

GEOESTADÍSTICA APLICADA

MARTÍN A. DÍAZ VIERA

Instituto de Geofísica, UNAM
Instituto de Geofísica y Astronomía, CITMA, Cuba
e-mail: mdiaz@tonatiuh.igeofcu.unam.mx

2002

ÍNDICE

Capítulos		Págs.
1	INTRODUCCIÓN	1
1.1	Origen, definición y objeto de estudio	1
1.2	Etapas del análisis geoestadístico	1
1.3	La Geoestadística, su posición y su relación con respecto a otras ramas de la estadística.	2
1.4	Campos de aplicación	2
2	CONCEPTOS BÁSICOS	3
2.1	Variables Regionalizadas	3
2.2	Función Aleatoria	3
2.3	Función de distribución y momentos de una función aleatoria	4
2.4	Funciones aleatorias estacionarias	5
2.5	Funciones aleatorias intrínsecas	6
2.6	Funciones aleatorias no estacionarias	7
3	ANÁLISIS ESTRUCTURAL	9
3.1	El Semivariograma. Definición	9
3.2	Estimadores del semivariograma	9
3.3	Consideraciones prácticas para el cómputo del semivariograma muestral	12
3.4	Formas generales del semivariograma muestral	13
3.5	Condiciones que deben cumplir los modelos del semivariograma	15
3.6	Modelos autorizados para el semivariograma	17
3.6.1	Modelos transitivos o acotados	17
3.6.2	Modelos no acotados	21
3.7	Semivariogramas Anisotrópicos	25
3.8	Modelación del variograma experimental	26
3.8.1	Métodos de ajuste	26
3.8.2	Criterios de Selección de los Modelos	27
3.9	Validación del modelo del semivariograma	28
4	KRIGING	31
4.1	El mejor estimador lineal insesgado	31
4.2	Planteamiento del problema general	31
4.3	Ecuaciones del Kriging ordinario	33
4.4	Ecuaciones del Kriging ordinario en forma matricial	35
4.5	Clasificación de los diferentes tipos de Kriging	37
4.6	Ejemplos de Kriging lineal más usuales	38
4.6.1	Kriging lineal con valores esperados conocidos: Kriging Simple	38
4.6.2	Kriging lineal con valor esperado estacionario pero desconocido: Kriging Ordinario	39
4.6.3	Kriging lineal en presencia de tendencia : Kriging Universal	40
4.7	Kriging por Bloques	41
4.8	Kriging en presencia de no estacionaridad	42
4.8.1	Modos habituales de manejar la no estacionaridad	43
4.8.2	Aspectos prácticos del Kriging Universal	43
4.8.3	El modelo de las Funciones Aleatorias Intrínsecas de orden k	46

4.8.4	Kriging con mediana suavizada	47
4.8.5	Kriging Residual	49
4.9	Kriging Indicador	51
4.9.1	La Función Indicador	52
4.9.2	Variograma Indicador	54
4.9.3	Estimación de la función de distribución de probabilidad acumulativa	55
4.10	Kriging logarítmico	57
5	GEOSTADISTICA MULTIVARIADA	59
5.1	Introducción	59
5.2	Definiciones de los momentos cruzados de segundo orden	59
5.3	Propiedades de los momentos cruzados de segundo orden	60
5.4	Cokriging en el caso estacionario	60
5.5	Cokriging en el caso de funciones aleatorias intrínsecas	63
5.6	Cokriging en el caso no estacionario	64
5.7	Estimación de los momentos cruzados de segundo orden	65
5.8	Modelación de los momentos cruzados de segundo orden	66
5.9	Modelo de correogionalización lineal	68
5.10	Análisis estructural multivariado	69
5.11	Validación del modelo de semivariograma cruzado	71
5.12	Ejemplos de Aplicaciones del Cokriging	71
5.12.1	El caso de dos variables correlacionadas	71
5.12.2	Estimación de combinaciones lineales de variables	73
5.12.3	El problema de variables pobremente muestreadas	75
5.13	Modelo de correlación intrínseco	76
5.14	Análisis de Componentes Principales basado en el Modelo Lineal de Correogionalización	79
5.15	Dificultades y Aspectos Prácticos del Cokriging	80
5.16	Algunos métodos alternativos al Cokriging	80
5.16.1	Kriging combinado con Regresión Lineal	80
5.16.2	Kriging con una Deriva Externa	81
5.17	Kriging Factorial	82
5.18	Esquema General del Análisis de Correogionalización	85
5.19	Manejo de procesos espacio-temporales	86
6	SIMULACIÓN CONDICIONAL	88
6.1	Introducción	88
6.2	Objetivos de la simulación	89
6.3	Condicionamiento	90
6.4	¿ Simulación o estimación ?	91
6.5	La teoría de la simulación condicional	92
6.6	Métodos de simulación de funciones aleatorias más conocidos	94
6.7	Métodos del tipo Gaussiano	95
6.8	Método matricial	96
6.9	Método espectral	98
6.10	Método de bandas rotantes	101
6.11	Métodos secuenciales	103
6.12	Métodos del tipo Recocido Simulado	104
Anexos		
A	CONCEPTOS BÁSICOS DE ESTADÍSTICA	106
A.1	Estadística univariada	106

A.2	Estadística multivariada	116
A.3	Regresión lineal y mínimos cuadrados	122
B	GLOSARIO DE TÉRMINOS GEOESTADÍSTICOS	124
	BIBLIOGRAFÍA BÁSICA	130

CAPITULO 1: INTRODUCCIÓN

1.1 Origen, definición y objeto de estudio

En los años 60, Matheron acuñó el término de *Geoestadística*. Reconocido como el padre de esta disciplina, Matheron formalizó y generalizó matemáticamente un conjunto de técnicas desarrolladas por D. G. Krige (1941) que explotaban la *correlación espacial* para hacer predicciones en la evaluación de reservas de las minas de oro en Sudáfrica. Él definió a la Geoestadística como "la aplicación del formalismo de las *funciones aleatorias* al reconocimiento y estimación de fenómenos naturales" (Matheron, 1962). Si bien más adelante nos detendremos en el concepto de una función aleatoria, baste decir de momento que puede visualizarse como la asociación de una variable aleatoria a cada punto del espacio.

La geoestadística es una rama de la estadística aplicada que se especializa en el análisis y la modelación de la variabilidad espacial en ciencias de la tierra. Su objeto de estudio es el análisis y la predicción de fenómenos en espacio y/o tiempo, tales como: ley de metales, porosidades, concentraciones de un contaminante, etc. Aunque el prefijo geo- es usualmente asociado con geología, sin embargo la geoestadística tiene sus orígenes en la minería.

1.2 Etapas del análisis geoestadístico

Actualmente, la geoestadística es un conjunto de técnicas usadas para analizar y predecir valores de una propiedad distribuida en espacio o tiempo. En contraposición con la estadística clásica o convencional, tales valores no se consideran independientes, por el contrario se suponen de manera implícita que están correlacionados unos con otros, es decir que existe una dependencia espacial. Intuitivamente esto indica que mientras más cercanos estén situados dos puntos están más correlacionados y mientras más separados hay menos relación entre estos.

El proceso de estimación y modelación de la función que describe la correlación espacial es conocido como "análisis estructural". Una vez realizado el análisis estructural, la

predicción de valores en puntos no muestrales se puede hacer aplicando la técnica de interpolación "kriging" o simulándolos a través de “simulaciones condicionales”.

En resumen, a grosso modo un análisis geoestadístico está compuesto por tres etapas: (a) el análisis exploratorio de los datos, (b) el análisis estructural y (c) las predicciones (kriging o simulaciones)

La primera etapa, conocida como análisis exploratorio de datos, está basada en técnicas estadísticas convencionales que nos permiten obtener todo un conjunto de información, desconocida a priori sobre la muestra bajo estudio, que es imprescindible para realizar “correctamente” cualquier análisis estadístico y en particular un análisis geoestadístico.

1.3 La Geoestadística, su posición y su relación con respecto a otras ramas de la estadística.

Muchos de las ideas de la geoestadística han sido inspiradas en su hermana menor: las series cronológicas o series de tiempo. Se puede advertir que los objetivos del Análisis de Series de Tiempo son similares a los de la Geoestadística. Mientras que el Análisis de Series Temporales está orientado al estudio de procesos unidimensionales con datos muestrales uniformemente distribuidos, la Geoestadística se ocupa del estudio de fenómenos con datos distribuidos de forma arbitraria en el espacio y tiempo, por lo que la metodología de ésta última tiene un carácter mucho más general.

En un marco más amplio, la geoestadística es una disciplina que pertenece a la estadística espacial.

1.4 Campos de aplicación

La geoestadística ha sido ampliamente aplicada en diversas ramas de las ciencias aplicadas y en las ingenierías, entre otras tenemos: petróleo, minería, pesca, geofísica marina, hidrogeología, medio ambiente, estudios forestales, salud pública, ingeniería civil, procesamiento de imágenes, cartografía, finanzas, ciencias de materiales, meteorología, edafología, etc.

CAPITULO 2: CONCEPTOS BÁSICOS

2.1 Variables Regionalizadas

La Geoestadística es la aplicación de la Teoría de las Variables Regionalizadas a la estimación de procesos o fenómenos geológicos en el espacio (Matheron 1962).

Se nombra como *variable regionalizada* $z(\underline{x})$ a la variable distribuida en el espacio de manera tal que presenta una estructura espacial de correlación.

Una definición más rigurosa matemáticamente equivalente consistiría en decir que una variable regionalizada es una variable aleatoria z definida en un punto del espacio \underline{x} . Donde en el caso más general \underline{x} es un punto en el espacio tridimensional, es decir $\underline{x} = (x_1, x_2, x_3)$.

2.2 Función Aleatoria

Si a cada punto \underline{x} que pertenece a un dominio en el espacio le hacemos corresponder una variable aleatoria $z(\underline{x})$, que en sentido general pueden ser dependientes, entonces el conjunto de variables aleatorias espacialmente distribuidas $\{z(\underline{x}), \underline{x} \in \Omega\}$ será una función aleatoria $Z(\underline{x})$.

Al tomar una muestra de una función aleatoria $Z(\underline{x})$, a la que llamaremos realización, se obtendrá una función espacial discreta $Z' = \{Z(\underline{x}_i), \underline{x}_i \in \Omega, i = 1, \dots, n\}$ la cual constituye una variable regionalizada. Es decir una realización de una función aleatoria $Z(\underline{x})$ es una variable regionalizada Z' .

En lo sucesivo no haremos distinción entre la función aleatoria y su realización.

2.3 Función de distribución y momentos de una función aleatoria

Sea $Z(\underline{x})$ una función aleatoria definida en \mathbb{R}^3 , entonces el vector aleatorio $\{Z(\underline{x}_1), Z(\underline{x}_2), \dots, Z(\underline{x}_n)\}$ se caracteriza por su función de distribución de probabilidad n -variada:

$$F_{Z(\underline{x}_1), Z(\underline{x}_2), \dots, Z(\underline{x}_n)}(z_1, z_2, \dots, z_n) = \Pr[Z(\underline{x}_1) \leq z_1, Z(\underline{x}_2) \leq z_2, \dots, Z(\underline{x}_n) \leq z_n] \quad (2.1)$$

El conjunto de todas las distribuciones para todo valor de n y para cualquier selección de puntos en \mathbb{R}^3 constituye la ley espacial de probabilidad de la función aleatoria $Z(\underline{x})$. Esta función en la práctica es imposible de determinar y sólo se puede esperar inferir los primeros momentos de la distribución de $Z(\underline{x})$. En las aplicaciones en geoestadística lineal resulta suficiente estimar los momentos hasta de segundo orden, no obstante en la mayoría de los casos la información disponible no permite inferir momentos de orden superior.

