



# INSTITUTO SUPERIOR MINERO METALÚRGICO

FUNDADO EL 1 DE NOVIEMBRE 1976

## ESTUDIOS DE DOCTORADO

EL INSTITUTO SUPERIOR MINERO METALÚRGICO DE MOA CONVOCA A ESTUDIOS DE DOCTORADO EN LOS PERFILES DE:

### GEOLOGÍA

Tectónica - Cartografía geológica, Geología regional, Estratigrafía, Geodinámica, Hidrogeología, Hidrogeoquímica, Cartografía, Ingeniería geológica y geoambiental, Mineralogía/Petrología, Geoquímica, Geomatemática/Informática aplicada, Economía geológica, Búsqueda y Prospección de yacimientos minerales, Geología del petróleo, Geología del Caribe, Geotecnia, Mecánica de suelos y rocas, Estabilidad de terrenos.

### MEDIO AMBIENTE

Rehabilitación de áreas minadas, Geología ambiental, Residuales, Calidad de aguas, Evaluación de impactos ambientales, Evaluación de peligros y riesgos geológicos y ambientales, Contaminación atmosférica, suelos, aguas, Cartografía ambiental, Gestión ambiental, Calidad del medio físico receptor, Estudios del medio ambiente y georrecursos, Ordenamiento ambiental, territorial y de áreas minadas, Impacto de la minería y la metalurgia sobre el medio ambiente, Sistema de información geográfica GIS, Ingeniería ambiental.

### ELECTROMEQUÍMICA

Transporte minero, Electrificación de minas, Supervisión y control de redes eléctricas, Soldadura y conformación de explosivos, Tecnologías y equipos de control de plantas industriales, Conservación energética en la industria minero-metalúrgica, Automatización aplicada a la minería, Diseño de elementos de máquina, Técnicas de protección y seguridad del trabajo en plantas, Máquinas y accionamiento eléctrico automatizados, Automatización de procesos metalúrgicos, Procesos de automatización y control de plantas industriales, Fuentes de energía no convencionales, Tecnología y equipos de plantas minero-metalúrgicas (bombas, ventiladores, compresores y otros).

### METALURGIA

Metalurgia extractiva, Hornos, Procesos metalúrgicos, Soldadura y conformación con explosivos, Investigación de procesos de preparación de mineral, Metalogenia de los macizos ofiolíticos, Evaluación geólogo-tecnológica de la materia prima mineral, Beneficio de los minerales, Técnicas de protección y seguridad del trabajo en plantas, Lixiviación ácida a presión en el proceso metalúrgico, Perfeccionamiento de tecnologías de producciones metalúrgicas, Procesos de automatización y control de plantas industriales, Perfeccionamiento de tecnologías metalúrgicas, Tecnología y equipos de plantas minero-metalúrgicas (bombas, ventiladores, compresores y otros).

### MINERÍA

Explotación a cielo abierto, Explotación de canteras, Voladuras, Geomecánica, Construcción subterránea, Seguridad en áreas minadas, Mineralogía, Transporte minero, Informática aplicada a la actividad minera, Economía de minas, Explotación racional de los georrecursos, Estabilidad de laderas y excavaciones, Topografía minera, Geotecnia, Mecánica de suelos y rocas, Estabilidad de terrenos.

PARA MÁS INFORMACIÓN DIRÍJASE A:  
Dr. Secundino Marrero, Director de Relaciones Institucionales  
ISMMM, Las Coloradas s/n, Moa, Holguín, Cuba.  
CP. 83329. Tel.: (53) (24) 6 6678, 6 4214  
Fax: (53) (24) 6 2290, 6 2326  
E-mail: marrero@ismm.edu.cu

## Contribución a la práctica del análisis variográfico y de la estimación por kriging

Aristides Alejandro Legrá Lobaina<sup>1</sup>  
René Luciano Guardiola Romero<sup>2</sup>

<sup>1</sup>Profesor Asistente, Departamento de Matemáticas, Instituto Superior Minero Metalúrgico, Las Coloradas s/n, Moa, Holguín.  
<sup>2</sup>Profesor Auxiliar, Departamento de Matemáticas Instituto Superior Minero Metalúrgico, Las Coloradas s/n, Moa, Holguín.

**RESUMEN:** En este trabajo se presentan los elementos principales relacionados con la práctica de la búsqueda del variograma experimental, del ajuste del variograma teórico, determinación de la propiedad de isotropía o anisotropía del fenómeno mediante la zona de influencia y la estimación mediante el kriging puntual. Se hace énfasis en los algoritmos fundamentales y se comentan las situaciones que, según el criterio de los autores, pueden crear dificultades al investigador. El enfoque presentado ha sido llevado a la práctica en los modelos geométricos y modelos geoquímicos del depósito mineral Punta Gorda de Moa, provincia Holguín, Cuba.

**Palabras clave:** variograma, kriging, zona de influencia, anisotropía, efecto pepita, estacionaridad

**ABSTRACT:** In this work are presented the principal elements related to practice of the search of the experimental variogram, of the adjustment of the theoretical variogram, determination of the isotropic or anisotropic property of the phenomenon through the influence zone and the estimate through the punctual kriging. It is made emphasis in the fundamental algorithms and are commented the situations that, according to the criterion of the authors, they could create difficulties to the investigators. The presented approach has been carried to practice in the geometric models and geochemistry models of the mineral deposit Punta Gorda of Moa, province Holguín, Cuba.

**Keywords:** variogram, kriging, influence zone, anisotropy, nugget effect, stationary.

### INTRODUCCIÓN

La geoestadística es actualmente un instrumento potente con que cuentan los geólogos y mineros para desarrollar con un alto margen de seguridad sus actividades. Además, independientemente de que existen detractores y críticos que han emitido criterios que hacen discutibles desde el punto de vista del rigor matemático algunos procedimientos geoestadísticos, Mercks (1992), no existen dudas de que, además de los profesionales de la geología y de la minería, los inversionistas han depositado su confianza en esta rama de las matemáticas aplicadas, Chica (1988), Cihl (1982), Ferguson (1988), etc.

Desde el punto de vista de sus orígenes, la geoestadística tiene dos contextos bien definidos: la práctica (desarrollada principalmente en Sudáfrica por D. G. Krige, Sichel, Wijs y otros a finales de la década del 50 y principios de los años 60 a partir de la explotación de las minas de oro y de uranio) y la teórica (cuyos artífices fueron el francés G. Matheron y sus colaboradores J. Serra, A. Journé, Ch. Huijbregts, etc basándose en la teoría de los procesos estocásticos), Alfonso (1989), Chica (1988) y García (1988).

Durante más de 40 años se han desarrollado nuevas teorías e innumerables aplicaciones que ya incluyen no solo las ciencias geólogo-mineras sino que también han modelado y resuelto problemas tan aparentemente diferentes como los que aparecen en la meteorología y en la economía, García Guerra (1988), no obstante, sin tratar de simplificar demasiado el asunto, podemos decir que toda la práctica de la geoestadística comienza con la modelación de la variabilidad de uno o varios parámetros de un fenómeno y finaliza con la obtención, a partir de estimaciones mediante kriging, de resultados que pueden ser materiales, por ejemplo, una modelación o un mapa, o conceptuales, por ejemplo, una conclusión.

Los éxitos que se pueden obtener con el uso de las técnicas que brinda la geoestadística se basan en tres aspectos principales:

1. Determinación de las herramientas más adecuadas para modelar y resolver un problema.
2. Análisis variográfico correcto (estudio de la variabilidad).
3. Interpretación adecuada de los resultados que permita volver a los pasos 1 y 2 para precisar las herramientas y el análisis.

Los aspectos 1 y 3 no pueden ser automatizados y el segundo aspecto solo será automatizado en las cuestiones relacionadas con cálculos y gráficos que permitan interactuar con los datos y modelos de manera rápida y fiable.

