción varía, respectivamente, como el cuadrado y el cubo de la dimensión de las instalaciones.

En resumen, utilizaremos las tres funciones económicas siguientes:

- $V(m)$: valor contenido en la tonelada de mineral de ley m
- $p(t)$: gastos de explotación por tonelada de mineral
- $I(t)$: monto de las inversiones

Resulta cómodo incluir un parámetro suplementario, la duración de vida de la mina N, relacionada con el tonelaje $T(m)$ y al ritmo de explotación t por medio de la relación siguiente:

$$(8) \quad N = \frac{T(m)}{t}$$

El resultado de la explotación futura se representa, de manera cómoda, por la expresión del valor actualizado de los beneficios futuros. Designando por i a la tasa de actualización, y por B_i al valor actualizado de los beneficios futuros a esta tasa, se tiene, evidentemente:

$$(9) \quad B_i = [V(m) - p(t)]t \frac{1 - e^{-iN}}{i} - I(t)$$

y, en el caso particular, muy importante en el cual se toma una tasa i igual a 0, el beneficio no actualizado toma la forma simple:

$$(10) \quad B_i = [V(m) - p(t)]T(m) - I(t)$$

En todos los casos, el beneficio puede ser considerado como una función de los parámetros m y t , o si se prefiere (considerando las ecuaciones (1), de x y de t , ley de corte y producción anual. Un criterio posible para guiar la elección de estos dos parámetros consiste en maximizar este beneficio.

5-1-2 Ecuaciones del óptimo.

Conociendo la relación tonelaje/ley, y los diferentes parámetros económicos examinados anteriormente, se debe elegir para la ley de corte x y el ritmo anual t , los valores x_0 y t_0 que maximizan el beneficio futuro. Aquí aparece una primera pregunta: ¿Se debe optimizar el beneficio futuro actualizado B_i de la expresión (9), (y en este caso ¿qué tasa i se debe adoptar?) o, al contrario, el beneficio B no actualizado de la expresión (10)? No podemos entrar aquí a una discusión detallada (ver [5] y [8]) que requiere conocimientos de ciencias económicas. Indiquemos solamente que el uso de una tasa de actualización no nula solamente es legítimo cuando se comparan entre sí operaciones de igual duración. Pero, precisamente, si se comparan dos proyectos de explotación de un mismo depósito los cuales tienen diferentes ritmos de producción anuales, se trata de operaciones de duración diferente, luego, no comparables entre ellos. Para hacer una comparación significativa, se debe pensar en un gran número de depósitos admitiendo las mismas características que el depósito real; se fija un objetivo global, el cual sería, por ejemplo, la producción anual de un tonelaje total dado de mineral; y hacer la comparación de resultados en la hipótesis en que se explotan simultáneamente n_1 depósitos al ritmo t_1 , o n_2 depósitos al ritmo t_2 ($n_1 t_1 = n_2 t_2$). Nos damos cuenta entonces que el óptimo corresponde a la maximización del beneficio no actualizado de la expresión (10).

Para maximizar B , se deben anular las derivadas parciales de B en t (ritmo) y en x (ley de corte). Según (4) y (1) se tiene:

$$(11) \quad \begin{cases} T \frac{dp}{dt} + \frac{dI}{dt} = 0 \\ x = \frac{p}{b} \end{cases}$$

La primera ecuación optimiza el ritmo de producción (con ley de corte x dada). La segunda ecuación muestra que la ley de corte x debe ser elegida según la regla marginalista, según la cual se explota una tonelada de mineral cuando paga sus gastos (las inversiones se amortizan en las partes ricas del depósito).

Este sistema se resuelve fácilmente por aproximaciones sucesivas. En particular, si se adoptan las expresiones (5) y (7) de $p(t)$ e $I(t)$, el sistema queda:

$$\begin{cases} t = \left(\frac{a_1 T(x)}{\gamma C_1} \right)^{\frac{1}{1+\gamma}} \\ x = \frac{1}{b} \left(a_0 + \frac{a_1}{t} \right) \end{cases}$$

Para $\gamma=1$ este sistema proporciona una regla empírica, que se encuentra a veces en la literatura, según la cual el ritmo de explotación anual debe ser proporcional a la raíz cuadrada del tonelaje de las reservas.

5-2 RECONOCIMIENTO OPTIMO PARA EL DIMENSIONAMIENTO DE LA EXPLOTACION

5-2-1 Descripción del problema.

Como lo habíamos indicado al comienzo de este estudio, todavía no introduciremos, en esta segunda parte, el riesgo de ruina, es decir el riesgo, nunca nulo, de poner en explotación un depósito que, en efecto, no sería rentable. Admitiremos que la prueba, o al menos la casi certitud de esta rentabilidad se ha podido establecer. Es necesario entonces disponer de un proyecto de explotación y de tratamiento preciso y definitivo. Los parámetros de dimensionamiento: producción anual t , y ley de corte x , se calculan con la ayuda de las ecuaciones (11). Sin embargo, para resolver estas ecuaciones, no conocemos la verdadera relación tonelaje/ley $T(x)$, sino más bien una estimación de ella $T_E(x)$, la cual es diferente si el reconocimiento ha sido más o menos intenso. Por consiguiente, los parámetros t_E y x_E que vamos a determinar van a diferir de los valores t_0 y x_0 que corresponderían al óptimo para la verdadera función $T(x)$ desconocida. Naturalmente, se puede disminuir esta pérdida efectuando trabajos suplementarios de reconocimiento, los cuales permitirían obtener una mejor estimación de $T(x)$, pero estos trabajos cuestan caro. Existe un óptimo, correspondiente a un volumen de trabajos de reconocimiento tales que la suma de la pérdida y el costo de los

trabajos de reconocimiento sea mínima. Este óptimo tiene el mismo carácter técnico que en el párrafo anterior: los trabajos de reconocimiento se consideran aquí como un parámetro técnico auxiliar, el cual sirve para determinar el tamaño óptimo de la explotación. Entonces se deberá razonar con una tasa $i = 0$, como en la sección 6-1. Debido a que la función aproximada $T_E(x)$ conduce a una solución vecina del óptimo real (t_0, x_0) , esperamos observar una pérdida de segundo orden (lo cual no significa que esta pérdida sea despreciable) cuyo valor esperado podrá expresarse en función de las varianzas de estimación del tonelaje y de la ley.