Momentos de la distribución de $Z(\underline{x})$

- El momento de primer orden de $Z(\underline{x})$ es la esperanza matemática definida como:

$$m(\underline{x}) = E[Z(\underline{x})] \quad (2.2)$$

- Los momentos de segundo orden considerados en geoestadística son:

i) La varianza de $Z(\underline{x})$

$$\sigma^2(\underline{x}) = \text{Var}[Z(\underline{x})] = E[\{Z(\underline{x}) - m(\underline{x})\}^2] \quad (2.3)$$

ii) La covarianza de dos variables aleatorias $Z(\underline{x}_i)$ y $Z(\underline{x}_j)$ definida como:

$$C(\underline{x}_i, \underline{x}_j) = E[\{Z(\underline{x}_i) - m(\underline{x}_i)\}\{Z(\underline{x}_j) - m(\underline{x}_j)\}] \quad (2.4)$$

Esta función es también conocida como función de autocovarianza.

iii) El semivariograma $\gamma(\underline{x}_i, \underline{x}_j)$ que se define como:

$$2\gamma(\underline{x}_i, \underline{x}_j) = \text{Var}[Z(\underline{x}_i) - Z(\underline{x}_j)] \quad (2.5)$$

$$\gamma(\underline{x}_i, \underline{x}_j) = \frac{1}{2} E\left[\left\{Z(\underline{x}_i) - Z(\underline{x}_j)\right\}^2\right] \quad (2.6)$$

También conocido como función de semivarianza. Además el variograma es $2\gamma(\underline{x}_i, \underline{x}_j)$ pero con frecuencia se usa el término indistintamente para designar a $\gamma(\underline{x}_i, \underline{x}_j)$.

Se debe notar que tanto la varianza como el variograma son siempre positivos, mientras que la covarianza puede tomar valores negativos.

2.4 Funciones aleatorias estacionarias

Se dice que una función aleatoria es *estrictamente estacionaria* si su función de distribución Ec.(2.1) es invariante a cualquier traslación respecto a un vector \underline{h} o lo que es equivalente, la función de distribución del vector aleatorio $\{Z(\underline{x}_1), Z(\underline{x}_2), \dots, Z(\underline{x}_n)\}$ es idéntica a la del vector $\{Z(\underline{x}_1 + \underline{h}), Z(\underline{x}_2 + \underline{h}), \dots, Z(\underline{x}_n + \underline{h})\}$ para cualquier \underline{h} .

Puesto que, como se planteó anteriormente, usualmente se trabaja sólo con los momentos hasta de segundo orden resulta práctico limitar la hipótesis de estacionaridad a estos primeros momentos.

Se dice que una función aleatoria es *estacionaria de segundo orden* si se cumple que:

i) su valor esperado existe y no depende de \underline{x} ,

$$E[Z(\underline{x})] = m; \quad \forall \underline{x} \quad (2.7)$$

ii) para cualquier par de variables aleatorias $Z(\underline{x})$ y $Z(\underline{x} + \underline{h})$, su covarianza existe y sólo depende del vector de separación \underline{h} ,

$$C(\underline{h}) \equiv C(\underline{x} + \underline{h}, \underline{x}) = E[Z(\underline{x} + \underline{h})Z(\underline{x})] - m^2 \quad (2.8)$$

La estacionaridad de la varianza implica que la varianza existe, es finita y no depende de \underline{x} , es decir

$$\sigma^2 = C(\underline{0}) = \text{Var}[Z(\underline{x})] \quad (2.9)$$

Así mismo bajo esta hipótesis el semivariograma también es estacionario y se cumple que:

$$\gamma(\underline{h}) \equiv \gamma(\underline{x} + \underline{h}, \underline{x}) = \frac{1}{2} E[\{Z(\underline{x} + \underline{h}) - Z(\underline{x})\}^2] \quad (2.10)$$

Además existe una relación directa entre el semivariograma y la función de covarianza

$$\gamma(\underline{h}) = C(\underline{0}) - C(\underline{h}) \quad (2.11)$$

En este caso resulta suficiente usar una de las dos funciones para caracterizar la dependencia espacial.

2.5 Funciones aleatorias intrínsecas

Existen funciones aleatorias $Z(\underline{x})$ que representan a fenómenos físicos que muestran una capacidad casi ilimitada de variación, por lo que para estas funciones no están definidas la varianza ni la covarianza. Sin embargo existen casos en que sus incrementos o diferencias $Z(\underline{x} + \underline{h}) - Z(\underline{x})$ tienen una varianza finita. En otras palabras, esto quiere decir que las diferencias son estacionarias de segundo orden.

Por lo tanto las *funciones aleatorias intrínsecas* son aquellas que cumplen las siguientes condiciones:

i) El valor esperado de las diferencias es

$$E[Z(\underline{x} + \underline{h}) - Z(\underline{x})] = 0 \quad (2.12)$$

ii) La varianza de las diferencias es

$$Var[Z(\underline{x} + \underline{h}) - Z(\underline{x})] = 2\gamma(\underline{h}) \quad (2.13)$$

Estas condiciones se conocen como *Hipótesis Intrínseca*.

Es evidente que una función aleatoria estacionaria de segundo orden es siempre intrínseca. Lo contrario no se cumple. A las funciones que cumplen con la hipótesis intrínseca se les considera como *débilmente estacionarias*.

2.6 Funciones aleatorias no estacionarias

Las *funciones aleatorias no estacionarias* son aquellas cuya esperanza matemática depende de \underline{x} :

$$E[Z(\underline{x})] = m(\underline{x}) \quad (2.14)$$

A $m(\underline{x})$ se le conoce como función de deriva o tendencia.

Si consideramos a la función aleatoria $Z(\underline{x})$ como la suma de una componente determinística $m(\underline{x})$ y de un residuo $R(\underline{x})$ estacionario con media nula, es decir:

$$Z(\underline{x}) = m(\underline{x}) + R(\underline{x}) \quad (2.15)$$

Entonces vemos que el semivariograma de $Z(\underline{x})$ depende de \underline{x}

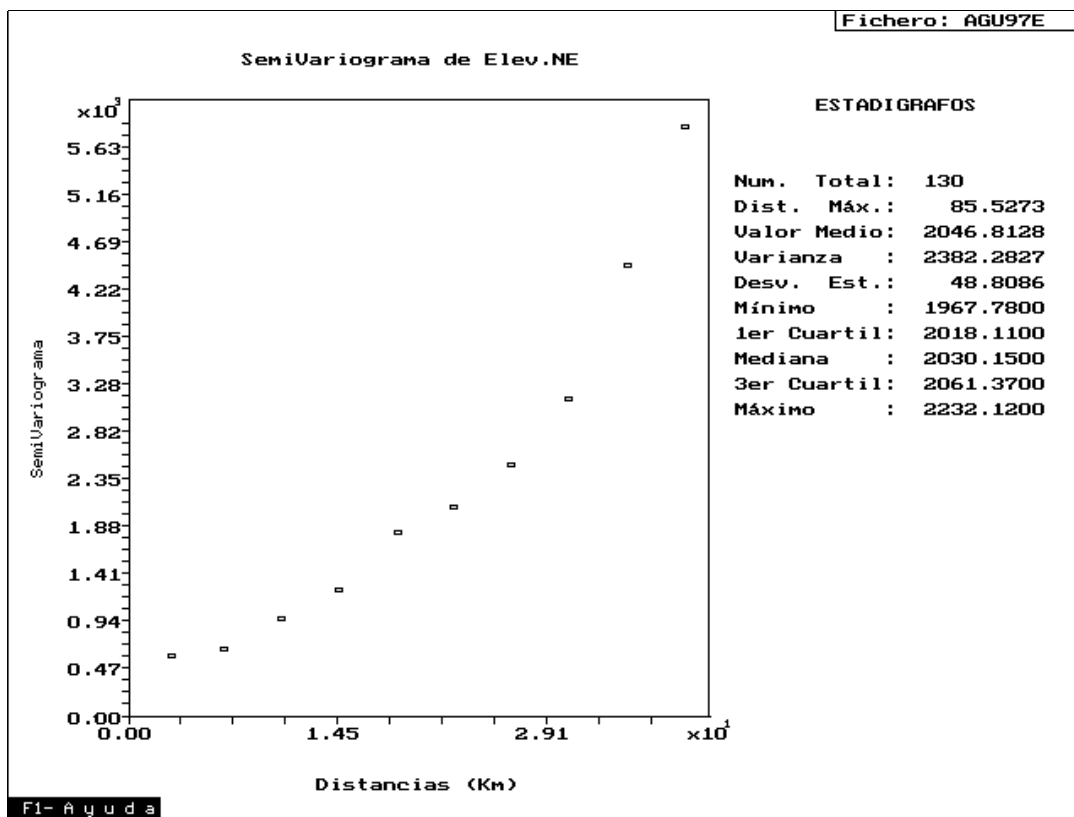
$$\gamma(\underline{x} + \underline{h}, \underline{x}) = \gamma_R(\underline{h}) + \frac{1}{2} \{m(\underline{x} + \underline{h}) - m(\underline{x})\}^2 \quad (2.16)$$

En el caso en que la deriva sea lineal $m(\underline{x}) = m_0 + \underline{m}_1 \cdot \underline{x}$ el semivariograma no depende de \underline{x}

$$\gamma(\underline{h}) = \gamma_R(\underline{h}) + \frac{1}{2}(\underline{m}_1 \cdot \underline{h})^2 \quad (2.17)$$

pero crece con el cuadrado de \underline{h} , lo cual puede servir de indicador para la detección de la existencia de tendencia.

Existe un enfoque que considera a las funciones aleatorias no estacionarias como *intrínsecas de orden k*. Esto quiere decir que si se toman las diferencias de un orden k apropiado éstas resultan ser estacionarias.



CAPITULO 3: ANÁLISIS ESTRUCTURAL

El análisis estructural es uno de los tópicos mas importantes de la geoestadística puesto que se encarga de la caracterización de la estructura espacial de una propiedad o fenómeno regionalizado. Es el proceso en el marco del cual se obtiene un modelo geoestadístico para la función aleatoria que se estudia.

En pocas palabras podemos decir que el análisis estructural consiste en estimar y modelar una función que refleje la correlación espacial de la variable regionalizada a partir de la adopción razonada de la hipótesis más adecuada acerca de su variabilidad. Esto quiere decir, que en dependencia de las características de estacionaridad del fenómeno se modelará la función de covarianzas o la de semivarianzas.

Por su importancia y generalidad estudiaremos el proceso de estimación y modelación de la función de semivarianzas o semivariograma.

3.1 El Semivariograma. Definición.

El semivariograma, conocido tambien como variograma, es la herramienta central de la geoestadística. Dada una variable regionalizada $Z(\underline{x})$ que cumpla la Hipótesis Intrínseca entonces existe la función semivarianza y se define como sigue:

$$\gamma(\underline{h}) = \frac{1}{2} Var[Z(\underline{x}) - Z(\underline{x} + \underline{h})] = \frac{1}{2} E\left[\{Z(\underline{x}) - Z(\underline{x} + \underline{h})\}^2\right] \quad (3.1)$$

El semivariograma es una función que relaciona la semivarianza con el vector \underline{h} conocido como "lag", el cual denota la separación en distancia y dirección de cualquier par de valores $Z(\underline{x})$ y $Z(\underline{x} + \underline{h})$.

3.2 Estimadores del semivariograma.

La forma de estimación mas común del semivariograma viene dada por

$$\tilde{\gamma}(\underline{h}) = \frac{1}{2N(\underline{h})} \sum_{i=1}^{N(\underline{h})} [Z(\underline{x}_i + \underline{h}) - Z(\underline{x}_i)]^2 \quad (3.2)$$

donde $N(\underline{h})$ es el número de pares $Z(\underline{x}_i)$ y $Z(\underline{x}_i + \underline{h})$ separados a una distancia $h = |\underline{h}|$.

Debido a que $\tilde{\gamma}(\underline{h})$ es esencialmente una media muestral, tiene todas las desventajas comúnmente asociadas a este tipo de estimador como es la no robustez.

Sus características son las siguientes:

- Es un estimador no paramétrico.
- Es óptimo cuando se dispone de una malla regular de muestreo que sea representativa y la distribución sea normal. En estas condiciones el sesgo es el mínimo posible.

En la práctica a menudo el empleo de este estimador produce variogramas experimentales erráticos, lo cuál se debe a desviaciones del caso ideal para la aplicación del mismo.

Estas desviaciones pueden ser enumeradas por su importancia en el orden siguiente:

- Desviaciones en la distribución.

$Z(\underline{x})$ y $Z(\underline{x} + \underline{h})$ no pueden ser representadas apropiadamente por una distribución binormal. Con frecuencia $Z(\underline{x})$ es sesgada y con grandes colas.

- Heterocedasticidad.

Se dice que los datos son heterocedásticos, si el diagrama del valor medio m contra la desviación estándar σ , calculadas en vecindades móviles, muestra que la dispersión de los valores (medidos por σ) está relacionada con su magnitud (medida por m). Por la tanto, por ser el semivariograma una medida de la dispersión, su magnitud está ligada a la magnitud de los valores de los datos. Cuando se da este tipo de fenómeno, se dice que existe un efecto proporcional.

- Desviaciones en el muestreo.