En este trabajo se presentan algunos elementos relacionados con la realización práctica del análisis variográfico y de la estimación por kriging puntual así como algunas ideas vinculadas a la automatización de los procedimientos que se vinculan a estos procesos. El enfoque presentado ha sido llevado a la práctica en la modelación geométrica y geoquímica del yacimiento laterítico Punta Gorda de Moa, provincia Holguín.

## DESARROLLO

## Algunos aspectos básicos de la teoría de los procesos estocásticos

Sea  $U$  una población de infinitos valores  $X \in A \subset R$ , que miden el comportamiento de un fenómeno  $F$  según varía el parámetro  $t \in B \subset R^n$  ( $t$  puede referirse al espacio y al tiempo). A la variable  $X$  la consideraremos una variable aleatoria.

Si medimos a  $X$  para los valores  $t \in M = [t_1, t_2] \subset B$  entonces se tiene una realización que denotamos  $X(t)$  y es una muestra del conjunto de todas las funciones posibles que representan a  $F$  en el intervalo  $M$ .

Si obtenemos  $k$  realizaciones  $S = \{X_1(t), \dots, X_k(t)\}$  en  $M$ , las mismas, en general, serán distintas y al conjunto  $S$  se le denomina *serie*. Gráficamente una serie de 4 funciones tiene el siguiente aspecto:

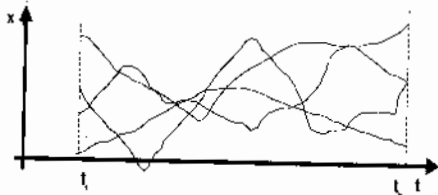


FIGURA 1. Serie de 4 funciones.

Si se fija el valor de  $t$  tomando  $t = t^*$ , entonces para cada función posible  $X(t)$  se obtiene un valor  $X = X(t^*)$ ; puesto que es infinito el conjunto  $Q$  de estos valores entonces a él se le puede asociar una función de distribución  $P[X(t^*)]$ . Si los valores de  $X$  se obtienen de una serie entonces el conjunto de valores  $\{X_1, \dots, X_k\}$  es una muestra tal como la entendamos en la estadística clásica.

Gráficamente se puede tener la siguiente interpretación:

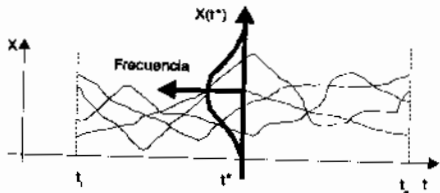


FIGURA 2. Una muestra en una serie de 4 funciones.

Entonces, se define como proceso estocástico o aleatorio (PE) al conjunto ordenado de todas las funciones  $\{X(t)\}$  para  $t \in B$ .

Un PE se denomina *estacionario con respecto a la media* si se cumple que  $E\{X(t)\} = C_1$  para cualquier valor  $t \in B$  y se dice *estacionario con respecto a la varianza* si  $Var\{X(t)\} = C_2$  para cualquier valor de  $t \in B$ .

Cuando un PE no es estacionario respecto a la media, entonces para cada valor de  $t$  se obtienen, en general, diferentes valores de  $E\{X(t)\}$ , ellos son la imagen de una función  $T(t)$  que se denomina *tendencia* del PE.

Un PE se denomina *ergódico con respecto a la media* cuando el valor medio en  $t$  de cualquier realización  $X(t)$  es igual al valor medio de los valores de  $E\{X(t)\}$  donde  $t \in M = [t_1, t_2] \subset B$ . Análogamente se puede definir un PE ergódico con respecto a la varianza.

Si un PE es ergódico con respecto a una variable estadística entonces basta una realización para evaluar dicha característica para todo el proceso.

Cuando un PE es ergódico pero no es estacionario con respecto a la media entonces para una realización  $X(t)$  y para la tendencia  $T(t)$  se define la *realización residual*  $R(t) = X(t) - T(t)$  y se cumple que  $E\{R(t)\} = 0$ .

Un problema de primera importancia es evaluar el grado de aleatoriedad de una serie. Si aceptamos que el PE es estacionario y ergódico con media 0 y se tiene una realización  $X(t)$  entonces se define la función de autocovarianza:

$$R_{xx}(\tau) = \lim_{T \rightarrow \infty} \left\{ \frac{1}{2T} \int_{-T}^T X(t)X(t+\tau) dt \right\} \quad \text{donde } \tau \text{ es un desplazamiento arbitrario de } t.$$

Se cumple que  $R_{xx}(0) = \sigma_x^2$  y se define la función de autocorrelación  $C_{xx}(\tau) = R_{xx}(\tau) / \sigma_x^2$ .

Tiene especial interés para el desarrollo que se verá más adelante mencionar un tipo de serie estacionaria no autocorrelacionada llamada *Ruido Blanco* que cumple que  $R_{xx}(\tau) = \sigma_x^2$  para  $\tau = 0$  y  $R_{xx}(\tau) = 0$  para  $\tau \neq 0$ .

Una consecuencia de adicionar un ruido blanco  $S(t)$  a otra serie  $X(t)$  es la aparición de una discontinuidad de la función de autocorrelación en el origen. En este caso se escribe:

$$C_{xx}(\tau) = \frac{R_{xx}(\tau)}{R_{xx}(0)} = \begin{cases} \frac{R_{SS}(\tau)}{R_{SS}(0) + R_{xx}(0)} & \text{para } \tau = 0 \\ \frac{R_{xx}(\tau)}{R_{xx}(0)} & \text{para } |\tau| > 0 \end{cases}$$

La separación RC que se debe producir a partir de un punto arbitrario  $t_0$  para que los valores de la realización observada  $X(t \pm RC)$  sean estadísticamente independientes de  $X(t_0)$  se denomina *radio de correlación* o *de influencia* de la realización. El radio de correlación puede obtenerse teóricamente a partir de la función de autocovarianza pero en la práctica esto es una tarea de grandes dificultades puesto que la función de autocovarianza debe determinarse a partir de datos discretos; sin embargo se han estudiado varios casos que aparecen con frecuencia en la práctica, entre ellos vale la pena destacar el caso en que la función está dada por la adición de un ruido blanco (es un proceso no autocorrelacionado o sea el radio de correlación es teóricamente nulo) mas otro proceso de otro tipo, que genera una discontinuidad de la función de autocorrelación en el origen. El radio de correlación de un ruido blanco

es teóricamente 0 y en la práctica es menor que el intervalo utilizado en el muestreo.

## Comentarios sobre los conceptos básicos de la geoestadística lineal.

La geoestadística, surgió como aplicación de la teoría de los PE al estudio de fenómenos y procesos geológicos y mineros.

La variable aleatoria  $X$  en este caso depende de coordenadas espaciales o sea  $t = (x, y, z)$  por lo que se le llama *variable aleatoria regionalizada*;  $B \subset R^3$ .

Otros conceptos básicos son los de campo y soporte geométricos y principalmente el *variograma* (o función estructural) y la *zona de influencia*. El variograma (a veces llamado *semivariograma*) se define, generalmente, como  $\gamma(h) = \gamma(x_i, x_j) = \frac{\text{var}(X(t_i) - X(t_j))}{2}$  donde  $h$  es

la distancia entre  $t_i$  y  $t_j$  y contiene junto con la zona de influencia (en este caso es una región cerrada del plano o del espacio) la información necesaria sobre la parte estructural del fenómeno estudiado.

Los fenómenos geológicos no siempre cumplen con los requerimientos generales que se necesitan para aplicar la teoría de los PE, vale destacar que estos fenómenos no siempre son estacionarios con respecto a la media y a la varianza, por otra parte la información que se dispone sobre el fenómeno es, la mayoría de las veces, solo una de las infinitas realizaciones teóricamente posibles por lo que es necesario admitir la ergodicidad del fenómeno.