El problema es relativamente delicado de formular, examinaremos solamente el caso particular simple de los depósitos "todo o nada", para un estudio más completo habrá que consultar la bibliografía ([5], [8]).

5-2-2 Caso de los depósitos de tipo "todo o nada".

Por depósito de tipo "todo o nada" designamos a un depósito, cuya ley no es necesariamente constante, sino en el cual no es posible hacer una selección: o bien se explota todo, o bien no se explota⁽¹⁾. La relación tonelaje/ley no juega ningún papel: en este caso el tonelaje y la ley son impuestos por la naturaleza, y el ser humano no tiene ninguna elección posible. En este caso, el error cometido sobre la ley media no tiene incidencia en el dimensionamiento: el corte se impone, el único parámetro que hay que determinar es la producción anual t , que solo depende del tonelaje T . Si se diferencia la primera ecuación (11), se observa que el error δT de la estimación del tonelaje implica una desviación δt entre el ritmo adoptado y el ritmo óptimo verdadero, desviación cuya expresión es:

$$\delta t = \frac{p' \delta T}{Tp'' + I''}$$

Respecto del beneficio (no actualizado) óptimo, esta divergencia implica una pérdida la cual se obtiene por el desarrollo limitado siguiente:

$$P = -\delta[(V - p)T - I] = (Tp' + I')\delta t + \frac{1}{2}(Tp'' + I'')\delta t^2$$

Según la ecuación (11), el término de primer orden es nulo, porque estamos cerca del óptimo. Al considerar la expresión de δt , se encuentra:

⁽¹⁾ Este tipo de depósitos, se encuentra en la práctica, más a menudo de lo que se cree. Un depósito todo o nada no está, obligatoriamente, rodeado de material estéril. Puede estar rodeado de una franja mineralizada no explotable, geoestadísticamente heterogénea con el mineral. Se observa entonces el fenómeno de umbral de ley, que los mineros conocen bien: los cálculos de tonelaje hechos con diferentes leyes de corte conducen, prácticamente, a resultados idénticos, y, en el plano, los límites entre paneles explotables e inexplotables prácticamente no se desplazan.

$$P = \frac{1}{2} \frac{p^2}{Tp + I} (\delta T)^2$$

En valor esperado, $(\delta T)^2$ proporciona la varianza σ_T^2 de la estimación del tonelaje, cuyo valor se puede calcular geoestadísticamente, en función de los trabajos de reconocimiento. Se tiene, entonces:

$$(12) \quad P = \frac{1}{2} \frac{p^2}{Tp + I} \sigma_T^2$$

Esta pérdida P debe ser puesta en balance con el costo R de los trabajos de reconocimiento. R es una función decreciente de σ_T^2 . Se obtendrá el volumen óptimo de reconocimiento al resolver la ecuación:

$$(13) \quad \frac{dR}{d\sigma_T^2} + \frac{1}{2} \frac{p^2}{Tp + I} = 0$$

Que expresa que la suma $P * R$ es mínima. En la práctica se expresará R y σ_T^2 en función de un mismo parámetro (número n de sondajes, distancia entre galerías de reconocimiento, etc...), o, eventualmente, de varios. Sean u_i estos parámetros, se debe resolver el sistema:

$$(14) \quad \frac{\partial R}{\partial u_i} + \frac{1}{2} \frac{p^2}{Tp + I} \frac{\partial \sigma_T^2}{\partial u_i}$$

En el caso en que se tiene, efectivamente, varios parámetros u_i (por ejemplo, en el reconocimiento de un filón por trabajos mineros, la distancia entre los niveles y la malla de canaletas), la solución del sistema (14) proporciona a la vez el volumen global óptimo de gastos que se deben destinar al reconocimiento, y la mejor distribución de estos entre las diferentes categorías de trabajos (niveles y canaletas).

A modo de ilustración, imaginemos un gran depósito sedimentario de hierro, de forma aproximadamente elíptica o rectangular, reconocido por una red de sondajes con malla adaptada, es decir implantados en los vértices de una malla rectangular, con una razón de malla igual a la razón de las dimensiones del depósito. La geoestadística nos indica que la varianza relativa σ_T^2/T^2 de la estimación del tonelaje es la suma de dos términos: el primero representa el efecto de borde, o si se quiere, el error de estimación de la superficie mineralizada, el cual admite una expresión de la forma

$$\frac{0.244}{n^{3/2}} \sqrt{\frac{D_1 D_2}{S}}$$

siendo n el número de sondajes útiles, S el área mineralizada, D_1 y D_2 los diámetros (más generalmente, las variaciones diametrales si el área S no es convexa) del área mineralizada. Para un contorno elíptico se toma

$$\sqrt{\frac{D_1 D_2}{S}} = \frac{2}{\sqrt{\pi}} = 1.13$$

y el primer término se reduce a $0.276/n^{3/2}$. El segundo término representa el error de la potencia media y admite una expresión de la forma A/n , donde A es ligeramente inferior a la varianza relativa de las potencias en el depósito, y puede ser considerada como una constante (en realidad es una función que decrece lentamente con n). Tomaremos un valor razonable, como $A=0.10$.

En resumen, la varianza relativa tiene el valor:

$$\frac{\sigma_T^2}{T^2} = \frac{0.276}{n^{3/2}} + \frac{A}{n}$$

Respecto de los valores económicos, tomaremos para $p(t)$ e $I(t)$ expresiones de la forma (5) o (6), y designaremos por C el costo unitario de sondaje. La suma que se desea minimizar, de la pérdida y del costo de reconocimiento, se escribe, en función del número n de sondajes como:

$$\frac{1}{4} \sqrt{a_1 C_1 T} \left[\frac{0.414}{n^{5/2}} + \frac{A}{n^2} \right] = C$$

Numéricamente, examinemos el caso en que $T = 1000$ millones de toneladas, $a_1 = 3 \cdot 10^6$, $c_1 = 10$ (ejemplo tomado de F. Blondel, [9]), $A=0.10$ y un costo unitario C de 2000 dólares por sondajes. Se encuentra $n = 60$ aproximadamente.

5-3 Optimización secuencial y término del reconocimiento.