Las observaciones $\{Z(\underline{x}_i), i = 1, \dots, n\}$ de la función aleatoria $Z(\underline{x})$ no están distribuidas espacialmente en forma aleatoria. Consecuentemente se produce un sesgo en el muestreo es decir en la elección de las localizaciones de las observaciones, tendientes a estar agrupadas en áreas de altos o bajos valores. Estas desviaciones pueden ser detectadas mediante un gráfico del número de muestras n contra el valor medio m para vecindades móviles de un tamaño determinado.

- Existencia de valores atípicos (*outliers*).

Normalmente las observaciones son correctas, pero a veces suelen aparecer observaciones discordantes que pueden ser debidas a errores humanos y/o mal funcionamiento de los instrumentos de medición. Si es posible constatar que un valor de estos es erróneo, lo mejor es eliminarlo, aunque esto no siempre es factible. De ahí que la eliminación de estas observaciones puedan causar una subestimación del potencial de un área.

Estimadores alternativos del variograma.

- El estimador de Cressie y Hawkins

Este estimador constituye una alternativa robusta al estimador tradicional y se define como:

$$\gamma^*(h) = \frac{1}{2} (0.457 + 0.494/N(h))^{-1} \left[\frac{1}{N(h)} \sum_{i=1}^{N(h)} (|Z(x_i + h) - Z(x_i)|)^{1/2} \right]^4 \quad (3.3)$$

Aunque el estimador de Cressie y Hawkins se considera óptimo en condiciones de normalidad, no obstante, automáticamente infravalora los datos contaminados, mientras que en el estimador tradicional los términos cuadráticos exageran la contaminación.

- Estimadores de ponderación univariados.

En este grupo de estimadores alternativos se incluyen tanto los basados en la función de distribución acumulativa, como aquellos basados en la función de densidad de probabilidad.

3.3 Consideraciones prácticas para el cómputo del semivariograma muestral.

Para el cómputo del semivariograma es necesario tener en cuenta algunas reglas prácticas que permiten elevar la eficiencia y la calidad de la estimación, independientemente del tipo de estimador que se utilice.

Estas reglas son las siguientes:

- En la estimación del semivariograma los pares de las observaciones se agrupan según la distancia dentro de un intervalo $h = |h|$ con una tolerancia $\pm \Delta h/2$ y dentro de una dirección θ con una tolerancia $\pm \Delta \theta/2$. El semivariograma así estimado es considerado suavizado o regularizado.

- El semivariograma muestral debe ser considerado solamente para pequeñas distancias por lo que generalmente, se estima para valores de $|h|$ menores que la mitad de la distancia máxima ($|h| < d_{\max}/2$).
- La elección del número de intervalos es arbitraria. No obstante se considera que un número máximo de 25 intervalos es suficiente para cualquier propósito, y un mínimo de 10 debe ser usado para determinar con precisión el rango y la meseta del semivariograma.
- El largo de los intervalos debe ser elegido de forma tal que el número de pares en cada intervalo sea lo suficientemente grande para que el estimado del semivariograma sea relativamente estable. Se considera que entre 30 y 50 pares satisfacen este requerimiento.
- Los valores estimados para cada intervalo se deben graficar contra la distancia promedio de todos los pares que se encuentran dentro de dicho intervalo.

3.4 Formas generales del semivariograma muestral.

En sentido amplio se considera por su forma que hay dos tipos principales de semivariogramas. En el primer tipo, la semivarianza se incrementa con el incremento del valor absoluto del intervalo $|h|$ hasta alcanzar un valor máximo a partir del cuál se mantiene relativamente constante y oscila alrededor del mismo. Estos semivariogramas son conocidos como de tipo transitivo. El valor del intervalo a partir del cual el semivariograma no se incrementa es conocido como alcance o rango (radio de correlación) y marca el límite de la dependencia espacial de la propiedad. La varianza máxima es conocida como "sill" o meseta del semivariograma y teóricamente debe coincidir con la varianza *a priori* σ^2 de la muestra de la función aleatoria $Z(\underline{x})$.

Una variable con este tipo de semivariograma no sólo cumple la hipótesis intrínseca, sino también es estacionaria de segundo orden. Esto significa que su valor esperado es constante

$$E[Z(\underline{x})] = m; \quad \forall \underline{x} \quad (3.4)$$

y la función de covarianza existe y está definida por

$$\begin{aligned} C(\underline{h}) &= \text{Cov}[Z(\underline{x} + \underline{h}), Z(\underline{x})] \\ &\equiv E[\{Z(\underline{x}) - m\}\{Z(\underline{x} + \underline{h}) - m\}] \end{aligned} \quad (3.5)$$

y el semivariograma está relacionado con las funciones de covarianza y de autocorrelación como sigue:

$$\gamma(\underline{h}) = C(\underline{0}) - C(\underline{h}) \quad (3.6)$$

$$\rho(\underline{h}) = C(\underline{h})/C(\underline{0}) = 1 - \gamma(\underline{h})/C(\underline{0}) \quad (3.7)$$

donde $\sigma^2 = C(\underline{0})$ es la varianza de la muestra, también conocida como varianza *a priori*, de la función aleatoria $Z(\underline{x})$.

El segundo tipo de semivariograma aparenta un incremento sin límites, es decir son no acotados, por esto no presentan una varianza *a priori* finita.

Un aspecto del semivariograma que es importante señalar es que por definición $\gamma(\underline{0}) = 0$ pero en la práctica el semivariograma muestral $\gamma^*(\underline{h})$ cuando $|\underline{h}|$ tiende a cero no necesariamente se anula.

Esto es conocido como efecto "nugget" o pepita, y el valor del semivariograma en cero $\gamma(\underline{0})$ es conocido como la varianza "nugget" o microvarianza. En principio esto puede ocurrir solamente si existen discontinuidades en la función aleatoria. En la práctica su existencia debe a la variación espacial que no puede explicar el variograma debido a la escala del muestreo. Para un material continuamente variable el efecto *nugget* se produce a partir de la contribución de los errores de medición y la variación a distancias mucho menores que el intervalo de muestreo más pequeño.

3.5 Condiciones que deben cumplir los modelos del semivariograma

En la selección de una función adecuada para ser ajustada a un semivariograma muestral, se debe tener en cuenta en la mayoría de los casos hasta tres elementos: un intercepto con la ordenada, una sección monótonamente creciente y una meseta. Sin embargo, no servirá cualquier modelo que aparente ajustar a los valores empíricos debido a la siguiente razón.

Supongamos que $Z(\underline{x})$ es una función aleatoria estacionaria de segundo orden de la cual obtenemos la variable regionalizada $\{Z(\underline{x}_i), i = 1, \dots, n\}$, donde su semivariograma y su función de covarianzas son $\gamma(\underline{h})$ y $C(\underline{h})$ respectivamente.

Sea Y una combinación lineal de $Z(\underline{x}_i)$ tal que

$$Y = \sum_{i=1}^n \lambda_i Z(\underline{x}_i) \quad (3.8)$$

donde $\lambda_i, i = 1, \dots, n$, son pesos arbitrarios.

La magnitud Y es también una variable regionalizada con varianza:

$$Var[Y] = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j C(\underline{x}_i, \underline{x}_j) \quad (3.9)$$

La varianza de Y puede ser positiva o cero, pero no negativa y la función de covarianza en la parte derecha de la expresión anterior debe asegurar que esta condición sea satisfecha.

La matriz de covarianzas

$$\begin{bmatrix} C(\underline{x}_1, \underline{x}_1) & C(\underline{x}_1, \underline{x}_2) & \dots & C(\underline{x}_1, \underline{x}_n) \\ C(\underline{x}_2, \underline{x}_1) & C(\underline{x}_2, \underline{x}_2) & \dots & C(\underline{x}_2, \underline{x}_n) \\ \dots & \dots & \dots & \dots \\ C(\underline{x}_n, \underline{x}_1) & C(\underline{x}_n, \underline{x}_2) & \dots & C(\underline{x}_n, \underline{x}_n) \end{bmatrix} \quad (3.10)$$

debe ser definida positiva, es decir, el determinante y todos sus menores principales son positivos o cero. Aquí las covarianzas están definidas como $C(\underline{x}_i, \underline{x}_j) = Cov[Z(\underline{x}_i), Z(\underline{x}_j)]$; $i, j = 1, \dots, n$.

La función de covarianzas $C(\underline{h})$ si existe debe ser positiva semidefinida por lo que serán permisibles funciones que cumplan este criterio.

Las propiedades que no tengan covarianzas *a priori* finitas no tendrán definida la función de covarianzas. En el caso de que cumplan la hipótesis intrínseca, entonces estará definido el semivariograma y se cumple que:

$$Var[Y] = C(\underline{0}) \sum_{i=1}^n \lambda_i \sum_{j=1}^n \lambda_j - \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \gamma(\underline{x}_i, \underline{x}_j) \quad (3.11)$$

Si los pesos λ_i suman 0, es decir $\sum_{i=1}^n \lambda_i = 0$ entonces:

$$Var[Y] = - \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \gamma(\underline{x}_i, \underline{x}_j) \quad (3.12)$$

Esto implica que $-\gamma(\underline{h})$ debe ser condicionalmente positiva semidefinida, con la condición $\sum_{i=1}^n \lambda_i = 0$. Esto es equivalente a decir que $\gamma(\underline{h})$ debe ser *condicionalmente negativa semidefinida*.

Como consecuencia de esta propiedad, se puede demostrar que el semivariograma debe tener un ritmo de crecimiento inferior a h^2 , es decir se debe cumplir que:

$$\lim_{h \rightarrow \infty} \frac{\gamma(h)}{h^2} = 0 \quad (3.13)$$

Cuando no se satisface esta condición puede ser un indicador de que la función aleatoria no sea estacionaria.

Por otro lado de la definición de semivariograma se infiere que éste debe anularse en el origen, es decir $\gamma(0) = 0$.

Las funciones que cumplen las condiciones anteriores son conocidas como **modelos autorizados** del semivariograma.

El hecho de probar si una función dada es aceptable o no, está relacionado con el examen de su Transformada de Fourier. Christakos [1984] obtuvo las condiciones que el espectro de la función debe reunir para ser un modelo autorizado.

Como una propiedad importante se debe destacar que cualquier combinación lineal de modelos autorizados es un modelo autorizado.

3.6 Modelos autorizados para el semivariograma

Según la forma del semivariograma se pueden dividir en transitivos y no acotados.

3.6.1 Modelos transitivos o acotados

Este grupo de modelos se deriva a partir de la noción de autocorrelación entre los valores promedios dentro de los bloques. La idea es que la función aleatoria, de la cual la propiedad medida es una realización, depende del grado de solapamiento de los dos bloques, es decir, una zona de transición.

Modelo Lineal con Meseta.

En una dimensión los bloques son simplemente líneas. Supongamos que estos son de longitud a y la distancia entre sus centros es h . Su solapamiento es entonces $a-h$, el cual, cuando se expresa como una proporción de a es $1-(h/a)$, siendo $h \leq a$.

La función de autocorrelación es entonces:

$$\rho(h) = 1 - h/a \quad (3.14)$$

y el semivariograma se expresa como

$$\gamma(h) = \begin{cases} S(h/a) & \text{para } 0 \leq h \leq a \\ S & \text{para } h > a \end{cases} \quad (3.15)$$

Modelo Circular

En dos dimensiones, los bloques son discos de diámetro a . El área de intersección de éstos, separados a una distancia h de sus centros es:

$$A = \frac{a^2}{2} \cos^{-1}\left(\frac{h}{a}\right) - \frac{h}{2} \sqrt{a^2 - h^2}; \quad \text{para } h \leq a \quad (3.16)$$

Expresado como una fracción del área del disco obtenemos la función de autocorrelación

$$\rho(h) = \frac{2}{\pi} \left\{ \cos^{-1}\left(\frac{h}{a}\right) - \frac{h}{a} \sqrt{1 - \left(\frac{h}{a}\right)^2} \right\} \quad (3.17)$$

y el semivariograma

$$\gamma(h) = \begin{cases} S \left\{ 1 - \frac{2}{\pi} \cos^{-1} \left(\frac{h}{a} \right) - \frac{2h}{\pi a} \sqrt{1 - \left(\frac{h}{a} \right)^2} \right\} & \text{para } 0 \leq h \leq a \\ S & \text{para } h > a \end{cases} \quad (3.18)$$

$$\text{gradiente} = 4S/(\pi a)$$

Este modelo no ha encontrado mucha aplicación en las ciencias de la tierra.