Esto último generalmente se admite de forma implícita; en opinión de los autores, la validez de esta decisión no puede buscarse solo en la realidad del fenómeno estudiado sino también en la disponibilidad de información que se posee.

Las hipótesis de la geoestadística más conocidas que se refieren a la estacionariedad son las de estacionariedad estricta, estacionariedad de segundo orden, condición intrínseca y la de los procesos cuasiestacionarios. A continuación se tiene una explicación breve de cada una de ellas, Alfonso (1989), Chica (1988):

1. **Estacionariedad estricta:** Para todo  $t \in B$  se tiene que  $E\{X(t)\} = C_1$  y  $Var\{X(t)\} = C_2$  o sea las funciones de distribución de probabilidades son iguales entre sí, independientemente del valor de  $t$  escogido.
2. **Estacionariedad de segundo orden:** En este caso  $E\{X(t)\} = C_1$ , y además existe la función de covarianza  $K(h) = E\{X(t)X(t+h)\} - C_1^2$  la cual es independiente de  $t$ .
3. **Condición intrínseca:** Se cumple que  $E\{X(t)\} = C_1$ , y existe  $Var\{X(t) - X(t+h)\} = 2\gamma(h)$  que solo depende de los valores de  $h$ . La estacionariedad de segundo orden implica la condición intrínseca (también llamada de homo-

geneidad) pero lo contrario no se cumple. Se puede demostrar que  $\gamma(h) = \sigma^2 - K(h)$  donde  $\sigma^2$  es la varianza de los datos.

4. **Cuasiestacionariedad:** En este caso se cumple la estacionariedad de segundo orden o la condición intrínseca para  $|h| < b$ , donde  $b$  representa las dimensiones de una región donde el fenómeno conserva cierta homogeneidad estadística.

La zona de influencia tiene una enorme importancia práctica ya que define las distancias y direcciones donde se mantiene la influencia de un dato respecto a otro. En  $R^2$  se acostumbra a representarla mediante una elipse y en  $R^3$  mediante un elipsoide aunque no necesariamente tienen que ser figuras de estas características. Desde el punto de vista geológico la zona de influencia tiene una interpretación precisa para cada caso que se analiza.

## Cuatro condiciones que deben cumplir los datos que expresan las mediciones del fenómeno

1. Puesto que los datos representan al fenómeno que se estudia, hay que garantizar que existan la cantidad suficiente como para que se obtengan los resultados deseados. Desde el punto de vista económico las mediciones son generalmente costosas y desde el punto de vista de la precisión de los resultados buscados, generalmente esta aumenta junto con el número de mediciones. Esta contradicción es bastante difícil de resolver y es, dentro de la geoestadística, un tema particular llamado *optimización de redes de exploración* y que exige el estudio particular de cada caso. El criterio general que se recomienda en este trabajo para determinar la red más adecuada es el siguiente:
  - a. Obtener, para una red poco densa de toda la región o para una red densa de una zona especialmente representativa, el variograma experimental y ajustar el variograma teórico (el como hacerlo será explicado más adelante).
  - b. Obtener mediante estimación por Kriging puntual redes cada vez más densas o extendidas a toda la región y calcular la media aritmética y la desviación estándar de los errores (estos errores solo dependen del variograma y de los datos conocidos). Comparando estos valores (por ejemplo, mediante porcentajes) con la media y la desviación estándar de los datos usados podemos valorar la calidad de nuestro modelo para cada nueva densidad de red. Nótese que mantenemos el modelo y se usan para la próxima red las mediciones que se van estimando.
  - c. Calcular los costos que se tendrían al realizar mediciones en cada una de las redes más densas y mediante un gráfico, por ejemplo de % error medio contra costo, o auxiliándose de técnicas analíticas tales

como la modelación de la relación entre las dos variables, tomar una decisión.

Otro factor que debe tenerse en cuenta para una posible densificación de la red es la presencia del efecto pepita, cuestión que analizaremos más adelante.

Por otra parte, es posible que la mejor solución sea determinar en cuales subregiones se presentan los mayores errores y solicitar la densificación de la red en estas zonas.

- La representatividad de los datos es tan importante como el aspecto anterior y no solo debe estar dada por el hecho de que incluya datos dentro de los rangos en que se manifiesta el fenómeno (esto es deseable) sino que en opinión de los autores deben ser mediciones tomadas en puntos geográficamente independientes de las características de la variable que se mide ya que representan una realización cualquiera de un proceso estocástico. Esto, generalmente, se considera resuelto definiendo una red espacial de muestreo lo más regular posible lo cual, además, garantiza que el error global de estimación (llamado también *error de cualidad*) sea mínimo, García Guerra (1988).
- Uno de los problemas que se pueden presentar es que las mediciones se hayan desarrollado sobre una red de dimensiones no uniformes o que se tengan dos o más redes diferentes de mediciones; por lo que deben distinguirse dos casos:
  - Se tienen dos o más redes regulares de diferentes dimensiones. Aquí se procede a determinar los variogramas en cada una de las redes y se realiza la *regularización* de cada uno de ellos, que es un proceso que tiene como objetivo convertir cada variograma correspondiente a una parte de los datos regulares pertenecientes a una red (soporte), en un variograma de soporte común para todas las mediciones. La regularización que se define como una integral estocástica conserva la media pero produce cambios importantes en el variograma, Chica Olmo (1988), Deraisme (1996), García Guerra (1988); el variograma de la variable regularizada se puede describir como una función del variograma de la variable original. El procedimiento de pasar del variograma regularizado al variograma original se denomina *desconvolución*. Detalles sobre la práctica de la regularización pueden verse en García Guerra (1988).

No obstante se debe aclarar que el término regularización puede emplearse en el sentido de obtener una red regular estimada por diferentes métodos (incluyendo los de interpolación); en este caso hablamos de *regularizar la red*.

- Que la red utilizada sea muy irregular o que exista redundancia en los datos. Cuando esto sucede se divide la región en subregiones regulares y se le

asigna al punto central de cada una de ellas un valor de la variable  $W$  que es una función de los valores  $W_i$  que pertenecen a la subregión (puede usarse la media aritmética, inverso de una potencia de la distancia, etc.); a este proceso se le llama *declustering*. En los casos de que una subregión no contenga ningún valor  $W_i$  deben cambiarse las dimensiones de las subregiones.

- Para ciertos procesos de simulación y de geostatística no lineal se hace necesario que los datos se distribuyan normalmente; si esto se cumple o no, debe comprobarse mediante los métodos conocidos de ajuste de distribuciones y en caso de que no suceda, pueden tomarse una de los siguientes caminos:
  - Buscar una función de transformación de los datos (en ocasiones se le llama *anatomosis gaussiana*, García Guerra (1988). Las dos formas más conocidas para realizar estas transformaciones son la *gráfica mediante el método de Monte Carlo* y la *analítica mediante los polinomios ortogonales de Hermite*, García Guerra (1988).
  - Generar, a partir de los datos dados, un conjunto de otros nuevos que conserven sus principales parámetros estadísticos (media, varianza, etc), que se distribuyan normalmente y que estén dentro de los rangos de los valores originales; a partir de estos nuevos valores se realizan los análisis. El método más conocido para obtener los valores simulados es el de *bandas rotantes*, García Guerra (1988).

#### Determinación del variograma experimental

El variograma real (VR) del fenómeno como expresión de su variabilidad es desconocido y solo sería posible determinarlo si se estudiara el fenómeno dentro de la continuidad espacial del dominio donde se desarrolla. En la práctica, se determina un variograma experimental (VE) que refleja de manera discreta al variograma real pero para los fines de los cálculos ajustamos un variograma teórico (VT) al variograma experimental. A continuación pasamos a explicar como se determina el variograma experimental.