5-3-1 El problema del término del reconocimiento.

Conviene ahora re-introducir el riesgo de ruina. En el capítulo anterior, hemos admitido que la rentabilidad del depósito había sido establecida, y habíamos determinado el mejor nivel de reconocimiento, tomando en cuenta solamente la pérdida por un mal dimensionamiento de las instalaciones. Se trataba de un óptimo puramente técnico, definido sin actualizar. El problema que abordamos ahora presenta la mayor importancia para la práctica del reconocimiento minero. Se resume en la pregunta fundamental: ¿En qué momento conviene detener el reconocimiento, y tomar una decisión - positiva o negativa - respecto de la explotación del depósito?. El problema no puede formularse de manera absoluta. Nos debemos contentar con el análisis de los elementos que dirigen la decisión relativamente al paso de una fase de la investigación a la fase siguiente. Tomando de nuevo el esquema secuencial de la figura 1, supongamos que se ha ejecutado una primera fase de trabajos de reconocimiento TR_1 . Esta fase ha proporcionado indicaciones cifradas sobre el tonelaje y la ley del depósito reconocido, y, los métodos de la geoestadística han permitido fijar la precisión según la cual estas informaciones representan la realidad. Por otra parte, un estudio de explotabilidad, ha podido definir para qué tonelajes y cuáles leyes, la explotación podría ser viable. En el punto crucial C de la figura 1 al cual hemos llegado, se tienen tres decisiones posibles:

1. Se decide no explotar y abandonar el depósito (se cierra)
2. Se decide explotar inmediatamente ⁽²⁾
3. Se decide hacer una segunda fase de reconocimiento TR₂.

Si se adopta la tercera de las decisiones, nos re-encontraremos, después de la ejecución de TR₂, en un nuevo punto crucial C', donde, de nuevo, en teoría, las tres decisiones son posibles, pero pueden ser tomadas a la luz de una información mejorada. En la práctica, en la mayoría de los casos, no será posible, de considerar una tercera fase TR₃ de reconocimientos. Nos limitaremos, en este estudio, al caso en que en C', la única elección posible consistirá en poner o no, el depósito en explotación.

Este problema secuencial se formula, necesariamente, de manera discontinua: En particular, la segunda fase de reconocimiento TR₂, considerada en la tercera decisión, será impuesta por la tecnología. Por ejemplo, si se ha reconocido un depósito tipo filón por galerías equidistantes de 60m, la segunda fase consistirá en trazar niveles intermedios, de manera de llevar a 30 metros la distancia entre niveles. En el caso de un depósito estratiforme reconocido por sondajes con malla cuadrada, la segunda fase consistirá en centrar la malla, es decir doblar el número de sondajes útiles. Esta segunda fase solo depende de un parámetro el cual varía en forma continua, y cuyo valor podría ser elegido arbitrariamente ⁽³⁾. Este valor se impone.

A veces se tendrá la ilusión de que una elección continua es posible. Por ejemplo, si se ha reconocido un depósito filoniano con tres niveles, y si se han muestreado estos tres niveles por sondajes percutantes horizontales, tomados de techo a muro en la formación, se podría pensar en mejorar la información disponible cerrando la malla de los percutantes sin hacer trabajos mineros suplementarios. La malla de los percutantes aparece así como un parámetro casi continuo. Pero se sabe que la mejora aportada por los percutantes suplementarios sería casi siempre

⁽²⁾ Si no se actualiza, es indiferente explotar ahora o más tarde. Si se actualiza, es preferible explotar inmediatamente, salvo en el caso en que los precios de hoy están a un nivel anormalmente bajos y se espere su repunte en los años a venir. Como, en este estudio, no tomamos en cuenta la variabilidad de los precios, no consideraremos la solución de espera, cuya formulación sobre bases realistas es bastante delicada.

⁽³⁾ Supondremos, esencialmente, que el reconocimiento del depósito ha sido exhaustivo. Nos hemos podido equivocar en la evaluación del tonelaje y de la ley, pero la totalidad del depósito ha sido reconocido. Los trabajos de reconocimiento han sobrepasado los límites de la mineralización en el espacio, y no hay problemas de extensión. Por otra parte, no es posible tratar matemáticamente un problema de extensión. Un filón puede proplongarse más lejos del último nivel de reconocimiento, y la geología puede aportar argumentos en favor de esta extensión: pero estos argumentos no son cifrables. Ni la geoestadística, ni ningún artificio probabilístico pueden reemplazar una información ausente. En el caso de los filones, es, en efecto, muy raro que los trabajos mineros alcancen el término del depósito. Los cálculos de rentabilidad que se hacen en la práctica, no consideran un bloque valor posible más allá del último nivel. El tonelaje efectivamente reconocido debe justificar, por si mismo, la puesta en explotación. Haremos lo mismo en nuestro análisis teórico.

despreciable, la parte más grande de la varianza de estimación, la cual proviene de la extensión de las secciones (que se supone conocidas) a sus rebanadas de influencia a tres dimensiones, extensión sobre la cual la mayor densidad de percutantes queda sin acción.

La geoestadística permite calcular la precisión con la cual las reservas se conocen a la salida de los trabajos TR_1 , pero también permite predecir, de antemano, la nueva precisión cuando se habrá ejecutado la segunda fase TR_2 , tal como la tecnología lo impone. Para formular el problema secuencial, conviene atribuir un valor económico al suplemento de información aportado por los TR_2 , y comparar con el costo de los TR_2 mismos. Entre todas las decisiones posibles, se deberá elegir aquella que proporcione el mayor valor probable del beneficio futuro ⁽⁴⁾ en el punto C (es decir, sin tener en cuenta los gastos anteriores, pero teniendo en

⁴ Dejaremos voluntariamente abierta la pregunta de saber si el beneficio debe, o no, ser actualizado. En una economía colectivista, sin duda que no se actualizaría. En economía competitiva, el razonamiento de la primera parte, aún se aplica, al menos teóricamente: una gran sociedad, que se fija como objetivo de producir cada año una cantidad dada de metal al menor costo, debería buscar el mínimo de la suma de los gastos anuales los cuales contienen los gastos de explotación, las inversiones para reemplazar los yacimientos que llegan a su agotamiento ese año, y los trabajos de reconocimiento para preparar el reemplazo de los depósitos que se agotarán los años próximos. Este punto conduciría, en buena lógica, a adoptar la expresión no actualizada del beneficio en la optimización secuencial de los trabajos de reconocimiento (por otra parte, correlativamente, el parámetro b no correspondería al precio de mercado, sino más bien a un precio interno, definido, por la Sociedad de manera, precisamente, de optimizar su programa de producción global).