Modelo Esférico

Por analogía podemos derivar este modelo considerando el solapamiento de los volúmenes de dos esferas de diámetro a y h la distancia que separa sus centros. El volumen de intersección es:

$$V = \frac{\pi}{4} \frac{2a^3}{3} - a^2 h + \frac{h^3}{3}; \quad \text{para } h \leq a \quad (3.19)$$

Dividiendo por el volumen de la esfera ($\pi a^3/6$) se obtiene la función de autocorrelación

$$\rho(h) = 1 - \frac{3h}{2a} + \frac{1}{2} \left(\frac{h}{a} \right)^3 \quad (3.20)$$

y el semivariograma

$$\gamma(h) = \begin{cases} \frac{S}{2} \left\{ 3 \left(\frac{h}{a} \right) - \left(\frac{h}{a} \right)^3 \right\} & \text{para } 0 \leq h \leq a \\ S & \text{para } h > a \end{cases} \quad (3.21)$$

$$\text{gradiente} = 3S/(2a)$$

Debe señalarse que el modelo esférico es apropiado para el caso de tres dimensiones aunque se puede aplicar para casos de una y dos dimensiones. Lo inverso no se cumple.

El modelo circular sólo es apropiado para una y dos dimensiones. El lineal solamente para el caso unidimensional. Esto se debe a que solamente en estos casos estos modelos son funciones condicionalmente negativas semidefinidas.

Modelo Exponencial

Si el solapamiento de los bloques varía su tamaño de forma aleatoria, entonces el semivariograma resulta exponencial. En el caso isotrópico es:

$$\gamma(h) = S \left(1 - e^{-\frac{h}{r}} \right) \quad \text{para } h \geq 0 \quad (3.22)$$

gradiente = S/r

Se considera como rango efectivo $a = 3r$.

Los procesos autorregresivos de primer orden y de Markov producen modelos exponenciales.

Modelo Gaussiano

Está dado por la expresión:

$$\gamma(h) = S \left(1 - e^{-\left(\frac{h}{r}\right)^2} \right) \quad \text{para } h \geq 0 \quad (3.23)$$

donde r es un parámetro no lineal que determina la escala espacial de la variación, como en el caso exponencial. El rango efectivo se considera $a = \sqrt{3}r$, que corresponde al valor $0.95S$ del variograma.

Modelo de efecto *agujero*

El efecto *agujero* es indicativo de fenómenos con componentes periódicas. Las expresiones más comunes de modelos de semivariogramas son:

$$\gamma(h) = S \left(1 - \frac{\sin(h)}{h} \right) \quad \text{para } h > 0 \quad (3.24)$$

Este puede ser usado para representar procesos regularmente continuos y que muestran un comportamiento periódico, el cual es frecuentemente encontrado, donde existe una sucesión de zonas ricas y pobres. Es un modelo negativo definido en tres dimensiones.

Otra alternativa es:

$$\gamma(h) = S(1 - \cos(h)) \quad \text{para } h \geq 0 \quad (3.25)$$

Si el efecto agujero es observado muy acentuado en cierta dirección, por ejemplo la vertical, cuando el fenómeno es una sucesión pseudoperiódica de estratificación horizontal. Este modelo es negativo definido en una dimensión.

3.6.2 Modelos no acotados

Modelo Potencia

Existen casos en que la varianza aparenta incrementarse indefinidamente. Así también si se toma cada vez un menor intervalo de muestreo, siempre existe alguna variación que queda sin resolver. Un punto de partida para entender esta variación es el movimiento Browniano en una dimensión, en el cual:

$$Z(\underline{x}) - Z(\underline{x} + \underline{h}) = \varepsilon \quad (3.26)$$

es una variable aleatoria gaussiana, espacialmente independiente.

Su semivariograma es:

$$\gamma(h) = \frac{1}{2} h^\theta; \quad \text{para } 0 < \theta < 2 \quad (3.27)$$

Mandelbrot (1982) llamó al resultado de tales procesos **fractales**.

Existe una relación entre la dimensión fractal (Hausdorff-Besicovitch) y el parámetro θ dada por la expresión

$$D = 2 - (\theta/2) \quad (3.28)$$

donde podemos ver los siguientes casos extremos

- para $\theta=2$ es una parábola y $D=1$, no representa un proceso aleatorio, además

$$\gamma(h) = kh^2 \quad (3.29)$$

no es una función condicionalmente negativa definida.

- para $\theta=0$ representa el caso de ruido puro

Con frecuencia notamos que muchos semivariogramas aparentan ser de varianza nugget pura. Esto ocurre porque toda la varianza está dentro de un intervalo de muestreo mas pequeño.

Modelo de efecto pepita puro

Formalmente se puede definir como sigue:

$$\gamma(h) = S(1 - \delta(h)) \quad (3.30)$$

Modelo logarítmico (Modelo de Wigsian)

Se define como:

$$\gamma(h) = k \log(h) \quad (3.31)$$

Puede ser de utilidad cuando el semivariograma experimental se comporta linealmente si se usa una escala logarítmica para las distancias.

Combinación de modelos

1.- La combinación lineal de semivariogramas (funciones de autocovarianza) con coeficientes positivos representa un semivariograma (función de autocovarianza). Si consideramos la función aleatoria $Z(\underline{x})$ como una combinación lineal de n funciones aleatorias independientes $Y_i(\underline{x})$,

$$Z(\underline{x}) = \sum_{i=1}^n \lambda_i Y_i(\underline{x}) \quad (3.32)$$

Al sustituir en las definiciones de semivariograma y covarianza obtenemos:

$$\gamma(h) = \sum_{i=1}^n \lambda_i^2 \gamma_i(h) \quad (3.33)$$

$$C(h) = \sum_{i=1}^n \lambda_i^2 C_i(h) \quad (3.34)$$

donde $\gamma_i(h)$ y $C_i(h)$ son el semivariograma y la función de autocovarianza de la variable $Y_i(\underline{x})$.

2.- El producto de funciones de autocovarianza es también otra función de autocovarianza. Consideremos a $Z(\underline{x})$ como el producto de n funciones aleatorias independientes $Y_i(\underline{x})$, es decir

$$Z(\underline{x}) = \prod_{i=1}^n Y_i(\underline{x}) \quad (3.35)$$

De la definición de función de autocovarianza se puede deducir que

$$C(h) = \prod_{i=1}^n C_i(h) \quad (3.36)$$

y el semivariograma correspondiente será

$$\gamma(h) = \prod_{i=1}^n C_i(0) - \prod_{i=1}^n C_i(h) \quad (3.37)$$

A partir de estas dos propiedades se pueden encontrar modelos mas complejos mediante la combinación de los modelos simples antes estudiados.

De hecho muchos de los modelos anteriores los encontramos usualmente como combinaciones por ejemplo del tipo

$$\gamma(h) = \gamma_0(h) + \gamma_1(h) \quad (3.38)$$

donde $\gamma_0(h) = S_0(1 - \delta(h))$ representa el efecto nugget puro, mientras que $\gamma_1(h)$ es cualquier otro modelo de variograma que expresa la variación espacial.

Con frecuencia se encuentra que la combinación está formada por componentes (semivariogramas) que tienen diferentes rangos, lo cual se conoce estructura anidada. Los modelos con estructuras anidadas describen variaciones espaciales a diferentes escalas y que se deben por lo tanto, a factores de naturaleza diferente.

Por ejemplo:

$$\gamma(h) = \gamma_0(h) + \gamma_1(h) + \gamma_2(h) \quad (3.39)$$

donde

$$\gamma_0(h) = S_0(1 - \delta(h)) \quad (3.40)$$

$$\gamma_1(h) = \begin{cases} \frac{S_1}{2} \left\{ 3 \left(\frac{h}{a_1} \right) - \left(\frac{h}{a_1} \right)^3 \right\}; & h \leq a_1 \\ S_1; & h > a_1 \end{cases} \quad (3.41)$$

$$\gamma_2(h) = \begin{cases} \frac{S_2}{2} \left\{ 3 \left(\frac{h}{a_2} \right) - \left(\frac{h}{a_2} \right)^3 \right\}; & h \leq a_2 \\ S_2; & h > a_2 \end{cases} \quad (3.42)$$

3.7 Semivariogramas Anisotrópicos

Todos los modelos anteriormente presentados son isotrópicos, es decir, la variabilidad espacial no depende de la dirección. En la práctica existen numerosas situaciones en que la variación es anisotrópica, por lo que en cada dirección hay semivariograma diferente.

Si la anisotropía se puede tener en cuenta mediante una transformación lineal simple de las coordenadas, entonces se dice que la anisotropía es geométrica o afín.

La transformación se puede expresar mediante la siguiente fórmula:

$$\Omega(\theta) = \left(A^2 \cos^2(\theta - \phi) + B^2 \sin^2(\theta - \phi) \right)^{1/2} \quad (3.43)$$

Esta función puede ser aplicada como un factor al argumento h de la función semivarianzas en el caso de los modelos transitivos o al gradiente en los modelos no acotados.

En la práctica se estudian 4 direcciones, estimando los semivariogramas y determinando los rangos para los mismos, y luego se construye el gráfico direccional de los rangos para decidir si hay anisotropía geométrica presente o no.

En presencia de anisotropía geométrica, el gráfico direccional de los rangos forma una elipse, en la cual el eje menor B es el rango de la dirección de mas rápida variación y A , el eje mayor, está en la dirección en que la variabilidad es mas lenta. La relación $\lambda = A/B$ es una medida de anisotropía.

En caso de anisotropía si se desea diseñar una red óptima de muestreo se debe hacer en forma rectangular haciendo coincidir los lados con las direcciones de los ejes

principales y las longitudes de los mismos deben estar en la proporción λ , donde el lado menor le correspondería a la dirección del eje B .

Mardia (1980) ha sugerido pruebas de significación formales para la relación de anisotropía. Estos parecen no haber sido usados en la práctica geoestadística.

3.8 Modelación del variograma experimental

3.8.1 Métodos de ajuste

Algunos geoestadísticos ajustan los modelos de forma visual, pero esta práctica no es fiable y es preferible usar algún procedimiento estadístico para estos fines. Con frecuencia es usada la aproximación por Mínimos Cuadrados.

- Mínimos Cuadrados Ordinarios (MCO).

Asume que los residuos están normalmente distribuidos y son independientes y que las semivarianzas estimadas poseen igual varianza.

- Mínimos Cuadrados Generalizados (MCG).

Se evitan muchos de los problemas de los MCO, pero requiere de la inversión de grandes matrices, lo que hace el procedimiento muy complicado.

- Mínimos Cuadrados Ponderados.

Se minimiza la siguiente expresión:

$$W(\tau, \gamma_k) = \sum_{j=1}^k \frac{[\gamma(h_j) - \gamma(h_j, \tau)]^2}{Var[\gamma(h_j)]} \quad (3.44)$$

donde

τ es el vector de los parámetros del modelo,

$Var[\gamma(h_j)]$ es la varianza del valor estimado del semivariograma en h_j .

Desafortunadamente esta varianza es desconocida casi siempre. Con frecuencia se escoge en su lugar el número de pares de observaciones $m(h_j)$ con que se estima la misma, suponiendo que es inversamente proporcional a la varianza de los estimados.

Cressie (1985) ha demostrado teóricamente que un mejor procedimiento sería reemplazar $Var[\gamma(h_j)]$ por $[\gamma(h_j, \tau)]^2/m(h_j)$ dando un peso considerablemente mayor a las semivarianzas esperadas menores. La diferencia en los resultados de la aplicación de estos dos tipos de pesos es pequeña y se ha confirmado que cuando se utiliza $m(h_j)$ se le asigna pesos muy pequeños a los estimados de los intervalos h_j menores.

Laslett ha sugerido un mejoramiento. Sus pesos están dados por:

$$\frac{m(h_j)\gamma(h_j)}{[\gamma(h_j, \tau)]^3} \quad (3.45)$$

y estos son actualizados en cada iteración del proceso cuando $\gamma(h_j, \tau)$ es reestimada. Teóricamente esta forma de ponderación debe dar mejor convergencia y en la práctica se ha encontrado que así es.