Sea  $A$  el conjunto de  $n$  datos dados como pares  $(P_i, W_i)$ ,  $i=1, \dots, n$  donde  $P_i$  es un punto del plano o del espacio y  $W_i$  es el valor en  $P_i$  de la variable que se analiza. El variograma debe expresarse de alguna forma el comportamiento de la variabilidad de  $W$  en cierto subconjunto del plano o del espacio el cual recibe el nombre de *campo geométrico* (CG); el subconjunto del CG en la cual se realizan las mediciones se denomina *soporte*; este concepto es fundamental pues de sus dimensiones depende la modelación que se obtenga (aunque todas las modelaciones obtenidas a partir de diferentes soportes están relacionadas por sus parámetros Chica (1996), García Guerra (1988), es imprescindible conocer que se

gún aumentan las dimensiones del soporte disminuye la varianza experimental.

Esta variabilidad puede ser estudiada por sus valores medios en función de las distancias entre los puntos y se expresa mediante el variograma, por tanto el variograma es una función de  $h$ . El valor de la distancia  $h$  puede ser determinado de varias maneras pero es la distancia euclidiana la que expresa (en las escalas en que se desenvuelven la geología y la minería) de manera más exacta las distancias entre los objetos del mundo real; es evidente que  $\gamma(0)=0$ .

Generalmente los valores de  $\gamma$  crecen o se mantienen aproximadamente constantes en la medida que aumentan los valores de  $h$  pues el variograma refleja un fenómeno bastante común relacionado con el hecho de que la variabilidad entre los valores de  $W$  no disminuye a medida en que los puntos de medición se alejan entre sí. Un variograma con un crecimiento lento indica una gran continuidad del comportamiento de  $W$  y por el contrario un crecimiento rápido del variograma indica discontinuidad en el comportamiento de la variable.

Por el contrario, la influencia de un valor de  $W$  con respecto a otro disminuye generalmente a medida en que los puntos de medición se alejan.

La primera cuestión a analizar para obtener el VE es precisar cual es la distancia básica o paso  $h_0$  (en los textos en inglés aparece como *lag Annelis* (1989), Deraisme (1996), Myers (1991), Pannatier (1996), a partir de la cual se definen las diferentes distancias  $2h_0, 3h_0, \dots, kh_0$ . Puesto que en la práctica las distancias que se obtienen entre los pares de puntos de  $A$  no coinciden exactamente con los valores de  $h$ , entonces se definen  $k+1$  intervalos disjuntos dos a dos  $[0, h_0], (h_0, h_0], \dots, (h_{k-1}, h_k]$  y las distancias entre los pares de puntos de  $A$  se incluyen en uno de estos intervalos.

Llamemos  $D$  a la máxima distancia entre todas las distancias entre parejas de puntos de  $A$ , entonces  $h_0 \leq D$ ; si  $h_0$  es muy grande entonces  $k$  será pequeño y se pierde mucha información sobre el fenómeno (el variograma se presenta gráficamente muy suavizado) y si por el contrario  $h_0$  es demasiado pequeño entonces no se obtiene una buena apreciación de las características más importantes de la variabilidad del fenómeno (el gráfico del variograma aparece como una sucesión de valores alto y bajos). En la práctica los autores recomiendan que para comenzar se definan todos los rangos de distancia posibles que sean iguales e incluyan al menos una pareja de puntos y se busquen de 6 a 10 rangos que con mayor frecuencia aparecen y de ellas se tome la menor como  $h_0$ . No obstante, ha sucedido que se ha tenido necesidad de rectificar el valor de  $h_0$  después de encontrarlo de esta manera debido a que las redes han sido especialmente complejas por sus irregularidades. Si las redes son regulares entonces el valor de  $h_0$  es bastante fácil de determinar pero ha de tenerse extremo cuidado si existe mucha diferencia entre

dos de las medidas del rectángulo (en el caso del plano) o del paralelepípedo (en el espacio) pues esto puede conducir a que se asigne a  $h_0$  un valor no adecuado. En García Guerra (1988), se expresa que debe tomarse la moda de los intervalos o sea aquella que tiene mayor cantidad de pares de puntos.

La segunda cuestión a tener en cuenta es que se necesita conocer cuantos pares de puntos están incluidos en cada intervalo de  $h$ . Puede darse el caso de que varios intervalos tengan cada uno de ellos varios miles de pares de puntos y en otro apenas aparezcan unas decenas. Esto debe evitarse ya que este último intervalo no tiene el mismo peso que los demás y puede tener un valor de  $g$  que no se corresponde con el comportamiento general de la variabilidad. En estos casos los autores recomiendan unir dos intervalos en uno de mayor longitud o redefinir  $h_0$ .

En tercer lugar, para cada valor de  $h$  al cual le corresponden  $n(h)$  pares de puntos se calcula el valor  $g(h)$ . Este cálculo se realiza generalmente por la fórmula

$$\gamma(h) = \frac{1}{2n(h)} \sum_{i=1}^{n(h)} (W(P_i) - W(P_j))^2$$

donde  $P_i$  y  $P_j$  son dos puntos de  $A$  que están a una distancia  $h$  uno del otro.

Se conocen otras fórmulas para este cálculo, Deraisme (1996) tales como las de Hawkins de 1980, Armstrong-Delfiner de 1980 y de Omre de 1984 las cuales fueron creadas para mejorar la solución de ciertos problemas específicos.

En cuarto lugar se precisa editar (en el sentido de seleccionarlos) los intervalos que se utilizarán en el ajuste y por tanto en las estimaciones. Algunos autores, como García Guerra (1988) y Gutiérrez Montano (1993), plantean que basta tomar el 25%, el 33% o el 50% de los primeros intervalos puesto que si la distancia entre dos puntos  $P_1$  y  $P_2$  es muy grande el valor de  $W_1$  no influye en el valor de  $W_2$ . Este razonamiento es en la mayoría de los casos correcto pero los autores de este trabajo opinan que siempre deben tomarse más del 50% de los intervalos ya que el variograma debe aportar información que sea posible de comparar con toda la información geológica o minera que se disponga. Además, si el variograma presenta a partir de cierto punto  $h_1$  un valor aproximadamente constante  $M_0$  entonces se verá que basta tomar hasta el próximo intervalo después del que incluye a  $h_1$ .

El quinto paso es el de calcular los parámetros principales de los datos y del variograma. Para los datos se calcularán, para  $W$  la Media Aritmética  $M_0$  y la varianza  $\sigma^2$  y para los puntos  $P_i$  se determinarán la distancia media  $d_n$  y la distancia máxima  $D$ .

Para el variograma se determinarán:

- Efecto pepita:** Este parámetro que se cuantifica como un valor no negativo al que designaremos como  $C_0$  tiene especial interés; en caso de que  $C_0=0$  entonces se dice que no hay efecto pepita. Desde el punto de vista teórico surge como un ruido blanco [1] y gráfica-

mente se manifiesta como una discontinuidad del variograma en el origen pues  $\gamma(0)=0$  y  $\gamma(0^+)=C_0$ . El origen del efecto pepita puede tener en la práctica dos causas fundamentales, Alfonso (1989), Chica (1988):

- a. Errores en las mediciones.
- b. A que la escala de mediciones sea inferior a las dimensiones de las zonas donde W manifiesta variaciones sustanciales.

Según Chica Olmo (1993), en la práctica es bastante difícil discernir entre los dos motivos y solo un buen conocimiento del fenómeno podrá ayudar a hacerlo.

El valor de  $C_0$  puede ser determinado prácticamente a partir del análisis de dos primeros puntos del variograma experimental buscando la intersección de la recta que ellos determinan con el eje de las ordenadas y García Guerra (1988). Este no es un criterio absoluto y en algunas ocasiones los autores de este trabajo han tomado con estos fines la recta mínimo cuadrada a partir de los primeros tres o cuatro puntos. Sin embargo estos valores de  $C_0$  siempre deberán ser cuidadosamente revisados.

Es evidente que si el valor de  $C_0$  obtenido de esta forma es negativo entonces puede tomarse como 0 o puede revisarse todo el proceso.