Este razonamiento nos parecía peremptorio, en el caso del estudio de la optimización técnica realizada en las dos primeras partes, porque conducía a una descripción correcta del comportamiento de los explotadores. Al contrario, en la optimización secuencial, el comportamiento real de los explotadores en economía competitiva, corresponde a una tasa no nula. En efecto, se busca, un criterio de decisión relativo a una puesta en explotación eventual, y - en economía competitiva - se decide explotar un depósito cuando el valor actualizado del beneficio futuro es positivo. Entonces utilizaremos la expresión (9), la cual corresponde a un beneficio correspondiente a una tasa i no nula. Será extremadamente fácil pasar al caso $i=0$ al reemplazar en los resultados obtenidos, las exponenciales e^{-iN} por la unidad, y las expresiones $(1-e^{-iN})/i$ por N .

Los economistas dan, con razón, gran importancia a la elección del criterio adoptado. Además de la expresión (9) del beneficio, actualizado o no, otros criterios pueden ser utilizados, los cuales conducen a prácticas más o menos malthusianas y a diversas variedades de "floreo". Por ejemplo el punto de vista estrictamente economicista se fijaría como objetivo maximizar la razón B_i/I . Igualmente, el punto de vista, un poco ingenio del técnico puro conduciría a minimizar los costos del kilo de metal, etc... De una manera general, J. Lesourne, [10] indica que las razones, tales como B/I , constituyen malos criterios económicos. En el problema que nos ocupa, la única elección posible consiste en hacer, o no hacer, $i=0$, en la expresión (9) del beneficio.

cuenta, para la tercera decisión, el costo de la segunda fase de reconocimientos TR_2 , dado que en C, TR_2 no ha sido ejecutado).

5-3-2 La relación de aditividad.

Supongamos que, después de una primera fase de trabajos, hemos llegado al punto crucial C de la figura 1. Disponemos entonces de una primera estimación Z^* de la magnitud Z que nos interesa (tonelaje o ley del depósito). Si decidimos hacer una segunda fase de trabajos de reconocimiento, estos trabajos no serán implantados al azar, sino que ocuparán una posición geométrica preferencial relativa a los trabajos de la primera fase (se centrará la malla de los sondajes, se trazará un nivel suplementario en la altura media entre dos galerías existentes, etc...). Designemos por Y^* la estimación a la cual nos conducirían los trabajos de la segunda fase si ellos existieran solos (es decir si los trabajos de la primera fase no existieran), e introduzcamos las varianzas y la covarianza siguientes:

$$\begin{cases} \sigma_1^2 = D^2(Z_1 - Z) \\ \sigma_y^2 = D^2(Y^* - Z) \\ \sigma_{1,y} = E[(Z_1 - Z)(Y^* - Z)] \end{cases}$$

las cuales solo dependen del variograma y de la geometría de los trabajos. A la salida de la segunda fase (si decidimos ejecutarla), dispondremos en realidad de los resultados de las dos fases de trabajos, entonces formaremos el estimador Z_2^* que realiza el krigeado de Z a partir de Z_1^* y de Y^* es decir:

$$Z_2 = \lambda Y^* + (1 - \lambda) Z^*$$

La varianza de estimación de Z por Z_2^* es:

$$D^2(Z_2^* - Z) = \lambda^2 \sigma_y^2 + 2\lambda(1 - \lambda)\sigma_{1,y} + (1 - \lambda)^2 \sigma_1^2$$

se minimiza para:

$$\lambda = \frac{\sigma_y^2 - \sigma_{1,y}}{\sigma_1^2 + \sigma_y^2 - 2\sigma_{1,y}} = \frac{\sigma_y^2 - \sigma_{1,y}}{D^2(Z_1^* - Y^*)}$$

y vale entonces:

$$(15) \quad \sigma_2^2 = D^2(Z_2^* - Z) = \sigma_1^2 - \lambda^2 D^2(Z_1^* - Y^*)$$

Introduzcamos la varianza:

$$D^2(Z_1^* - Z_2^*) = \lambda^2 D^2(Z_1^* - Y^*)$$

que representa el error que se comete al estimar el futuro estimador Z_2^* (el estimador que se formará una vez efectuada la segunda fase) con la ayuda del resultado Z_1^* de los primeros trabajos. La relación (15) se escribe:

$$(16) \quad D^2(Z_1^* - Z_2^*) = \sigma_1^2 - \sigma_2^2$$

Así, la varianza $D^2(Z_1^* - Z_2^*)$ representa exactamente la ganancia de información que deben aportar los nuevos trabajos, y la geoestadística permite calcularla antes que la segunda fase sea realizada (debido a que solo depende del variograma y de la geometría de los trabajos). Entonces, basta con encontrar el equivalente económico de esta ganancia de información para realizar la optimización secuencial: se efectuará, o no, la segunda fase de trabajos según que el valor económico del suplemento de información que esta fase aporte sobrepase, o no, el costo de su ejecución.

En la bibliografía ya citada, se encontrará la formulación de este problema en el caso general en el cual intervienen los dos, el tonelaje y la ley. Nos contentaremos aquí de tratar el caso simple en el cual la ley interviene sola.

5-3-3 Caso en el cual solamente interviene la ley

Sea el caso en el cual el tonelaje T del depósito puede ser considerado como conocido con una aproximación suficientemente buena como que la rentabilidad solo dependa de la ley media (desconocida) Z del depósito: Si m_0 representa la ley límite (límite de rentabilidad, y no ley de corte: el depósito será explotado o no según que se tenga $Z \geq m_0$ o $Z < m_0$), el beneficio futuro (si se explota) puede ponerse en la forma:

$$B = bT(Z - m_0)$$

Designemos por B_1 , B_2 , B_3 el beneficio correspondiente a cada una de las tres decisiones posibles: abandonar el depósito (B_1), explotar inmediatamente (B_2), postergar la decisión y ejecutar una segunda fase de trabajos de reconocimiento (B_3).