3.8.2 Criterios de Selección de los Modelos

Un compromiso entre la bondad de ajuste y la complejidad del modelo puede ofrecerlo el Criterio de Información de Akaike (AIC), que se define como

$$AIC = -2 \ln(\text{máx. verosimilitud}) + 2 (\text{núm. de parámetros}) \quad (3.46)$$

Y se puede estimar usando

$$(AIC)^* = \left\{ n \ln \left(\frac{2\pi}{n} \right) + n + 2 \right\} + n \ln(R) + 2p \quad (3.47)$$

Debido a que la cantidad que se encuentra entre llaves en la expresión anterior Ec.(3.47) es constante $(n \ln(2\pi/n) + n + 2 = \text{const})$ independientemente del tipo de modelo, entonces para fines prácticos se calcula:

$$\tilde{A} = n \ln(R) + 2p \quad (3.48)$$

que es un estimador simplificado del Criterio de Información de Akaike. Donde n es el número de valores estimados $\{\gamma^*(h_i), i=1, \dots, n\}$ del variograma muestral, R es la suma residual de los cuadrados de las diferencias entre los valores experimentales $\gamma^*(h_i)$ y los del modelo ajustado $\gamma(h_i)$, es decir $R = \sum_{i=1}^n (\gamma(h_i) - \gamma^*(h_i))^2$, mientras que p es el número de parámetros del modelo de variograma ajustado $\gamma(h)$.

Se considera que el modelo que presenta el menor AIC es el mejor.

3.9 Validación del modelo del semivariograma

Para validar el modelo obtenido de variograma se puede proceder de varias maneras. Un método que resulta atractivo por su sencillez y eficiencia es el *leave one out* que consiste en sacar un elemento de la muestra y estimar el valor en ese punto usando Kriging con el modelo de variograma obtenido. De forma análoga se actúa para el resto de los elementos de la muestra. Como resultado se obtiene un mapa de las diferencias $Z(\underline{x}_i) - Z^*(\underline{x}_i), i=1, \dots, n$ entre el valor real y el estimado.

De forma tal que si el modelo del semivariograma refleja adecuadamente la estructura espacial implícita en el conjunto de datos, entonces los valores estimados deben ser cercanos a los valores observados.

Esta "cercanía" puede ser caracterizada según los siguientes estadígrafos:

a) $\frac{1}{n} \sum_{i=1}^n \{Z(\underline{x}_i) - Z^*(\underline{x}_i)\}$ cercano a 0.

b) $\frac{1}{n} \sum_{i=1}^n \{Z(\underline{x}_i) - Z^*(\underline{x}_i)\}^2$ pequeño.

c) $\frac{1}{n} \sum_{i=1}^n \left\{ \frac{Z(\underline{x}_i) - Z^*(\underline{x}_i)}{\sigma_i} \right\}^2$ cercano a 1.

d) La correlación muestral de $Z(\underline{x}_i)$, $Z^*(\underline{x}_i)$ sea cercana a 1.

e) La correlación muestral de $Z(\underline{x}_i)$, $\{Z(\underline{x}_i) - Z^*(\underline{x}_i)\}/\sigma_i$ es cercana a cero.

Siendo:

$Z(\underline{x}_i)$ - los valores muestrales de la propiedad en x_i ,

$Z^*(\underline{x}_i)$ - los valores estimados de la propiedad en el punto x_i ,

σ_i - es la desviación estándar de la estimación en el punto x_i .

Idealmente todos los estadígrafos anteriores deben satisfacerse simultáneamente, pero en la práctica una mejoría en uno de ellos puede degradar a otro. Por lo que es recomendable hacer un análisis integral de los estadígrafos de las diferencias.

Mediante el histograma de los errores normalizados se compila una lista con los puntos que poseen grandes errores. Esto es útil para la identificación de valores atípicos (outliers), datos sospechosos o anomalías de otra naturaleza.

La eliminación de las localizaciones con elevados errores normalizados y el recálculo del semivariograma y su modelación puede producir mejorías significativas.

El procedimiento de *leave one out* es un caso particular del método conocido como Jackknifing, ya que de disponer de suficientes datos espacialmente distribuidos de forma homogénea de la función aleatoria $Z(\underline{x})$, éstos se podrían dividir en dos muestras: $Z_1(\underline{x})$ y $Z_2(\underline{x})$. La primera se usaría para estimar el variograma, mientras que la segunda serviría para validarlo. Es decir, se estimarían como en el método anterior los valores empleando

kriging con el modelo de variograma obtenido usando la primera muestra $Z_1(\underline{x})$ en los puntos correspondientes a las observaciones que pertenecen a la segunda muestra $Z_2(\underline{x})$ y se evaluarían los estadígrafos de las diferencias $Z_2(\underline{x}_i) - Z_2^*(\underline{x}_i)$ de manera análoga.

CAPITULO 4: KRIGING.

4.1 El mejor estimador lineal insesgado

El kriging es un término que ha sido acuñado para designar al "*mejor estimador lineal insesgado*" de un punto y al mejor promedio lineal móvil ponderado de un bloque.

Este nombre apareció alrededor de 1960 para nombrar una técnica creada en Francia por Matheron a partir de los trabajos de D. G. Krige quién fue probablemente el primero que hizo uso de la correlación espacial y del mejor estimador lineal insesgado en el campo de la evaluación de yacimientos minerales.

El kriging es una técnica de estimación local que ofrece el mejor estimador lineal insesgado de una característica desconocida que se estudia. La limitación a la clase de estimadores lineales es bastante natural ya que esto significa que solamente se requiere el conocimiento del momento de segundo orden de la función aleatoria (la covarianza o el variograma) y que en general en la práctica es posible inferir a partir de una realización de la misma.

4.2 Planteamiento del problema general

Sea $Z(\underline{x})$ una función aleatoria, la cual está definida en un soporte puntual y es estacionaria de segundo orden, con :

- un valor esperado,

$$E[Z(\underline{x})] = m; \quad \forall \underline{x} \quad (4.1)$$

donde m es una constante generalmente desconocida,

- una función de covarianza centrada

$$C(\underline{h}) = E[Z(\underline{x} + \underline{h})Z(\underline{x})] - m^2 \quad (4.2)$$

- un variograma

$$Var[Z(\underline{x} + \underline{h}) - Z(\underline{x})] = 2\gamma(\underline{h}) \quad (4.3)$$

donde al menos uno de estos dos momentos de segundo orden se supone conocido.

Cuando solamente existe el variograma, entonces la función aleatoria $Z(\underline{x})$ se considera intrínseca.

Los valores experimentales consisten en un conjunto de valores discretos $\{Z(\underline{x}_i), i = 1, \dots, n\}$. Con frecuencia estos valores están definidos en soportes puntuales o casi puntuales, en otros casos son los valores medios $Z_{V_i}(\underline{x}_i)$ definidos en los soportes V_i centrados en los puntos \underline{x}_i , donde los n soportes pueden ser todos diferentes.

La estimación del valor medio $Z_{V_i}(\underline{x}_i)$ en el dominio de V_i se define como

$$Z_{V_i}(\underline{x}_i) = \frac{1}{V_i} \int_{V_i} Z(\underline{x}) d\underline{x}$$

Es bueno destacar que bajo la hipótesis de estacionaridad el valor esperado de cada uno de estos datos es $E[Z(\underline{x}_i)] = m, \quad \forall i$.

El estimador lineal Z_k^* considerado es una combinación lineal de n valores de datos tal que:

$$Z^*(\underline{x}_k) = \sum_{i=1}^n \lambda_i Z(\underline{x}_i),$$

donde $Z_k^* = Z^*(\underline{x}_k)$.

Los n coeficientes λ_i son calculados de manera tal que el estimador sea insesgado y que la varianza de la estimación sea mínima, entonces se dice que el estimador Z_k^* es óptimo.

4.3 Ecuaciones del kriging ordinario

Como ya hemos planteado el estimador kriging se considera óptimo ya que es insesgado y es el que minimiza la varianza de la estimación. A partir de estos dos criterios se derivan las ecuaciones del kriging.

- La condición de insesgadez.

Para obtener un valor esperado del error igual a cero resulta suficiente imponer la condición:

$$E[Z_k^*] = E\left[\sum_{i=1}^n \lambda_i Z(\underline{x}_i)\right] = m; \quad (4.4)$$

donde m es el valor esperado de la función aleatoria $Z(\underline{x})$.

Esto implica que

$$\sum_{i=1}^n \lambda_i E[Z(\underline{x}_i)] = E[Z_k^*] \quad (4.5)$$

entonces

$$\sum_{i=1}^n \lambda_i m = m \quad (4.6)$$

y finalmente

$$\sum_{i=1}^n \lambda_i = 1 \quad (4.7)$$

- La condición de que la estimación sea de mínima varianza

Para satisfacer esta condición hay que minimizar la siguiente función:

$$F = \sigma_e^2 - 2\mu \left(\sum_{i=1}^n \lambda_i - 1 \right) \quad (4.8)$$

donde

σ_e^2 - es la varianza de la estimación,

μ - un multiplicador de Lagrange.

Nótese que la función F a minimizar consta de la varianza de la estimación σ_e^2 e incluye la condición que garantiza el no sesgo de la estimación.

La varianza de la estimación se expresa de la siguiente manera:

$$\sigma_e^2 = Var[Z_k - Z_k^*] = E[(Z_k - Z_k^*)^2] \quad (4.9)$$

$$\sigma_e^2 = Var[Z_k] - 2Cov[Z_k, Z_k^*] + Var[Z_k^*] \quad (4.10)$$

Sustituyendo en esta última fórmula la expresión del estimador Z_k^* tenemos:

$$\sigma_e^2 = Var[Z_k] - 2Cov\left[Z_k, \sum_{i=1}^n \lambda_i Z(x_i)\right] + Var\left[\sum_{i=1}^n \lambda_i Z(x_i)\right] \quad (4.11)$$

desarrollando obtenemos

$$\sigma_e^2 = \sigma_{Z_k}^2 - 2 \sum_{i=1}^n \lambda_i \sigma_{Z_k Z_i} + \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \sigma_{Z_i Z_j} \quad (4.12)$$

Si hallamos las derivadas parciales de F respecto a los coeficientes desconocidos λ_i y con respecto a μ obtenemos el siguiente sistema de ecuaciones:

$$\begin{cases} \frac{\partial F}{\partial \lambda_i} = -2\sigma_{Z_k Z_i} + 2 \sum_{j=1}^n \lambda_j \sigma_{Z_i Z_j} - 2\mu = 0, & i = 1, \dots, n \\ \frac{\partial F}{\partial \mu} = \sum_{i=1}^n \lambda_i - 1 = 0 \end{cases} \quad (4.13)$$

De una manera más usual se escribe como sigue:

$$\begin{cases} \sum_{j=1}^n \lambda_j \sigma_{Z_i Z_j} - \mu = \sigma_{Z_k Z_i}, & i = 1, \dots, n \\ \sum_{i=1}^n \lambda_i = 1 \end{cases} \quad (4.14)$$

El sistema de ecuaciones así obtenido sirve para el cálculo del Kriging Puntual. Y la varianza del error de la estimación se puede calcular de una manera más simple si se sustituye el valor de μ :

$$\sigma_e^2 = \sigma_{Z_k}^2 - \sum_{i=1}^n \lambda_i \sigma_{Z_k Z_i} + \mu \quad (4.15)$$

4.4 Ecuaciones del Kriging ordinario en forma matricial

De forma análoga podemos representar el sistema de ecuaciones del Kriging en forma matricial como sigue:



$$\begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1n} & 1 \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2n} & 1 \\ \dots & \dots & \dots & \dots & \dots \\ \sigma_{n1} & \sigma_{n2} & \dots & \sigma_{nn} & 1 \\ 1 & 1 & \dots & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \dots \\ \lambda_n \\ -\mu \end{bmatrix} = \begin{bmatrix} \sigma_{k1} \\ \sigma_{k2} \\ \dots \\ \sigma_{kn} \\ 1 \end{bmatrix} \quad (4.16)$$

donde $\sigma_{ij} = \sigma_{Z_i Z_j}$

Nótese que bajo la hipótesis intrínseca las covarianzas pueden ser reemplazados por las semivarianzas, usando la expresión:

$$\sigma_{ij} = \sigma^2 - \gamma_{ij} \quad (4.17)$$

donde

$\sigma^2 = C(0)$, es la varianza total o *a priori* de la muestra.