2. **Alcance y meseta**: Estos parámetros solo existen si el fenómeno es estacionario y el variograma presenta a partir de cierto punto  $h_0 < D$  un valor aproximadamente constante  $M_0$ . Precisamente  $h_0$  es denominado alcance y  $M_0$  meseta. Teóricamente  $M_0$  coincide con la varianza experimental  $\sigma^2$  de los datos, Alfonso (1989), Chica (1988), Derisme (1996), Pannatier (1996), Rodríguez Morán (1988). Los variogramas que presentan valores de meseta y de alcance en ocasiones se les denominan de transición y aparecen con frecuencia ya que el alcance indica a partir de que distancia promedio desaparece la relación estadística entre los valores de W en dos puntos. En el caso en que el valor de  $C_0$  coincide con el valor de  $M_0$  y  $h_0=0$  entonces el fenómeno se denomina totalmente aleatorio. En la práctica, los valores del alcance y la meseta los decide el investigador a partir de los puntos del VE, su experiencia y del conocimiento del fenómeno, sin embargo se propondrá un método que puede facilitar el análisis:

a. Se definen tres coeficientes porcentuales, a saber:

- RE o sea rango de estabilidad que define el intervalo de valores de la variable g para los cuales se puede considerar que la misma es aproximadamente constante. En general este intervalo se define mediante  $[\sigma^2 - \epsilon, \sigma^2 + \epsilon]$ , donde  $\epsilon = RE \sigma^2 / 100$ . En la práctica puede comenzarse a trabajar con  $RE=20\%$ .

- RB o sea rango de búsqueda de estabilidad. Siendo k el número de intervalos o clase consideradas para h se define el índice del intervalo hasta donde se busca la estabilidad de y como IB = Redondear  $(RB \cdot k / 100)$  donde "Redondear" es la fun-

ción que redondea un número. En la práctica puede comenzarse a trabajar con  $RB=70\%$ .

- RA o sea rango de aleatoriedad pura. Este parámetro tiene en cuenta la posibilidad de que el crecimiento del variograma sea tan rápido como su estabilización por lo cual sea conveniente considerar que el fenómeno es totalmente aleatorio. En la práctica puede comenzarse a trabajar con  $RA=15\%$ .

b. Conociéndose el intervalo  $[\sigma^2 - \epsilon, \sigma^2 + \epsilon]$  ( $\sigma^2$  debe ser un valor cercano o igual a  $M_0$ , si este último existe), se comprueba desde el primer intervalo hasta el intervalo de índice IB si los valores de  $\gamma$  del VE están, a partir de cierto índice  $i$ , dentro de dicho intervalo y siendo así entonces un valor de h perteneciente al intervalo de orden  $i$ -ésimo definirá el valor de  $h_0$  y un valor conveniente de  $\gamma \in [\sigma^2 - \epsilon, \sigma^2 + \epsilon]$  definirá el valor de  $M_0$ . Estos valores de  $h_0$  y  $M_0$  deberán ser precisados por los investigadores.

c. Si se obtiene un índice  $i_0$  donde comience la estabilidad de g entonces deberá verificarse que  $h_{i_0} < RA \cdot h_{i_0}$  para decidir si se trata o no de un caso de aleatoriedad pura.

d. Si no se obtiene un índice  $i_0$  donde comience la estabilidad de g entonces se tratará de encontrar en el VE el primer valor de h tal que  $\gamma \geq \sigma^2$ . Si este valor existe entonces se debe considerar que se producen oscilaciones o no y si el variograma es, en general, creciente o tiene un comportamiento de crecimiento-decrecimiento o viceversa; en estos casos se puede sospechar de un modelo compuesto de varias estructuras (de los cuales se hablará más adelante). Si el valor no existe entonces se puede sospechar un fenómeno de gran continuidad y en estos casos se puede tomar para los fines prácticos a  $h_0 = h_{i_0} / 2$  y  $M_0$  como el valor de g correspondiente o analizar la posible existencia de una tendencia (no estacionariedad).

El alcance y la meseta junto con  $C_0$  intervienen en muchos modelos teóricos con los que se ajusta el VE. Denotando a  $C_1 = M_0 - C_0$ , entre ellos se puede mencionar:

- **Modelo esférico de matheron**:  $\gamma(0)=0$ ;  $\gamma(h) = C_0 + C_1(3^*h/(2^*h_0) - [(h/h_0)^3]/2)$  para  $0 < h \leq h_0$ ;  $\gamma(h) = C_0 + C_1$  para  $h > h_0$ .
- **Modelo exponencial de Formery**:  $\gamma(0)=0$ ;  $\gamma(h) = C_0 + [C_1/(1-e^{-h/h_0})](1-e^{-h/h_0})$  para  $0 < h \leq h_0$ ;  $\gamma(h) = C_0 + C_1$  para  $h > h_0$ .
- **Modelo parabólico de Gauss**:  $\gamma(0)=0$ ;  $\gamma(h) = C_0 + [C_1/(1-e^{-h/h_0})](1-e^{-h/h_0})$  para  $0 < h \leq h_0$ ;  $\gamma(h) = C_0 + C_1$  para  $h > h_0$ .
- **Efecto seno**:  $\gamma(0)=0$ ;  $\gamma(h) = C_0 + [C_1/(h/h_0)](1-\sin(\pi h/h_0))$ .

Debe destacarse que en los casos II y III se han hecho pequeñas variaciones a las ecuaciones originales para garantizar la continuidad de los modelos en  $h=h_0$ .

3. **Pendiente**: Solo es necesario buscarla si se supone que el variograma se puede explicar como una función lineal de la distancia o sea  $\gamma(h)=C_0 + P_1 \cdot h$ . En este caso  $P_1$  es la pendiente de la recta. En estos casos es más importante que la recta se correspondiera mejor a los primeros valores de h. Un caso que puede considerarse semejante al modelo de la recta es el modelo logarítmico de Wijs cuya expresión es  $\gamma(h) = C_0 + r \ln(h)$  para  $h > 0$  y  $\gamma(0)=0$ . En la práctica de los autores de este trabajo se ha adaptado este modelo al caso de que se presente un VE con alcance y meseta mediante la expresión  $\gamma(h) = C_0 + [C_1/\ln(1+h_0)] \cdot (\ln(1+h))$  si  $0 < h \leq h_0$  y  $\gamma(h) = C_0 + C_1$  para  $h > h_0$ .

4. **Potencia**: En ocasiones el variograma experimental puede explicarse como una función  $\gamma(h)=w^2 \cdot h^v$ . El parámetro v es la potencia y se demostró que  $v \in (0,2)$ . Los parámetros v y  $w^2$  se determinan a partir de que  $\ln(\gamma) = \ln(w^2) + v \ln(h)$ . Este modelo se ha analizado también para el caso en que  $C_0=0$  y en este caso toma la forma  $\gamma(h) = C_0 + w^2 \cdot h^v$ .

5. **Periodo y atenuación**: Uno de los modelos utilizados con cierta frecuencia es el llamado Efecto Coseno que se define como  $\gamma(0)=0$  y  $\gamma(h) = C_0 + C_1(1-\cos(2\pi h/T))$  si no presenta atenuación y en caso contrario  $\gamma(h) = C_0 + C_1(1-\cos(2\pi h/T))e^{-(h/h_0)^n}$  donde T es el periodo y atn es un factor de atenuación. Los valores de T y de atn son difíciles de obtener y el modo de hacerlo que se han empleado los autores de este trabajo es mediante la interacción con gráficos y criterios analíticos de bondad de ajuste usando computadoras.

Aunque los modelos de la recta, de Wijs, potencial y efecto coseno no presentan explícitamente alcance y mesetas pueden ser definidos por tramos y por tanto a partir de cierto valor de h precisar que el valor de g es constante.