En los dos primeros casos, se tiene, evidentemente:

$$(17) \quad \begin{aligned} E(B_1) &= 0 \\ E(B_2) &= bT(Z_1^* - m_0) \end{aligned}$$

Se ve que - como era de esperar - que si la elección solo se limitara a las dos primeras opciones, se explotaría o no, según que los trabajos de la primera fase proporcionaran un resultado Z_1^* , mayor o no, que la ley límite m_0 .

El cálculo de $E(B_3)$ es un poco más complejo. Pongámonos imaginariamente en el punto crucial C' , y anticipemos los resultados que deben proporcionar los trabajos de la segunda fase. En C' , se decidirá explotar si $Z_2^* \geq m_0$ (los trabajos de la segunda fase ya fueron realizados, su

costo no debe influenciar esta decisión), y la esperanza del beneficio, (después de C') será entonces:

$$bT(Z_2^* - m_0)$$

Si al contrario, se tiene

$$Z_2^* < m_0$$

se abandonará el depósito y la esperanza será nula. Designemos por:

$$G(z) = P(Z_2^* < z)$$

la probabilidad (evaluada en el punto crucial C) para que los trabajos de la segunda fase (no realizada aún) conduzcan a una estimación Z_2^* inferior a z , y por R el costo de los trabajos de la segunda fase (que no son aún ejecutados en C, y deben entonces ser considerados en la contabilidad). Se ve que se tiene:

$$(18) \quad E(B_3) = bT \int_{m_0}^{\infty} (z - m_0) dG(z) - R$$

La comparación de $E(B_1)$, $E(B_2)$, $E(B_3)$ - fórmulas (17) y (18) - permiten entonces adoptar la mejor decisión en el punto C.

Observación. En realidad no se conoce la ley de probabilidad $G(z)$ de la variable Z_2^* . Se sabe solamente que en el punto crucial C Z_2^* admite la esperanza Z_1^* y la varianza $D^2(Z_2^* - Z_1^*)$ la cual se sabe calcular por los métodos de la geoestadística: esta varianza es justamente la que, según (16), mide la ganancia de información que aportaría la segunda fase de trabajos. En las aplicaciones, se podrá obtener siempre un orden de magnitud numérico, suponiendo, por ejemplo, que $G(z)$ es una ley de Gauss (o una ley lognormal, o cualquier otra ley de probabilidad que dependa de dos parámetros) la cual tiene una esperanza Z_1^* y una varianza $D^2(Z_2^* - Z_1^*)$.

Ejemplo: Para ilustrar estas diversas fórmulas, nos daremos un ejemplo numérico. Un depósito de Plomo-Zinc ha sido reconocido por trabajos mineros y de superficie. El tonelaje es de 400.000 toneladas. Es posible que exista un bloque valor pero no se tomará en cuenta. El valor fino equivalente del Pb, del Zn y del Ag contenidos en una tonelada de mineral ha sido estimado como ⁽⁵⁾ a 3,000 unidades monetarias, de manera que estamos exactamente en el caso marginal. Considerando la incertidumbre de los precios futuros, haremos los cálculos para tres valores diferentes de m_0 , es decir 2,700, 3,000 y 3,300. La ley media se conoce con una desviación estándar $\sigma_1=0.15$. Una segunda campaña de reconocimiento, la cual consistiría en poner un nivel intermedio, costaría 40 millones, y

⁽⁵⁾ Como hay varios metales, es más simple razonar directamente sobre los valores finos contenidos, y no sobre las leyes.

proporcionaría una nueva ⁽⁶⁾ desviación estándar $\sigma_2=(1/2)\sigma_1$. Se deduce entonces la desviación estándar σ de la estimación de m_2 a partir de m_1 :

$$\sigma^2 = \sigma_2^1 - \sigma_1^2 = \frac{\sqrt{3}}{2} \sigma_1^2 = 0.13$$

El cálculo de las esperanzas (17) y (18) para los tres límites de explotabilidad conduce a los resultados numéricos siguientes (en unidades monetarias), razonando en el caso de una ley lognormal:

m_0	2,700	3,000	3,300
$E(B_1)$	0	0	0
$E(B_2)$	+120	0	-120
$E(B_3)$	136-40=+96	56-40=+16	20-40=-20

Si la ley límite se fija en 2,700 unidades monetarias, se debe explotar inmediatamente. Si es de 3,000, es decir exactamente el caso marginal, existe una ventaja, teóricamente no nula, pero numéricamente bastante débil (16 millones), al hacer una segunda campaña de reconocimiento. Esta circunstancia es bastante general: en el caso exactamente marginal la información aportada por los trabajos suplementarios presenta el máximo de interés. Este interés decrece bastante rápido cuando nos alejamos del caso marginal por exceso o por defecto. Cuando estamos prácticamente seguros que la ley real es superior (o inferior) a la ley límite de rentabilidad, se decide explotar (o cerrar), y la necesidad de hacer nuevos trabajos no se hace sentir.

⁽⁶⁾ Se supone que estamos en la zona lineal del variograma, de manera que la nueva varianza es igual a un cuarto de la antigua.

6 EJERCICIOS

6-1 EJERCICIOS SOBRE LOS METODOS TRANSITIVOS.

Ejercicio 1. (Covariograma exponencial). Sea en el espacio de una dimensión, una V. R. cuyo covariograma transitivo es $g(h) = e^{-\lambda|h|}$. Formar la expresión exacta de la varianza de estimación $\sigma^2(a)$, encontrar su parte principal, y concluir, en ausencia de Zitterbewegung.

$$\text{(solución: } \sigma^2(a) = a \left(\frac{2}{1 - e^{-\lambda a}} - 1 - \frac{2}{\lambda a} \right) \approx \frac{1}{6} \lambda a^2 \text{)}$$

Ejercicio 2. (Covariograma triangular). Sea en el espacio de una dimensión la V. R. $f(x)$ igual a 1 en el intervalo $(0, b)$.

1/ Formar el covariograma de $f(x)$. (solución: $g(h) = b - |h|$ para $|h| < b$ y 0 para $|h| \geq b$; establecer este resultado por un razonamiento geométrico).