Entonces sustituyendo la expresión de la covarianza en función de la semivarianza obtenemos el sistema del Kriging en función de las semivarianzas:



$$\begin{bmatrix} 0 & \gamma_{12} & \dots & \gamma_{1n} & 1 \\ \gamma_{21} & 0 & \dots & \gamma_{2n} & 1 \\ \dots & \dots & \dots & \dots & \dots \\ \gamma_{n1} & \gamma_{n2} & \dots & 0 & 1 \\ 1 & 1 & \dots & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \dots \\ \lambda_n \\ \mu \end{bmatrix} = \begin{bmatrix} \gamma_{k1} \\ \gamma_{k2} \\ \dots \\ \gamma_{kn} \\ 1 \end{bmatrix} \quad (4.18)$$

Y la varianza de la estimación se calcula como sigue:

$$\sigma_e^2 = \sum_{i=1}^n \lambda_i \gamma_{ki} + \mu \quad (4.19)$$

Estas últimas expresiones resultan ser la formulación del Kriging más común puesto que se aplica en el caso cuando la variable regionalizada cumple la hipótesis intrínseca la cual es una condición menos exigente que la estacionaridad de segundo orden.

4.5 Clasificación de los diferentes tipos de Kriging

I) Según la forma del estimador

- lineales:

Simple

Ordinario

Universal

Residual

- no lineales:

Disyuntivo

Indicador

Probabilístico

II) Según el soporte de la medición de los datos

- puntual

- en bloques

III) Kriging paramétrico y no paramétrico

- paramétrico:

Multigaussiano

Disyuntivo

Lognormal

- no paramétrico:

Simple

Ordinario

Universal

Residual

Indicador

Probabilístico

4.6 Ejemplos de Kriging lineal más usuales

Todos los estimadores "kriging" pueden ser interpretados como proyecciones de un valor desconocido $Z(\underline{x})$ en el conjunto de los posibles estimadores. Mientras mas amplio sea el conjunto en el cual es hecha la proyección mas cercano estará el estimador kriging correspondiente del valor desconocido y se necesitarán más requisitos.

4.6.1 Kriging lineal con valores esperados conocidos: Kriging Simple.

Sistema de ecuaciones:

$$\begin{cases} \sum_{j=1}^n \lambda_j \sigma'_{ij} = \sigma'_{i0}, & i = 1, \dots, n \\ \lambda_0 = m(\underline{x}_0) - \sum_{i=1}^n \lambda_i m(\underline{x}_i) \end{cases} \quad (4.20)$$

Estimador:

$$Z_0^* = \lambda_0 + \sum_{i=1}^n \lambda_i Z(\underline{x}_i) \quad (4.21)$$

Varianza de la estimación:

$$\sigma_{K_S}^2 = \sigma'_{00} - \sum_{i=1}^n \lambda_i \sigma'_{i0} \quad (4.22)$$

donde

$m(\underline{x}_i) = E[Z(\underline{x}_i)]$ es el valor esperado en el punto x_i ,

$\sigma'_{ij} = \sigma_{ij} - m(\underline{x}_i)m(\underline{x}_j)$ - covarianza centrada.

Requisitos:

- Conocer $n+1$ valores esperados $m(\underline{x}_i) = E[Z(\underline{x}_i)]$, $\forall i = 0, \dots, n$ de la función aleatoria $Z(\underline{x})$.
- Conocer la función de covarianzas σ_{ij} de la función aleatoria $Z(\underline{x})$.

4.6.2 Kriging lineal con valor esperado estacionario pero desconocido: Kriging Ordinario

Sistema de ecuaciones:

$$\begin{cases} \sum_{j=1}^n \lambda_j \sigma_{ij} - \mu = \sigma_{0i}, & i = 1, \dots, n \\ \sum_{i=1}^n \lambda_i = 1 \end{cases} \quad (4.23)$$

Estimador:

$$Z_0^* = \sum_{i=1}^n \lambda_i Z(x_i) \quad (4.24)$$

Varianza de la estimación:

$$\sigma_{K_0}^2 = \sigma_{00} - \sum_{i=1}^n \lambda_i \sigma_{i0} + \mu \quad (4.25)$$

Requisitos:

- Se requiere que el valor esperado $m(\underline{x}_i) = E[Z(\underline{x}_i)] = m, \forall i = 1, \dots, n$ de la función aleatoria $Z(\underline{x})$ sea constante
- Conocer la función de covarianzas σ_{ij} o el semivariograma γ_{ij} de la función aleatoria $Z(\underline{x})$.

4.6.3 Kriging lineal en presencia de tendencia : Kriging Universal

Sistema de ecuaciones:

$$\begin{cases} \sum_{j=1}^n \lambda_j \sigma_{ij} - \sum_{l=1}^L \mu_l \phi_l(\underline{x}_i) = \sigma_{0i}, & i = 1, \dots, n \\ \sum_{i=1}^n \lambda_i \phi_l(\underline{x}_i) = \phi_l(\underline{x}_0), & l = 1, \dots, L \end{cases} \quad (4.26)$$

Estimador:

$$Z_0^* = \sum_{i=1}^n \lambda_i Z(\underline{x}_i) \quad (4.27)$$

Varianza de la estimación:

$$\sigma_{K_U}^2 = \sigma_{00} - \sum_{i=1}^n \lambda_i \sigma_{i0} + \sum_{l=1}^L \mu_l \phi_l(\underline{x}_0) \quad (4.28)$$

Requisitos :

- Conocer la forma de la tendencia $m(\underline{x}) = E[Z(\underline{x})]$ de la función aleatoria $Z(\underline{x})$ expresada mediante funciones conocidas $\phi_l(\underline{x})$, usualmente polinomios.
- Conocer la función de covarianzas σ_{ij} o el semivariograma γ_{ij} de la función aleatoria $Z(\underline{x})$ sin tendencia, es decir $\{Z(\underline{x}) - m(\underline{x})\}$.

4.7 Kriging por Bloques

Hemos visto hasta el momento estimaciones en un punto, en el caso cuando se desea conocer el valor promedio sobre una región o bloque, se formulan las ecuaciones del Kriging de manera análoga al caso puntual.

En el kriging por bloques en lugar de estimar el valor en un punto \underline{x}_k se considera una región V_k de área A_k con centro en el punto \underline{x}_k . El estimador tiene la siguiente forma:

$$Z_{V_k}^* = \sum_{i=1}^n \lambda_i Z(\underline{x}_i) \quad (4.29)$$

En las ecuaciones del Kriging puntual el vector del miembro derecho las semivarianzas γ_{ij} son reemplazadas por las semivarianzas promedios con respecto al bloque V_k que se expresan como:

$$\gamma_{V_k, i} = \frac{1}{A_k} \int_{V_k} \gamma(\underline{x} - \underline{x}_i) d\underline{x} \quad (4.30)$$

Entonces las ecuaciones del Kriging por bloques serán:

$$\begin{bmatrix} 0 & \gamma_{12} & \dots & \gamma_{1n} & 1 \\ \gamma_{21} & 0 & \dots & \gamma_{2n} & 1 \\ \dots & \dots & \dots & \dots & \dots \\ \gamma_{n1} & \gamma_{n2} & \dots & 0 & 1 \\ 1 & 1 & \dots & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \dots \\ \lambda_n \\ \mu \end{bmatrix} = \begin{bmatrix} \gamma_{V_k,1} \\ \gamma_{V_k,2} \\ \dots \\ \gamma_{V_k,n} \\ 1 \end{bmatrix} \quad (4.31)$$

Y la varianza de la estimación se expresa como

$$\sigma_{V_k}^2 = \sum_{i=1}^n \lambda_i \gamma_{V_k,i} + \mu - \gamma_{V_k,V_k} \quad (4.32)$$

donde

$$\gamma_{V_k V_k} = \frac{1}{A_k^2} \int_{V_k} \int_{V_k} \gamma(\underline{x} - \underline{y}) d\underline{x} d\underline{y} \quad (4.33)$$

La estimación por bloques resulta más suave que la estimación puntual.

4.8 Kriging en presencia de no estacionaridad

El mejor predictor sin suponer linealidad es $E[Z_k / Z_1, \dots, Z_n]$. Para calcularlo es necesario conocer las distribuciones conjuntas de $(n+1)$ dimensiones lo cual es imposible en la práctica. Cuando la función aleatoria $Z(\underline{x})$ es un proceso Gaussiano el mejor predictor es el estimador lineal. En lo adelante se impondrá que fue hecha una transformación adecuada que convierte el problema en datos gaussianos (con posibles outliers aditivos), por supuesto, la predicción se necesita en la escala original y al invertir las transformaciones se deben hacer correcciones para el sesgo y la varianza. Cuando este tipo de transformación no es obvia han sido propuesto un número de enfoques como son:

- disyuntivo (Matheron 1976)
- multigaussiano (Verly 1983)

- indicador (Journel 1983)
- probabilístico (Suliman 1984)

El objetivo fundamental de este capítulo es estudiar como aplicar el kriging en presencia de no estacionaridad (tendencia).

4.8.1 Modos habituales de manejar la no estacionaridad

Supongamos que exista no estacionaridad en la media

$E[Z(\underline{x})] = m(\underline{x})$, donde $m(\underline{x})$ - es una función espacial y se conoce como tendencia o deriva.

Han sido propuestas dos vías del kriging en presencia de deriva no constante :

- 1) Se supone que $m(\underline{x})$ puede ser representado como un polinomio de orden finito k .
- 2) Se postula que una cierta diferencia de orden finito k de la función aleatoria $Z(\underline{x})$ es débilmente estacionaria (o intrínseca de orden k).

El primer método es conocido como **Kriging Universal** (Matheron 1969), mientras que el segundo es el de las **Funciones Aleatorias Intrínsecas de orden k** (Matheron 1973).

4.8.2 Aspectos Prácticos del Kriging Universal

En este método deben ser conocidos a priori el orden k del polinomio que mejor describe o explica la tendencia y la función de semivarianzas o variograma γ de la función aleatoria $Z(\underline{x})$ sin tendencia.

Los supuestos anteriores nos conducen en la práctica a dos problemas :

- a) el orden k del polinomio nunca es conocido, hay que adivinarlo.
- b) El variograma γ tampoco es conocido y hay que estimarlo a partir de los residuales (datos – deriva), es decir $R(\underline{x}) = Z(\underline{x}) - m(\underline{x})$.

Matheron señaló (1971) que tales estimadores son sesgados y que a menos que el variograma de los residuos sea conocido, digamos que a partir de su estimación en una dirección donde no se perciba la deriva, el kriging universal tiene serias dificultades operacionales.

Como ya se ha visto el kriging universal puede ser expresado de manera muy simple. El problema real es que cuando existe una sola realización de la función aleatoria no estacionaria $Z(\underline{x})$ resulta imposible estimar el variograma. Lo que uno puede hacer es intentar eliminar la deriva $m(x)$ y trabajar con los residuos $R(\underline{x}) = Z(\underline{x}) - m(\underline{x})$ los cuales son estacionarios o al menos intrínsecos. Esto significa que la deriva tiene que ser estimada a partir de los valores muestrales. Aquí comienzan las dificultades.

Debido a que este estimador tiene que ser no sesgado y de varianza mínima uno obtiene un conjunto de ecuaciones para calcular los coeficientes de la deriva expresada como un polinomio. Pero para reducir las ecuaciones que dan este estimador óptimo es necesario conocer el variograma de la función aleatoria, lo cual es precisamente el último propósito del estudio. Sin embargo, un estimador no sesgado simple es válido y puede ser derivado a partir de los métodos de mínimos cuadrados del análisis de superficie de tendencia.

Nosotros somos ahora capaces de calcular un variograma experimental de los residuos estimados $\gamma_R^*(\underline{h})$ pero Matheron (1969) ha mostrado que tal variograma difiere del variograma de los residuales verdaderos $\gamma_R(\underline{h})$ y que el sesgo es una función de la forma del estimador de $m(\underline{x})$.

Se puede ver que en la estimación del variograma es donde está la clave del problema. No hay solución directa conocida, uno puede solo hacer un conjunto de suposiciones y tratar de verificarlos mediante el método de prueba y error, por lo que entonces el Kriging Universal se transforma en un procedimiento heurístico.

El proceso entero puede ser resumido como sigue :

- Se necesita conocer el variograma $\gamma(\underline{h})$ de $Z(\underline{x})$, pero como éste no puede ser directamente estimado uno trata de estimar el variograma de los residuos $\gamma_R(\underline{h})$.
- Para esto se requiere seleccionar un tamaño (un radio r_v) de la vecindad y dentro de ésta se supone un tipo de deriva $m(\underline{x})$, en general se toma un polinomio en \underline{x} de cierto orden k . Es decir, $m^*(\underline{x}) = p_k(\underline{x})$, donde $p_k(\underline{x})$ es un polinomio de orden k .
- Los coeficientes del polinomio de la deriva son estimados haciendo una suposición del tipo del variograma $\gamma_R(\underline{h})$.
- Con los coeficiente de la deriva podemos obtener los residuos $R(\underline{x}_i)$ en los puntos muestrales \underline{x}_i .
- Entonces se calcula el variograma experimental de los residuos $\gamma_R^*(\underline{h})$.
- El variograma teórico esperado de los residuos es calculado $\gamma_R(\underline{h})$.
- El variograma teórico supuesto al inicio $\gamma_R(\underline{h})$ y el experimental resultante de los residuos $\gamma_R^*(\underline{h})$ son comparados en algún sentido de su bondad de ajuste según la norma $\|\gamma_R^*(\underline{h}) - \gamma_R(\underline{h})\|$ que se elija.