En la literatura consultada se proscriben el uso del método de ajuste por el principio de los mínimos cuadrados o no se menciona (lo mismo sucede con las interpolaciones) como un posible modelo. Las dos causas principales que se aducen son:

- 1. No toda función f(h) es un variograma, ya que para que esto suceda debe cumplirse que -f(h) sea definida

positiva o sea que si  $w(P) = \sum \lambda_i W(P_i)$  se cumpla que para todo conjunto de puntos  $P_1, \dots, P_n$  y de números reales  $\lambda_1, \dots, \lambda_n$  la condición  $\text{Var}(W(P)) = \sum \sum \lambda_i \lambda_j (P_i, P_j) \geq 0$

junto con la condición  $\sum \lambda_i = 0$ . Esto es sumamente complejo de demostrar para cada caso.

Se puede demostrar una propiedad importante para los variogramas:

$$\lim_{|h| \rightarrow \infty} \frac{\gamma(h)}{|h|^2}$$

Cuando esta propiedad no se cumple en el VE (efecto de parábola) debe pensarse en la existencia de una tendencia (drift).

- 2. Al ajustarse una función f(h) a un VE, deben considerarse dos cuestiones:

- a. El ajuste a los primeros intervalos es más importante que el ajuste global.
- b. Los métodos analíticos no tienen en cuenta el número de pares necesarios para realizar un buen ajuste (este número según García Guerra (1988), no debe ser menor que 30).

Siempre deben tenerse en cuenta todos los factores anteriores, no solo como elementos teóricos sino que, además, deben formar parte de todo análisis variográfico real.

Finalmente se quiere señalar que una de las formas más comunes de realizar un buen ajuste a partir de modelos conocidos es creando un modelo compuesto por varias estructuras en diferentes intervalos de h (un variograma definido por tramos) o una combinación de variogramas para todo el intervalo. Para esta última posibilidad hay tres casos interesantes:

1. **Estructuras imbricadas**: Cuando el comportamiento de la variabilidad del fenómeno real depende de los cambios de escalas de la distancia. El variograma se define como una combinación de variogramas elementales donde cada uno de ellos se ha obtenido a partir de una escala diferente de h.

2. **Estructuras de periodicidad**: El variograma presenta variaciones periódicas que definen relaciones crecientes y decrecientes de  $\gamma$  con respecto a h. En estos casos puede usarse un modelo como el de Efecto Coseno o combinarse varios variogramas de este y otros tipos.

3. **Efecto de pozo o de hueco**: Se produce a partir de cierto valor de h un decrecimiento de  $\gamma$  y luego se estabiliza su comportamiento (aunque puede presentarse más de una oscilación).

### Ajuste del variograma teórico

Tres elementos contribuyen notablemente a realizar un ajuste adecuado de un VT a un VE.

El primer elemento está dado por el conocimiento que se tenga de los diferentes modelos teóricos de variogramas, tanto de sus parámetros y ecuaciones como de sus gráficos. Esto se complementa con un software que permita ir ajustando dinámicamente y visualmente el modelo teórico al VE.

El segundo elemento es el uso del llamado IGF (Indicative Goodness to Fit) o sea indicador de bondad de ajuste, Pannatier (1994), que está dado por:

$$IGF = \frac{1}{T} \sum_{k=1}^T \sum_{i=1}^{U(i)} \frac{h_k}{\sum_{j=1}^{U(i)} \text{Pares}(j)} \frac{\text{Pares}(i)}{h_i} [v_i - \gamma(h_i)]^2$$

Donde T es el número de estructuras que forman el modelo, L(T) es el número de intervalos que intervienen en la estructura T, Pares(i) es el número de pares que intervienen en el intervalo i, h<sub>i</sub> es la distancia máxima de h, h<sub>i</sub> es la media de la distancia para el intervalo de índice i, γ<sub>i</sub> es el valor del VE en el intervalo de índice i y γ(h) es el valor del VT en h. En este caso mientras mas cercano a 0 sea el IGF, podremos considerar como mejor el ajuste del VT al VE. Este indicador no considera la forma de la curva del VT.

El tercer elemento está dado por cuestiones relacionadas con la estimación por kriging:

a. **Validación cruzada:** Se estima cada punto P<sub>i</sub> de los datos a partir del variograma obtenido y usando solo el resto de los datos. Las diferencias entre los valores estimados de W y los valores originales son buenos indicadores de la eficiencia del modelo.

b. **Errores de estimación:** Para fines prácticos lo mas importante es que los errores de estimación sean mínimos. Es posible obtener una red dos veces mas densa que la que contiene los datos y obtener los errores de estimación para cada punto, los cuales desde el punto de vista porcentual con respecto a los valores estimados, pueden dar una idea general y local bastante precisa de la eficiencia del modelo.

Para fines prácticos de estimación para ciertos casos donde el VE es no decreciente en todos los intervalos de h, se han utilizado como modelo teórico un *spline* lineal (gráficamente es una línea quebrada) que por ser un interpolador exacto garantiza que el IGF sea nulo.

Otro modelo de VT que se han usado en algunos casos de funciones no decrecientes con alcance y meseta determinados ha sido un ajuste mínimo cuadrado condicionado lo cual consistió en buscar, usando el *princípio de los mínimos cuadrados*, los coeficientes reales K<sub>1</sub>, K<sub>2</sub>, K<sub>3</sub> que mejor ajustan la función variograma γ(h) = K<sub>1</sub>e<sup>-αh</sup> + K<sub>2</sub>e<sup>-βh</sup> + K<sub>3</sub>e<sup>-γh</sup>; α, β y γ son valores reales diferentes entre sí dos a dos (en la práctica se obtienen buenos resultados para 0.1, 0.25 y 0.6 respectivamente) y además se le imponen condiciones para que dicha función pase por los puntos (0, C<sub>0</sub>) y (h<sub>0</sub>, M<sub>0</sub>). En estos casos el IGF ha sido muy pequeño y los resultados de las estimaciones satisfactorios.

**Determinación de la zona de influencia y su relación con la anisotropía**

Hasta ahora se ha hablado de "medir la variabilidad de W" pero no se ha mencionado un problema de importancia fundamental y es el hecho de que los fenómenos geológicos y mineros que estudia la geostatística no se comportan de la misma forma en todas las zonas ni en todas las direcciones. Esto se expresa mediante los conceptos de *comportamiento isotrópico*

o *anisotrópico* de W. Un fenómeno se dice *anisotrópico* cuando presenta direcciones particulares de variabilidad, Chica Olmo (1988), esto quiere decir que, en un punto, la influencia que se recibe desde otros puntos puede tener intensidades diferentes en diferentes direcciones y además puede suceder que, en algunas direcciones, a partir de cierta distancia no exista ninguna influencia. Todo esto, generalmente se describe mediante una zona de influencia con forma de *elipse* (caso del plano) o de *elipsoide* (caso del espacio) de *anisotropía*; la longitud de los radios en cada dirección está determinada por los alcances y la dirección del mayor alcance con respecto al semieje positivo OX define (caso del plano) el ángulo α o (caso del espacio) los ángulos α y β de anisotropía, este último con respecto al plano o XY. Los ángulos mencionados tienen dominio [0°, 180°] y [-90°, 90°] respectivamente.

La anisotropía puede ser detectada obteniendo los variogramas en diferentes direcciones planas o espaciales. En la práctica se distinguen tres tipos de anisotropía:

1. **Anisotropía geométrica:** En estos casos, los variogramas presentan el mismo valor de meseta pero diferentes alcances en diferentes direcciones y mediante un factor de ponderación que tenga en cuenta los alcances mínimos y máximo y los ángulos de anisotropía puede resolverse el problema.

Sabemos que la distancia euclidiana define a

$$h_p = d(P_1, P_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

para el caso del plano y

$$h_e = D(P_1, P_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}$$

en el caso del espacio.