2/ Calcular la varianza de estimación para una malla a pequeña, mediante la fórmula de aproximación. (solución: $(1/6)a^2$).

3/ Se dispone de $n+2$ sondajes $S_0, S_1, S_2, \dots, S_n, S_{n+1}$ en una malla regular, en la cual el primer y el último sondaje son negativos y los otros positivos. Se admite que cada una de las dos extremidades del intervalo del cual se quiere estimar la longitud b puede caer en cualquier parte del segmento (S_0, S_1) y (S_n, S_{n+1}) respectivamente, de manera independiente una de la otra, con una densidad uniforme. La longitud b es entonces una variable aleatoria. Probar que $E(b) = na$; $D^2(b) = a^2/6$: se encuentra bien la varianza de estimación tal como la proporciona la fórmula de aproximación, pero no se obtiene el término fluctuante.

d/ Calcular el valor exacto de la varianza de estimación de b /

(Solución: poner $b = na + \varepsilon a$ con $0 \leq \varepsilon < 1$. Si la primera muestra positiva es en x ($0 \leq x < a$), se tiene la estimación $(n+1)a$ para $x \leq \varepsilon a$, y na para $\varepsilon a < x < a$. El punto x está implantado al azar en el segmento $(0, a)$ con una densidad uniforme de probabilidad. Esta estimación es una variable aleatoria de esperanza b y de varianza $(\varepsilon - \varepsilon^2)a$. Esta expresión considera el Zitterbewegung. Cuando ε (el cual, en la práctica, es siempre desconocido) es asimilado a una variable aleatoria, con distribución uniforme en $(0, 1)$, el valor medio de esta expresión coincide con la expresión aproximada de b / (en la cual se excluye en Zitterbewegung).

Ejercicio 3. (Varianza de estimación en el caso aleatorio) - Sea, en el espacio de dos dimensiones una V.R. $f(x)$ de covariograma $g(h)$ y de soporte S . Si el rectángulo de malla (a_1, a_2) es grande respecto de S , a lo más, un solo sondaje es positivo, y todo sucede como si esta muestra única fuera sorteada al azar dentro de un rectángulo de lados a_1, a_2 conteniendo S . Si la muestra cae en un punto $x \in S$, se toma el estimador $Q^*(x) = a_1 a_2 f(x)$, y 0 si $x \notin S$. $Q^*(x)$ es así una variable aleatoria. Probar directamente que:

$$E(Q^*) = \int f(x) dx$$

$$E((Q^*)^2) = a_1 a_2 \int (f(x))^2 dx$$

$$D^2(Q^*) = a_1 a_2 g(0) - \int g(h) dh$$

Ejercicio 4. (covariograma transitivo de la esfera). Encontrar el covariograma $K(h)$ de una esfera de diámetro D en R^3 .

(Solución:
$$K(h) = \frac{\pi D^3}{6} \left(1 - \frac{3|h|}{2D} + \frac{1|h|^3}{2D^3} \right) \quad \text{para } h \leq D.$$

Partir de la relación:

$$K'(h) = \frac{\pi}{4} (D^2 - h^2)$$

e integrar, observando que $g(0) = \frac{\pi}{6} D^3$. Aplicación a la estimación del volumen de esta esfera.

6-2 EJERCICIOS SOBRE LAS F. A. INTRINSECAS.

Ejercicio 1. Se elige al azar un origen x_0 en $(0, a)$, luego se divide la recta en segmentos de longitud a mediante los puntos de subdivisión $x_0 + ka$ (k entero, positivo o negativo). Sea $Y(x)$ la F.A. que toma en cada uno de estos segmentos de longitud a un valor aleatorio constante Y , sorteado al azar de manera independiente de un segmento a otro, según una ley de probabilidad de media m y de varianza σ^2 . Probar que la probabilidad de que dos puntos x y $x+h$ pertenezcan al mismo segmento es $1 - |h|/a$ para $|h| \leq a$ y 0 para $|h| > a$. Deducir el variograma de $Y(x)$ (Solución: $|h|\sigma^2/a$ para $|h| \leq a$ y σ^2 para $h > a$).

Ejercicio 2. La misma pregunta que en el ejercicio precedente, pero las longitudes de los segmentos sucesivos siguen la misma ley exponencial $\exp(-\lambda l)$ (dicho de otra manera, los puntos de discontinuidad constituyen un proceso de Poisson)

(Solución: la probabilidad para que x y $x+h$ estén en el mismo segmento es $\exp(-\lambda h)$. Se deduce: $\gamma(h) = \sigma^2 (1 - \exp(-\lambda h))$).

Ejercicio 3. Se considera el variograma h^λ ($0 < \lambda < 2$) en el espacio de una dimensión. Calcular las funciones auxiliares χ y F , la varianza de extensión elemental de un punto a su segmento de influencia (dispositivo centrado), la varianza de extensión a este segmento del promedio de sus dos extremidades (dispositivo cerrado). Comparar y discutir.

$$\chi(h) = \frac{h^\lambda}{\lambda+1} ; F(h) = \frac{2h^\lambda}{(\lambda+1)(\lambda+2)}$$

$$\frac{2a^\lambda}{\lambda+1} \left[\frac{1}{2^\lambda} - \frac{1}{\lambda+2} \right] ; a^\lambda \left[\frac{2}{\lambda+2} - \frac{1}{2} \right]$$

Se tiene: $\sigma_{E'}^2 > \sigma_E^2$, para $\lambda > 1$ e inversamente, con igualdad para $\lambda = 1$: para $\lambda > 1$, alta continuidad y la muestra única pero bien ubicada es superior a las dos muestras mal ubicadas; para $\lambda < 1$, estamos próximos del caso aleatorio puro, y las dos muestras, a pesar de estar mal ubicadas, proporcionan siempre mejores resultados que la muestra única; para $\lambda = 1$ los dos dispositivos son equivalentes.