Si el ajuste es razonable (no existen pruebas aún para la bondad de ajuste) la vecindad tomada, el tipo de deriva y el variograma supuestos son correctos. En caso contrario, uno de los parámetros del procedimiento es cambiado y el proceso comienza otra vez. El

algoritmo es simple, pero desde el punto de vista práctico puede consumir una gran cantidad de tiempo y no ofrece ninguna garantía de que exista convergencia.

Otro criterio que permitiría discriminar en qué medida son correctos o no los supuestos del procedimiento podría ser el empleo del método de validación cruzada (ver sección 3. del modelo (combinación de variograma y deriva) que se obtiene en cada paso. En este sentido una elección razonable sería tomar el modelo que arroje el mejor resultado en términos de los estadígrafos (valor medio y desviación estándar) de los errores $Z(\underline{x}_i) - Z^*(\underline{x}_i)$ producto de la validación cruzada de los valores estimados $Z^*(\underline{x}_i)$ con los valores reales $Z(\underline{x}_i)$ de la muestra.

4.8.3 El modelo de las Funciones Aleatorias Intrínsecas de orden k

El segundo método de kriging en presencia de tendencia es el de las Funciones Aleatorias Intrínsecas de orden k , el cual posee como supuesto un modelo más general que el primero, pero en la práctica se reduce a adivinar el orden k y estimar una función de covarianzas generalizada a partir de las diferencias de orden k .

Su utilidad ha estado limitada debido a :

- a) no ha sido posible de forma general estimar la covarianza generalizada de forma no paramétrica.
- b) no es fácil interpretar los parámetros de la covarianza generalizada.
- c) el orden k de las diferencias tiene que ser adivinado (se toma usualmente 2).
- d) los efectos de frontera conducen a una reducción drástica de aquellos puntos que pueden ser estimados por este método.

4.8.4 Kriging con mediana suavizada

Parte de la siguiente premisa

$$\begin{array}{lcl} \text{Valores observados} = & \text{variación en gran escala} & + \text{variación en pequeña escala} \\ & (\text{deriva}) & (\text{error}) \end{array}$$

El enfoque de trabajo propuesto es dejar que la posición espacial de los datos determine la escala de variación (la deriva es a gran escala). Esto concuerda con el análisis exploratorio de datos, en que las decisiones sobre el modelo se basan más en lo que se ve que en lo que pudiera haberse visto.

Se puede ajustar con exactitud una superficie a través de los puntos observados, pero hay poco poder predictivo en este enfoque. Se quiere extraer la deriva *aparente* y tratar de modelar el resto mediante estacionaridad débil.

El enfoque *spline* es similar, donde se permite que el tamaño de la rejilla determine la variabilidad a gran escala, conjuntamente con las condiciones adicionales acerca de la diferenciabilidad y la suavidad de la función.

Para el tipo de aplicaciones en mente, sin embargo, es la variabilidad local la que se desea modelar, interpretar y en consecuencia explotar con el kriging.

Se toma un enfoque analítico :

$$\text{datos} = \text{ajuste} + \text{residuo}$$

Los residuos son tomados como datos originales y se realiza el mismo proceso :

$$\text{residuos} = \text{nuevo ajuste} + \text{nuevo residuo}$$

y así sucesivamente.

Supongamos que la localización de los datos es un subconjunto de puntos de una malla finita rectangular L , la cual no tiene que estar obligatoriamente completa. Cuando los datos están irregularmente distribuidos, cada dato se le asigna el nodo más cercano de una malla imaginaria superpuesta.

Cressie (1984) sugirió que el ajuste ϕ_{ij} en la localización (i,j) sea obtenido mediante la mediana de los valores observados:

$$\phi_{ij} = a + r_i + c_j$$

$$R_{ij} = Z_{ij} - \phi_{ij}$$

donde :

r_i - efecto fila, c_j - efecto columna, R_{ij} - residuos.

$$\text{med}_i (R_{ij}) = \text{med}_j (R_{ij}) = 0, \text{med}_i (r_i) = \text{med}_j (c_j) = 0.$$

En la práctica se obtuvo buenos resultados aplicando esta metodología seguida de una estimación por kriging ordinario. Esto se debió a que la variación a gran escala está basada en la mediana la cual es resistente a los outliers, y por otro lado debido a que los residuos tienen media 0, esto hace que el variograma sea menos sesgado. El resultado final es un estimador kriging robusto y resistente en presencia de no estacionaridad.

En resumen, el kriging con Mediana se ejecuta de la siguiente forma :

$$Z_{ij} = \phi_{ij} + R_{ij}$$

donde:

ϕ_{ij} - es la deriva estimada por la interpolación mediante Mediana suavizada.

R_{ij} - es una variable regionalizada (error de la estimación).

4.8.5 Kriging Residual

La hipótesis principal del kriging residual propuesto por Gambolati y Volpi (1978 y 1979) consiste en considerar conocido el orden de la deriva o tendencia $m(x)$, la cual se estima usando mínimos cuadrados ordinarios $m^*(x)$, y a partir de ésta se obtienen los residuos $R(x)$ a los que se le aplica el kriging ordinario.

- Kriging residual directo

El algoritmo se puede resumir en los siguientes pasos:

1. Suponer conocido el orden k de la deriva en base a razonamientos físicos.
2. Ajustar mediante mínimos cuadrados ordinarios la deriva $m_k^*(\underline{x})$.
3. Calcular los residuos $R(\underline{x}) = Z(\underline{x}) - m_k^*(\underline{x})$.
4. Estimar y modelar el semivariograma de los residuos $\gamma_R(\underline{h})$.
5. Aplicar kriging ordinario a los residuos $R(\underline{x})$ usando el semivariograma $\gamma_R(\underline{h})$.
6. Se obtiene la estimación en un punto no observado como $Z^*(\underline{x}) = m_k^*(\underline{x}) + R^*(\underline{x})$.

El proceso del kriging residual es internamente inconsistente ya que el método de los mínimos cuadrados ordinarios supone que los residuos sean espacialmente independientes, mientras que la existencia del semivariograma lo contradice.

- Kriging residual iterativo

El algoritmo es similar al kriging residual directo y fue propuesto por Neuman y Jacobson (1984). Su principal diferencia radica en que en su primera parte hace una búsqueda sistemática del modelo de tendencia que produzca residuos estacionarios, es decir, los puntos 1 al 4 se ejecutan cíclicamente incrementando el orden del modelo de tendencia hasta que el variograma de los residuos sea estacionario. Mientras que en su segunda parte se aplica mínimos cuadrados generalizados usando la matriz de covarianzas de los residuos (\underline{C}_R) , en lugar de mínimos cuadrados ordinarios, para la reestimación del modelo de

tendencia $m_k^*(\underline{x})$ con el objetivo de eliminar la inconsistencia de la estimación del método directo.

Esquema general del algoritmo:

1. Suponer que el orden de la deriva es $k = 1$.
2. Ajustar mediante mínimos cuadrados ordinarios la deriva $m_k^*(\underline{x})$.
3. Calcular los residuos $R(\underline{x}) = Z(\underline{x}) - m_k^*(\underline{x})$.
4. Estimar el semivariograma de los residuos $\gamma_R^*(\underline{h})$.
5. Comprobar que el variograma estimado sea estacionario. Si lo es ir al paso 6, en caso contrario incrementar el orden del modelo de tendencia $k = k + 1$ e ir al paso 2.
6. Calcular la matriz de covarianzas $(C_R)_{ij} = \sigma_R^2 - \gamma_R^*(\underline{x}_i - \underline{x}_j)$
7. Ajustar mediante mínimos cuadrados generalizados la deriva $m_k^*(\underline{x})$.
8. Calcular los residuos $R(\underline{x}) = Z(\underline{x}) - m_k^*(\underline{x})$.
9. Estimar el semivariograma de los residuos $\gamma_R^*(\underline{h})$.
10. Comprobar que el variograma estimado $\gamma_R^*(\underline{h})$ y el modelo de deriva $m_k^*(\underline{x})$ alcanzan “valores estables”. Si es así pasar al paso 11, en caso contrario ir al paso 6.
11. Modelar el semivariograma de los residuos $\gamma_R(\underline{h})$.
12. Aplicar kriging ordinario a los residuos $R(\underline{x})$ usando el semivariograma $\gamma_R(\underline{h})$.
13. Se obtiene la estimación en un punto no observado como $Z^*(\underline{x}) = m_k^*(\underline{x}) + R^*(\underline{x})$.

Existen otras modificaciones, como la de Samper (1986), propuestas al método del kriging residual que consisten esencialmente en estimar la matriz de covarianzas no a partir del variograma muestral, sino directamente a partir de los residuos. Para efectos prácticos, estas modificaciones pueden convertir al procedimiento en impracticable. Incluso, en la mayoría de los casos resulta suficiente la estimación con mínimos cuadrados ordinarios, obviando los pasos 6-10 del algoritmo anterior.

4.9 Kriging Indicador

La teoría y el desarrollo de estimadores no paramétricos de distribuciones espaciales, es similar a la teoría y el desarrollo de estimadores no paramétricos del valor medio local (Journel, 1983).

La mayor diferencia entre estos dos tipos de estimadores está en el estimado de las variables y en los tipos de datos usados. Una variable aleatoria puede ser definida en cada y cualquier locación \underline{x}_i dentro de la región de interés. El conjunto de estas variables aleatorias determina una función aleatoria. Como un conjunto de variables aleatorias, la función aleatoria expresa el comportamiento local y regional de la variable de interés. En un punto dado en el espacio, la función aleatoria $Z(\underline{x})$ es una variable aleatoria, pero sobre una región, la función aleatoria incorpora la estructura de correlación espacial completa de cualquier subconjunto de variables aleatorias. Entonces la función aleatoria describe los aspectos aleatorios y estructurados de la variable.

Para cualquier tipo de estimación, algún conocimiento de la ley espacial de la función aleatoria es necesario antes que un estimador óptimo pueda ser definido.

Debido a que los datos son de sólo una realización y porque cualquier propiedad de la función aleatoria no puede ser inferida a partir de una realización simple sin un modelo, un supuesto de estacionaridad debe ser incluido en el modelo para permitir la inferencia de las propiedades de la función aleatoria. El tipo de estacionaridad invocada para la estimación no paramétrica de distribuciones espaciales es la estacionaridad de la distribución bivariada de $Z(\underline{x})$ y $Z(\underline{x} + \underline{h})$ para varios valores del vector de distancia \underline{h} , que es:

$$F_{Z(\underline{x}_1), Z(\underline{x}_2)}(z_1, z_2) = F_{\underline{x}, \underline{x} + \underline{h}}(z_1, z_2) = F_{\underline{h}}(z_1, z_2) \quad (4.34)$$

distribución bivariada de dos variables aleatorias $Z(\underline{x})$ y $Z(\underline{x} + \underline{h})$ depende de la magnitud del vector \underline{h} que separa a estas variables aleatorias y no depende de las localizaciones particulares: $\underline{x}_1 = \underline{x}$ y $\underline{x}_2 = \underline{x} + \underline{h}$.

Esto implica estacionaridad de la distribución univariada, es decir, que las variables aleatorias que representan la variable de interés de cada localización particular $Z(\underline{x}_1)$, $Z(\underline{x}_2)$, ..., $Z(\underline{x}_n)$ están idénticamente distribuidas, con distribución $F_Z(z)$.

Con la estacionaridad de la distribución bivariada establecida, los estimadores no paramétricos de la distribución de probabilidad pueden ser desarrollados.

4.9.1 La Función Indicador

La distribución espacial de una variable aleatoria en un soporte puntual dentro de una región A puede ser definida matemáticamente (Journel 1983) como:

$$\phi(A, z_c) = \frac{1}{A} \int_{\underline{x} \in A} i(\underline{x}, z_c) d\underline{x} \quad (4.35)$$

donde

$\phi(A, z_c)$ - es la distribución espacial de valores puntuales dentro de la región A para un valor umbral z_c , es decir la proporción de valores dentro de la región A menores que z_c , es un indicador definido como:

$$i(\underline{x}, z_c) = \begin{cases} 1, & \text{si } Z(\underline{x}) \leq z_c \\ 0, & \text{si } Z(\underline{x}) > z_c \end{cases} \quad (4.36)$$

donde

$Z(\underline{x})$ - es el valor observado en el punto \underline{x} ,

z_c - es el valor umbral o de corte.