Lo que necesitamos es obtener un valor ponderado de la distancia h (al que distinguiremos por h<sub>p</sub><sup>p</sup> y h<sub>e</sub><sup>p</sup>) de manera que en la dirección de (caso del plano) el ángulo α o (caso del espacio) de los ángulos α y β y en sus respectivas direcciones perpendiculares se tengan los mismos valores de la distancia ponderada.

Si denotamos en el caso del plano a A<sub>1</sub> como el radio de la dirección principal y A<sub>2</sub> como el radio de la dirección perpendicular, esto se puede conseguir mediante la fórmula:

$$h_p^p = h_a \left[ \frac{1}{A_1} + \left( \frac{1}{A_2} - \frac{1}{A_1} \right) |\sin(\alpha - \theta)| \right]$$

donde θ es el ángulo que forman los puntos P<sub>1</sub> y P<sub>2</sub>.

En el caso del espacio denotamos a A<sub>1</sub> como el radio mayor en el plano XY; A<sub>2</sub> como el radio perpendicular a A<sub>1</sub> en el nuevo plano horizontal y A<sub>3</sub> como el radio perpendicular al nuevo plano horizontal. Entonces se tiene la fórmula:

$$h_e^p = h_a \left[ \frac{1}{A_1} + \left( \frac{1}{A_2} - \frac{1}{A_1} \right) |\sin(\alpha - \theta)| + \left( \frac{1}{A_3} - \frac{1}{A_1} \right) |\sin(\gamma - \mu)| \right]$$

donde θ es el ángulo que forma la proyección del segmento que une los puntos P<sub>1</sub> y P<sub>2</sub> en el plano XY con respecto al eje OX; η = β + 90° y μ es el ángulo que forma el segmento que une a P<sub>1</sub> y P<sub>2</sub> con respecto a la proyección de dicho segmento en el plano XY.

Una forma clásica de realizar esta transformación en el plano puede verse en Chica Olmo (1988).

2. **Anisotropía de efecto proporcional:** Se manifiesta mediante variogramas de iguales alcances y diferentes mesetas. En estos casos se toma un único variograma γ<sub>0</sub>(h) y para los cálculos se multiplica por un factor que es función de la dirección: D(dirección); o sea γ(h, dirección) = D(dirección) γ<sub>0</sub>(h).

3. **Anisotropía zonal:** Esta puede manifestarse de dos formas: a. Variogramas de diferentes alcances y mesetas: En este caso existe estacionaridad y deben combinarse los dos casos anteriores.

b. Algunos variogramas no presentan mesetas: Este caso hay que analizarlo con extremo cuidado pues varias son las posibles explicaciones que están relacionadas con las dimensiones de la red de muestreo, con la confección del variograma teórico y con la presencia de tendencias (*drift*).

En este trabajo generalmente se ha analizado en el plano variogramas en 5 clases o intervalos de direcciones: [0°, 30°], [30°, 60°], [60°, 90°], [90°, 120°] y [150°, 180°] y en lugar de elipses de anisotropía hemos utilizado *splines* lineales en coordenadas polares que pueden describir curvas cerradas mas complejas que una elipse. Para el caso del espacio se han utilizado *splines* bilineales. Kazvo (1996), Legrá (1997), que permiten describir superficies cerradas; los intervalos del ángulo α medidos para el plano XY son [0°, 45°], [45°, 90°], [90°, 135°] y [135°, 180°] y en el eje OZ, tomando como referencia el plano XY, medimos β en los intervalos [-90°, -45°], [-45°, 0°], [0°, 45°] y [45°, 90°]. Estos valores han permitido barrer todas las direcciones posibles de cada caso y además en periodos aceptables de tiempo de cálculo en computadora se han obtenido resultados que expresan con aproximaciones satisfactorias las características de los fenómenos.

Criterios mas complejos para el tratamiento de los intervalos de direcciones posibles pueden encontrarse en Pannatier (1994).

**Comentarios sobre la estimación mediante kriging**

Este método de estimación llamado también BLUE (*Best Linear Unbiased Estimator* o sea mejor estimador lineal insesgado) es una herramienta fácil de usar y solo requiere de medios para resolver Sistemas de ecuaciones lineales (SEL). El kriging mas conocido es el llamado puntual y el valor estimado se calcula, en general, como  $W = \sum_{i=1}^p a_i W_i$ , donde p es el número de datos que intervienen en la media ponderada.

Para obtener los valores de a, se distinguen cuatro casos, Chica Olmo (1988):

1. W es una función aleatoria estacionaria de esperanza conocida:

Sea C(h) la covarianza, esperanza M<sub>a</sub> y varianza σ<sup>2</sup>. Resolver el SEL cuadrado:

$$\left\{ \sum_{j=1}^p C(h_{ij}) a_j = C(h_{i0}), \text{ donde } j=1, \dots, p. \right.$$

donde h<sub>ij</sub> es la distancia entre P<sub>i</sub> y P<sub>j</sub>. Además h<sub>i0</sub> es la distancia entre P<sub>i</sub> y P siendo este último el punto donde se estima.

$$W = M_a + \sum_{i=1}^p a_i (W_i - M_a)$$

El error de estimación está dado por  $E = \sigma^2 - \sum_{i=1}^p a_i C(h_{i0})$

2. W es una función aleatoria estacionaria de esperanza desconocida:

Sea C(h) la covarianza y varianza σ<sup>2</sup>. Resolver el SEL cuadrado:

$$\left\{ \begin{aligned} \sum_{i=1}^p C(h_{ij}) a_i + \mu &= C(h_{j0}) \\ \sum_{i=1}^p a_i &= 1 \end{aligned} \right.$$

donde μ es un multiplicador de Lagrange y a la última ecuación se le llama Condición de No Sesgo.

$$W = \sum_{i=1}^p a_i W_i$$

El error de estimación está dado por

$$E = \sigma^2 - \sum_{i=1}^p a_i C(h_{i0}) + \mu$$

3. W es una función aleatoria intrínseca y no existe covarianza:

Sea γ(h) el variograma. Resolver el SEL cuadrado:

$$\left\{ \begin{aligned} \sum_{i=1}^p \gamma(h_{ij}) a_i + \mu &= \gamma(h_{j0}) \\ \sum_{i=1}^p a_i &= 1 \end{aligned} \right. \text{ y luego } W = \sum_{i=1}^p a_i W_i$$

El error de estimación está dado por  $E = \sum_{i=1}^p a_i \gamma(h_{i0}) + \mu$

4. W es una función aleatoria no estacionaria:

Este es el caso mas complejo y, sin dar los detalles, se dice que se han dado dos soluciones relacionadas con:

a. Búsqüeda con el modelo de kriging universal que plantea una descomposición de la variable W en dos componentes, una de ellas determinística, como combinación lineal de funciones independientes, que representa la tendencia del fenómeno y la otra aleatoria (parte residual). Entre varias críticas que se le han hecho a este método sobresale la que plantea que el variograma de la parte residual es una estimación sesgada del variograma verdadero.

Para los autores de este trabajo ha sido atractiva la idea de separar la parte determinística de la parte aleatoria y se ha usado de la siguiente manera:

Sea  $W=M(P)$  una función que describe la tendencia de  $W$ ; calculemos los puntos  $V_i = W_i - M(P)$ . Sea el variograma  $\gamma(h)$  de los puntos  $V_i$ . Resolver el SEL cuadrado:

$$\begin{cases} \sum_{i=1}^n \gamma(h_{ij}) a_i + \mu = \gamma(h_{j0}) \\ \sum_{i=1}^n a_i = 1 \end{cases} \quad \text{y luego } W = M(P) + \sum_{i=1}^n a_i V_i$$

El error de estimación del kriging está dado por

$$E = \sum_{i=1}^n a_i \gamma(h_{j0}) + \mu$$

Como puede apreciarse se trata de considerar una nueva variable regionalizada  $V$  y trabajar sobre ella en lugar de  $W$ ; la única dificultad que tiene este método está relacionada con la determinación de la función  $M$  pero esto ha sido resuelto mediante regularización de los datos y el uso de los splines bilineales y bicúbicos para el plano y los splines trilineales y tricúbicos para el espacio, Legrá (1997). La dificultad práctica de este método está dada por el hecho de que el variograma debe obtenerse después de conocerse  $M(P)$  y este último debe tomarse a partir de las características inconvenientes del variograma lo cual puede provocar un proceso laborioso y complejo.

b. El segundo método está relacionado con la teoría de las funciones intrínsecas de orden  $K$  que resuelve de manera satisfactoria los problemas de inferencia estadística.