Ejercicio 4. - Con el variograma h^λ en el espacio de una dimensión, calcular la varianza de estimación del segmento $L=na$, por n muestras puntuales, en los tres casos siguientes:

- malla regular (dispositivo centrado)

$$\frac{1}{n} \frac{2a^\lambda}{\lambda+1} \left[\frac{1}{2^\lambda} - \frac{1}{\lambda+2} \right]$$

- malla aleatoria estratificada

$$\frac{1}{n} \frac{2a^\lambda}{(\lambda+1)(\lambda+2)}$$

- malla aleatoria pura

$$\frac{1}{n} \frac{2L^\lambda}{(\lambda+1)(\lambda+2)}$$

Ejercicio 5. Sea, en el espacio de dos dimensiones, el variograma r^λ . Sea un depósito reconocido líneas de longitudes superiores a su equidistancia h .

a/ Calcular la varianza de extensión de una línea de largo l en su rectángulo $l \times h$ de influencia.

(Solución: la subida bajo potencia constante l transforma r^λ en $Ar^{\lambda+1}/l$ (o a $\log(r)$ en $\pi r/l$); de donde:

$$\sigma_E^2 = A \frac{2}{\lambda+2} \left[\frac{1}{2^{1+\lambda}} - \frac{1}{\lambda+3} \right] \frac{h^{1+\lambda}}{l} = B \frac{h^{1+\lambda}}{l}$$

b/ deducir la varianza de estimación (ponderar por los cuadrados de las longitudes de las líneas, poner $L = \sum l_i$ y $S = Lh$, poner entonces la varianza de estimación en la forma:

$$B \frac{S^{1+\lambda}}{L^{2+\lambda}}$$

Ejercicio 6. Mismo problema que en 5, pero se supone que además las líneas están estimadas a partir de muestras (canaletas) a la malla a . Calcular el término de línea.

(Solución:

$$\frac{C}{L} a^{1+\lambda} \text{ con } C = \frac{2}{\lambda+1} \left[\frac{1}{2^\lambda} - \frac{1}{\lambda+2} \right]$$

Optimización económica: sea w el costo del metro de galería, p_0 el costo de una muestra, y $l_0 = p_0/w$. Optimizar la precisión si el presupuesto de reconocimiento es fijo, o, lo que es lo mismo, optimizar el costo de reconocimiento si la precisión es fija.

(Solución: Todo consiste en minimizar $Ca^{1+\lambda} + Bh^{2+\lambda}$ con la condición $1/(hl_0) + 1/(ha) = C$. Se obtiene la relación siguiente:

$$B(2+\lambda)h^{1+\lambda} = \lambda Ca^{1+\lambda} + C(1+\lambda) \frac{a^{2+\lambda}}{l_0}$$

que proporciona la equidistancia h de los niveles en función de la malla de canaletas.

Ejercicio 7: Efecto de pepita al estado puro - Sean pepitas puntuales distribuidas en el espacio según un esquema de Poisson (definición: el número $N(v)$ de pepitas en un volumen v es una variable de Poisson de media λv ; si v y v' son disjuntos, $N(v)$ y $N(v')$ son independientes).

a/ Se toman volúmenes v , y se pone $Y(x) = N(v_x)$ en que v_x representa el trasladado de v implantado en el punto x . Probar que la covarianza de $Y(x)$ e $Y(x+h)$ es la varianza de $N(v_x \cap v_{x+h})$, es decir $\lambda K(h)$, siendo $K(h)$ el covariograma geométrico del volumen v .

b/ Se supone ahora que las pepitas tienen pesos aleatorios independientes (con media p_0 y varianza σ_p^2 , y se toma para $Y(x)$ la suma $P_1 + P_2 + \dots + P_N$ de los pesos de los $N = N(v_x)$ pepitas contenidas en v . Calcular la media y la varianza de $Y(x)$ (solución: $E(Y) = \lambda v p_0$, $D^2(Y) = \lambda v (p_0^2 + \sigma_p^2)$). Deducir la covarianza de $Y(x)$ e $Y(x+h)$ (reemplazar v por $K(h)$).

Ejercicio 8: Esquema de dilución. - Se tienen gérmenes poissonianos como en el ejercicio 9. Sea $f(x)$ una función. Se pone:

$$Y(x) = \sum_i f(x - x_i)$$

en que x_i representa las implantaciones de los diferentes gérmenes. Se trata entonces de una dilución de estos gérmenes por la función f . Encontrar la covarianza $K(h)$ de $Y(x)$ e $Y(x+h)$ ($\lambda g(h)$, con $g = f * \tilde{f}$).

(Hay que limitarse al caso en que f es a soporte compacto, considerar dos puntos x_0 y x_0+h y un domino acotado V que contiene los trasladados por x_0 y x_0+h del soporte de f . Se razonará suponiendo que el número n de puntos poissonianos caídos en V es fijo, luego se des-condicionará en n).

Ejercicio 9: Esquema esférico de una dimensión - Calcular las funciones auxiliares del esquema esférico:

$$\begin{aligned}\gamma(\ell) &= \frac{3\ell}{2a} - \frac{1}{2} \frac{\ell^3}{a^3} \quad \text{para } \ell \leq a \quad 1 \quad \text{para } \ell \geq a \\ \chi(\ell) &= \frac{3\ell}{4a} - \frac{1}{8} \frac{\ell^3}{a^3} \quad \text{para } \ell \leq a \quad 1 - \frac{3a}{8\ell} \quad \text{para } \ell \geq a \\ F(\ell) &= \frac{1}{2} \frac{\ell}{a} - \frac{1}{20} \frac{\ell^3}{a^3} \quad \text{para } \ell \leq a \quad 1 - \frac{3a}{4\ell} + \frac{1}{5} \frac{a^2}{\ell^2} \quad \text{para } \ell \geq a\end{aligned}$$

b/ Varianza de extensión elemental de una muestra en su segmento b de influencia, para $b \leq a$ y $b \geq 2a$. Interpretar.

$$\left(\frac{1}{4} \frac{b}{a} + \frac{3}{160} \frac{b^3}{a^3} \quad \text{y} \quad 1 - \frac{3a}{4b} - \frac{1}{5} \frac{a^2}{b^2} \right)$$

6-3 EJERCICIOS SOBRE EL KRIGEADO (7).

Ejercicio 10 (Kriging de un segmento de longitud l) a/ Se tiene en la recta una F.A.I. sin deriva cuyo variograma es $\gamma(h)$ y cuatro puntos de muestreo: $x_1, x_2=x_1+l, x_3=x_2+l, x_4=x_3+l$. Se desea krigear el segmento (x_2, x_3) de longitud l a partir de los valores Z_1, Z_2, Z_3, Z_4 de la realización en x_1, x_2, x_3, x_4 . Escribir el sistema (3-17) utilizando las funciones auxiliares χ y F del párrafo 2-5-2.