Debe destacarse que en las fórmulas se habla de  $p$  puntos que intervienen en la estimación, esto se debe a que cuando se hace kriging se utilizan solo aquellos puntos que por estar dentro de la zona de influencia pueden ser útiles para obtener el valor estimado. En ocasiones, debido a que los SEL que aparecen son de alto orden, conviene definir el número máximo de puntos que intervendrán y esto se logra mediante una reducción radial de la zona de influencia.

El kriging es un interpolador exacto y además es un estimador que garantiza que los valores estimados de  $W$  están acotados por el menor y el mayor valor de  $W$ , pero tal como se planteó anteriormente debe prestarse especial cuidado a los valores que se estiman y a los errores que se obtienen de acuerdo al variograma que se tenga. Debe resultar «sospechoso» cualquier variograma con zonas de convexidad hacia arriba (este es el llamado efecto de parábola). Se puede ilustrar con un ejemplo sencillo:

Sea el variograma que cumple que  $\gamma(0)=0$ ,  $\gamma(0,5)=0,5$ ,  $\gamma(1)=1$  y  $\gamma(1,5)=2$ . Considerando que se quiere estimar el valor de  $W$  para  $P=(1,5,0)$  a partir de los puntos

$(0,0,4)$  y  $(1,0,1)$ . Aplicando lo visto para el caso 3 de kriging puntual no es difícil obtener que  $a_1=-0,25$ ,  $a_2=1,25$  y  $\mu=0,75$  por lo que  $W=-0,5$  y  $E=1$ . En este caso, siendo positivo el valor de  $E$  no parece que esta estimación presente dificultades, sin embargo hay que hacer notar que si todos los valores de  $W$  son no negativos (cosa que no expresa el variograma) entonces el valor estimado no está acotado por el menor y el mayor valor de  $W$ ; esto advierte para que se desconfíe cuando aparece un valor negativo de  $a_i$ . El problema puede ser aún peor: si se cumple que  $\gamma(1,5)=10$  entonces  $a_1=-4,25$ ,  $a_2=5,25$  y  $\mu=4,75$  entonces se tiene que  $W=-11,75$  y  $E=-38,4375$ .

Una forma de resolver estos problemas es revisar el variograma y determinar la posible existencia de un drift pero además se debe estar atentos a las anomalías locales; una solución puede ser la de no usar los puntos que generan los valores negativos de los coeficientes  $a_i$ . Para esto, en el trabajo práctico, los autores de este trabajo han definido un rango de negatividad admisible que puede ser pequeño o por otra parte, se eliminan todos los puntos que generan coeficientes  $a_i$  con valores negativos.

Finalmente se debe mencionar que con variaciones metodológicas han surgido otras formas de estimar con kriging, Chica Olmo (1988), Deraisme (1996), Marcotte (1988), como por ejemplo *kriging universal* (ya mencionado), *co-kriging*, *kriging disyuntivo*, *análisis krigante*, *teoría de las funciones de recuperación*, *teoría de simulación de explotación*, *funciones aleatorias intrínsecas de orden  $K$*  (ya mencionado), etc. De la misma manera, en los últimos tiempos, han surgido otros conceptos más complejos que estudian nuevos aspectos de la geoestadística, Pannatier (1994).

### CONCLUSIONES

Dos son las conclusiones que se quieren enunciar:


1. El análisis variográfico junto con la estimación por kriging es actualmente una poderosa herramienta que permite resolver dos problemas comunes del profesional geólogo-minero: modelar la variabilidad de un variable aleatoria y estructural del plano o del espacio y realizar estimaciones de nuevos valores de estas variable o de valores relacionados con ella.
2. Aunque estas técnicas se han popularizado (gracias a sus éxitos prácticos, a la existencia de bibliografía teórica y práctica de diferentes niveles y sobre todo a la existencia de varios software que las incluyen), no se puede confiar en recetas y algoritmos más o menos ingeniosos sino que se debe conocer a fondo todo lo relacionado con ellas y sobre todo recordar que sobre cada caso que se estudie se puede escribir, por sus singularidades, otro manual de recetas prácticas.

### BIBLIOGRAFÍA

- ALFONSO ROCHE, JOSE R.: *Estadísticas en las ciencias geológicas*, tomo II, ISPJAE, La Habana, 1989.
- ANNIS, ALWYN E.: *Mineral Deposit Evaluation. A Practical Approach*, Chapman & Hall, London, 1991.
- CHICA OLMO, MARIO: *Análisis geoestadístico en el estudio de la explotación de los recursos minerales*, Universidad de Granada, España, 1988.
- DERAISME, JACKES Y CHANTAL DE FOUQUET: «The Geostatistical Approach for Reserves», en *Mining Magazine*, USA, May, 1996.
- DEHL, P. y M. DAVID: «Classification of Ore Reserve/Resources based on Geostatistical Methods», en *Bulletin CIM*, February, 1982.
- FERGUSON, GAVIN: «Estimación de recursos y reservas (Una guía sobre las preguntas principales que un inversionista debería formularse)», en *Mining Magazine* (edición en español), Londres, Mayo, 1998.
- GARCÍA GUERRA, PEDRO A.: *Geoestadística Operacional*, Dpto. Nacional de Producción Mineral, Brasil, 1988.
- GUTIERREZ MONTANO, AIDA y LILIAM M. BEYRA MATOS: «Introducción al análisis variográfico de yacimientos de corteza de interperismo», Trabajo de Diploma, ISMM, Moa, 1993.
- KAYU YAMAMOTO, JORGE y MARCELO MONTEIRO DA ROCHA: «Análisis y recomendaciones para el cálculo y clasificación de reservas

minerales», (traducción de Lázaro V. Guzmán Castillo y Jorge L. López Álvarez), Brasil, 1996.

- LEGRÁ LOBAINA, ARISTIDES A. y RAFAEL A. TRUJILLO CODRINI: «Algoritmo para la obtención del spline K-cúbico natural», trabajo presentado en el evento COMPUMAT 97, Cienfuegos, Cuba, noviembre de 1997.
- MARCOTTE, D. y M. DAVID: «Trend Surface Analysis as a Special Case of IIRF-k kriging», en *Mathematical Geology*, vol. 20, no. 7, 1988.
- MERKES, J.M.: «Geostatistics of Voodoo Statistics», en *Engineering and Mining Journal*, vol. 193, no. 9, Illinois, USA, September 1992.
- MYERS, DONALD E.: «Interpolation and Estimation with spatially located data», en *Chemometrics and Intelligent Laboratory Systems*, no. 11, Elsevier Science Publishers B.V., Amsterdam, 1991.
- MYERS, DONALD E.: «To Be or Not to Be... Stationary? That is the Question», en *Mathematical Geology*, vol. 21, no. 3, 1989.
- PANNATIER, YVANN: «VariWm 2.1», en *Phd Project*, University of Lausanne, Lausanne, Switzerland, 1994.
- RODRÍGUEZ MORA, OSVALDO y JULIO GÓMEZ HERRERA: *Método de regionalización de variables: kriging*, Centro de Investigaciones del Petróleo, MINBAS, Ciudad de La Habana, 1988. □



# Maestrías

ADMINISTRACIÓN DE EMPRESAS AGROPECUARIAS

GEOLOGÍA

CIENCIAS FORESTALES

# EDUCACIÓN

INFORMÁTICA APLICADA

CULTIVO DEL TABACO