(Solución: probar que $\lambda_1 = \lambda_4 = \lambda/2, \lambda_2 = \lambda_3 = (1-\lambda)/2$, lo cual es evidente por simetría, con λ solución del sistema:

$$\begin{aligned}\frac{\lambda}{2} \gamma(l) + \frac{1-\lambda}{2} \gamma(l) + \frac{\lambda}{2} \gamma(2l) &= \chi(l) + \mu \\ \frac{1-\lambda}{2} \gamma(l) + \frac{1-\lambda}{2} \gamma(2l) + \frac{\lambda}{2} \gamma(3l) &= 2\chi(2l) - \chi(l) + \mu\end{aligned}$$

y eliminar el parámetro de Lagrange μ , de donde:

$$\lambda \left[\frac{1}{2} \gamma(3l) - \gamma(2l) - \frac{1}{2} \gamma(l) \right] = 2\chi(2l) - 2\chi(l) - \frac{1}{2} \gamma(2l)$$

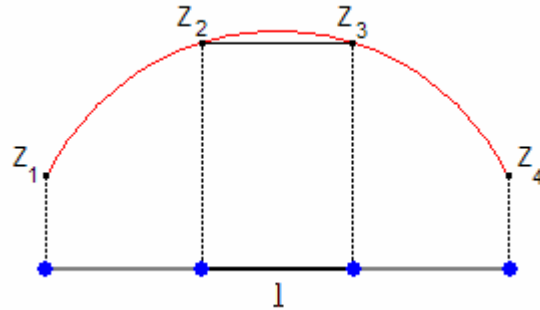
b/ Aplicar al caso $\gamma(h) = |h|^\alpha, 0 \leq \alpha < 2$, e interpretar.

(Solución: utilizar el ejercicio 5 del capítulo 2. Se encuentra:

$$\lambda = \frac{2^\alpha - \frac{4}{\alpha+1}(2^\alpha - 1)}{1 + 2^{\alpha+1} - 3^\alpha}$$

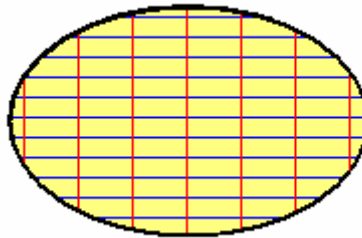
(7) Estos ejercicios fueron agregados por el traductor. Curiosamente, en este fascículo, no hay ejercicios sobre el kriging)

Para $\alpha=0$, $\lambda=1/2$ (efecto de pepita puro, el estimador óptimo es la media aritmética de las 4 muestras). Cuando α aumenta, λ decrece, se anula para $\alpha=1$ (propiedad markoviana del variograma lineal) y es negativo para $1<\alpha<2$: para $\alpha>1$, la realización presenta buenas propiedades de continuidad y el dibujo adjunto hace comprender que, para $(Z_1+Z_4)/2 < (Z_2+Z_3)/2$ (por ejemplo), el estimador debe ser mayor que $(Z_2+Z_3)/2$, luego $\lambda < 0$:



(comparar con el ejercicio 12 del capítulo 4)

Ejercicio 11 (Yacimiento reconocido por dos sistemas de líneas)- Sea un yacimiento reconocido por dos redes regulares de galerías paralelas a dos direcciones perpendiculares.



Sean Z_M y Z_T las leyes medias de las líneas, sea Z la ley media desconocida del yacimiento y sean $\sigma_T^2 = D^2(Z-Z_T)$ y $\sigma_M^2 = D^2(Z-Z_M)$ las varianzas de estimación del yacimiento por las solas líneas horizontales y las solas líneas verticales respectivamente. Se admite que (lo cual es verdadero con una buena aproximación) $E[(Z-Z_M)(Z-Z_T)] = 0$.

a/ Probar que el estimador óptimo (krigeado) de Z es $Z^* = \lambda Z_M + (1-\lambda) Z_T$ con:

$$\lambda = \frac{\sigma_T^2}{\sigma_T^2 + \sigma_M^2}$$

(ponderación por el inverso de las varianzas de estimación respectivas) y que la varianza correspondiente es:

$$\frac{\sigma_M^2 \sigma_T^2}{\sigma_T^2 + \sigma_M^2}$$

(Solución: los ponderadores son de la forma λ y $1-\lambda$ debido a la condición que los pesos suman 1. Partir luego de:

$$D^2(Z-Z^*) = D^2[\lambda(Z-Z_M) + (1-\lambda)(Z-Z_T)] = \lambda^2\sigma_M^2 + (1-\lambda)^2\sigma_T^2$$

y minimizar en λ .

BIBLIOGRAFIA.

- [1] CARLIER, A. Contribution aux méthodes d'estimation des gisements d'uranium. Tesis. Fontenay aux Roses, 1964
- [2] FORMERY, Ph. Cours de Géostatistique. Ecole des Mines de Nancy, 1965 (no publicado).
- [3] MATHERON, G. Traité de Géostatistique Appliquée. Tomo 1 (1962), Tomo 2 (1963). Ediciones Technip, Paris.
- [4] MATHERON, G. Les Variables Régionalisées et leur estimation (Tesis). Masson, Paris, 1965.
- [5] MATHERON, G. Osnovy Prikladnoi Geostatistiki. Ediciones MIR, Moscú, 1968.
- [6] MATHERON, G. Le Krigeage Universel. Les Cahiers du Centre de Morphologie Mathématique, Fascicule 1, Fontainebleau, 1968.
- [7] SERRA, J. Echantillonnage et estimation locale des phénomènes de transition minières. Tesis. Nancy, 1967.
- [8] FORMERY, Ph. y Matheron, G. Recherche d'optima dans la reconnaissance et la mise en exploitation des gisements miniers. Annales des Mines, Mayo y Junio 1963.
- [9] BLONDEL, F. Evolution de la notion de réserves dans l'Industrie Minérale. -Freiberg, 1959.
- [10] LESOURNE, J. Technique Economique de Gestion Industrielle. - Dunod, Paris.