La distribución espacial $\phi(A, z_c)$ puede ser considerada como una realización de una variable aleatoria $\Phi(A, z_c)$, la cual a su vez representa la transformación integral de la variable aleatoria $Z(\underline{x})$ siguiente:

$$\Phi(A, z_c) = \frac{1}{A} \int_{\underline{x} \in A} I(\underline{x}, z_c) d\underline{x} \quad (4.37)$$

donde

$I(\underline{x}, z_c)$ - es la función indicador aleatoria definida como:

$$I(\underline{x}, z_c) = \begin{cases} 1, & \text{si } Z(\underline{x}) \leq z_c \\ 0, & \text{si } Z(\underline{x}) > z_c \end{cases} \quad (4.38)$$

El valor esperado de la variable aleatoria $\Phi(A, z_c)$ puede ser determinado como:

$$E[\Phi(A, z_c)] = \frac{1}{A} E \left[\int_{\underline{x} \in A} I(\underline{x}, z_c) d\underline{x} \right] = \frac{1}{A} \int_{\underline{x} \in A} E[I(\underline{x}, z_c)] d\underline{x} \quad (4.39)$$

$$E[\Phi(A, z_c)] = \frac{1}{A} \int_{\underline{x} \in A} \{ (1) \Pr[Z(\underline{x}) \leq z_c] + (0) \Pr[Z(\underline{x}) > z_c] \} d\underline{x} \quad (4.40)$$

$$E[\Phi(A, z_c)] = \Pr[Z(\underline{x}) \leq z_c] = F_Z(z_c) \quad (4.41)$$

donde

$F_Z(z_c)$ - función de distribución acumulativa univariada de Z estacionaria.

La función de distribución espacial $\phi(A, z_c)$ puede ser vista como una realización de una función aleatoria $\Phi(A, z_c)$, cuyo valor esperado es igual al valor de la función de distribución acumulativa univariada estacionaria $F_Z(z_c)$. La distribución de la variable indicador en los puntos muestrales $i(\underline{x}, z_c)$, será la misma con sólo dos valores 0 y 1. Si el valor observado es menor que el valor de corte, el valor indicador será 1, en caso contrario, este será 0. Claramente, la variable indicador $i(\underline{x}, z_c)$, cambiará con el

incremento del valor de corte. Más muestras tienen valores menores que el valor de corte, por lo tanto mas variables indicador tienen el valor 1.

La distribución de probabilidad de variables indicador es una distribución de Bernoulli y sus momentos están dados por:

a) valor esperado

$$E[I(\underline{x}, z_c)] = (1) \Pr[Z(\underline{x}) \leq z_c] + (0) \Pr[Z(\underline{x}) > z_c] = F_Z(z_c) \quad (4.42)$$

b) función de covarianzas

$$\begin{aligned} C_I(\underline{h}, z_c) &= E[I(\underline{x}, z_c), I(\underline{x} + \underline{h}, z_c)] - \{E[I(\underline{x}, z_c)]\}^2 = \\ &= F_{\underline{x}, \underline{x} + \underline{h}}(z_c, z_c) - \{F_Z(z_c)\}^2 \end{aligned} \quad (4.43)$$

c) varianza

$$Var[I(\underline{x}, z_c)] = C(\underline{0}, z_c) = F_Z(z_c)[1 - F_Z(z_c)] \quad (4.44)$$

d) y la función de dsemivarianzas

$$\gamma_I(\underline{h}, z_c) = C_I(\underline{0}, z_c) - C_I(\underline{h}, z_c) = F_Z(z_c) - F_{\underline{x}, \underline{x} + \underline{h}}(z_c, z_c) \quad (4.45)$$

4.9.2 Variograma Indicador

Si la función aleatoria $Z(\underline{x})$ es acotada, es decir, si existe un valor Z_{\max} tal que $Z(\underline{x}) \leq Z_{\max} \quad \forall \underline{x}$, entonces se cumple que:

$$\begin{aligned} i(\underline{x}, z_c) &= 1, \quad \forall Z \geq Z_{\max} \\ \gamma_I(\underline{h}, z_c) &= 0, \quad \forall Z \geq Z_{\max} \end{aligned} \quad (4.46)$$

La función aleatoria indicador siempre posee varianza finita y acotada en el intervalo

$$0 \leq C_I(\underline{h}, z_c) \leq 0.25 \quad (4.47)$$

Por lo tanto el variograma indicador $\gamma_I(\underline{h}, z_c)$ siempre alcanza una meseta $S(z_c)$ que coincide con $C_I(\underline{0}, z_c)$ y de acuerdo con Ec. (4.43) según (Myers, 1984) se deduce que

a) $S(z_c)$ es una función creciente de z_c en el intervalo $(-\infty, z_M)$, donde

$$F_Z(z_M) = 0.5.$$

b) $S(z_c)$ es una función decreciente en el intervalo (z_M, ∞) ,

- c) Una vez conocida la meseta del variograma $S(z_c)$ se puede calcular $F_Z(z_c)$ usando la expresión:

$$F_Z(z_c) = 0.5 + \text{signo}(z_c - z_M) \sqrt{0.25 - S(z_c)} \quad (4.48)$$

Este puede ser interpretado como una probabilidad bivariada:

$$\gamma_I(\underline{h}, z_c) = 0.5 \{ \Pr[Z(\underline{x}) > z_c, Z(\underline{x} + \underline{h}) \leq z_c] \} + 0.5 \{ \Pr[Z(\underline{x} + \underline{h}) > z_c, Z(\underline{x}) \leq z_c] \} \quad (4.49)$$

Los variogramas de indicador $\gamma_I(\underline{h}, z_c)$ son estimados para cada valor de corte dado z_c , de la misma manera que se procede usualmente en la estimación de los variogramas de funciones aleatorias intrínsecas. Hay que hacer notar que éstos son mas robustos, es decir, son menos sensibles a la presencia de valores extremos (*outliers*)

4.9.3 Estimación de la función de distribución de probabilidad acumulativa

El propósito de transformar los datos originales mediante las variables indicador es para usar las variables indicador para estimar la función de probabilidad acumulativa $F_Z(z_c)$. Las funciones de probabilidad estimadas $\phi^*(A, z_c)$ se obtienen mediante combinaciones lineales de la función indicador. Esta función es la proporción exacta de valores, dentro de cualquier área A , de una variable $Z(\underline{x})$, menores que el valor de corte z_c . Si la medida de probabilidad es usada, el valor de $\phi^*(A, z_c)$ puede ser tomado como la probabilidad de que un parámetro estimado sea menor que el valor de corte.

La forma del estimador de $\phi(A, z_c)$ está dado por:

$$\phi^*(A, z_c) = \sum_{\alpha=1}^n \lambda_{\alpha}(z_c) i(x_{\alpha}, z_c) \quad (4.50)$$

con la condición para el no sesgo,

$$\sum_{\alpha=1}^n \lambda_{\alpha}(z_c) = 1 \quad (4.51)$$

donde $\lambda_{\alpha}(z_c)$ son los pesos.

El kriging simple puede ser aplicado para hallar los pesos λ_α usando las variables indicador $i(\underline{x}, z_c)$ y el variograma indicador.

La varianza de la estimación del Kriging indicador es:

$$\sigma_{KI}^2(A, z_c) = \sum_{\alpha=1}^n \lambda_\alpha(z_c) \gamma_I(\underline{x}_\alpha, A, z_c) - \bar{\gamma}_I(A, A, z_c) \quad (4.52)$$

donde

$$\gamma_I(\underline{x}_\alpha, A, z_c) = \frac{1}{A} \int_A \gamma_I(\underline{x}_\alpha - \underline{x}, z_c) d\underline{x} \quad (4.53)$$

y

$$\bar{\gamma}_I(A, A, z_c) = \frac{1}{A^2} \int_A \int_A \gamma_I(\underline{x} - \underline{y}, z_c) d\underline{x} d\underline{y} \quad (4.54)$$

es el valor medio de $\gamma_I(\underline{x}_j, A, z_c)$ cuando \underline{x} recorre todo el dominio A .

Para cada región, $\phi^*(A, z_c)$ es una función del valor de corte z_c . Si es aplicada una serie de valores de corte, será obtenida una serie de estimados. Con el incremento del valor de corte, $\phi^*(A, z_c)$ se incrementa, debido a que el porcentaje de los puntos con valores menores que el valor de corte se incrementa. Debido a que $\phi^*(A, z_c)$ es una función de probabilidad acumulativa, las siguientes relaciones de orden deben cumplirse:

$$a) \phi^*(A, -\infty) = 0, \quad y \quad \phi^*(A, +\infty) = 1,$$

$$b) 0 \leq \phi^*(A, z_c) \leq 1 \quad \forall A, z_c$$

$$c) \phi^*(A, z_c) \text{ no es decreciente, es decir}$$

$$\phi^*(A, z_c) \leq \phi^*(A, z'_c) \quad \text{si } z_c \leq z'_c$$

Si ocurre alguno de los siguientes casos :

- a) $\phi^*(A, z_c)$ estimada por KI es decreciente.
- b) $\phi^*(A, z_c)$ tiene valores negativos o mayores que 1,

las relaciones de orden no se cumplen. En resumen, si una distribución estimada tiene problemas con las relaciones de orden, no es una función de distribución válida.

Las ecuaciones del Kriging indicador se resuelven en la práctica para un número finito N , usualmente menor que 10, de valores de corte z_c . Si la estimación de cada valor de $\phi^*(A, z_c)$ se realiza por separado no se puede garantizar que se cumplan las relaciones de orden arriba expuestas.

Como alternativa (Myers, 1984) se plantea:

- a) Resolver el sistema de ecuaciones para $z_c = z_M$ y luego utilizar los coeficientes $\lambda_\alpha(z_c)$ para otros valores de z_c .
- b) Utilizar Cokriging indicador, es decir el estimado de $\phi^*(A, z_c)$ se obtiene a partir de los indicadores $i(x_\alpha, z_\beta)$, $\alpha = 1, \dots, n$ y $\beta = 1, \dots, N$. Esta alternativa es más exacta pero a la vez es más costosa en cuanto a cómputo.

4.10 Kriging logarítmico

El Kriging se considera óptimo cuando $Z(\underline{x})$ tiene una distribución normal. Cuando no se cumple esta condición se requiere realizar una transformación sobre $Z(\underline{x})$ de manera tal que siga una distribución normal. Por fortuna los estimadores asociados a la distribución normal son bastante robustos, es decir que su propiedad de optimalidad no se ve fuertemente afectada en el caso en que la distribución se aleje ligeramente de la normal.

La metodología a seguir puede ser resumida en pocos pasos. Sea $Y = f(Z)$ la variable transformada, entonces la aplicación del Kriging consiste en:

- 1.- Transformar todos los datos: $Y(\underline{x}_i) = f(Z(\underline{x}_i))$, $i=1, \dots, n$.
- 2.- Estimar y modelar el semivariograma de $Y(\underline{x})$.
- 3.- Aplicar las ecuaciones del Kriging a la variable Y para obtener Y^* y $Var[Y - Y^*]$.

El cálculo de Z^* y de $Var[Z - Z^*]$ a partir de los resultados de Y puede no ser trivial.

Lo que debe tenerse bien presente es que $Z^* \neq f^{-1}(Y^*)$

El ejemplo más común es cuando la distribución es lognormal. Entonces aplicando una transformación logarítmica $Y = \log(Z)$ será normal.

Al aplicar la metodología habitual del Kriging a Y y luego si intentamos emplear la teoría en sentido estricto transformando hacia atrás para obtener el valor estimado para Z^* , sería:

$$\tilde{Z} = \exp\left(Y^* + \frac{1}{2}\sigma_{KY}^2\right) \quad (4.55)$$

y su varianza

$$Var[\tilde{Z}] = m_Z^2 \exp(\sigma_Y^2) [1 - \exp(-\sigma_{KY}^2)] \quad (4.56)$$

Existen las siguientes dificultades según Journel y Huijbregts (1978). En la práctica \tilde{Z} no siempre satisface la condición de sesgo nulo, ya que puede diferir fuertemente de la media obtenida a partir de los valores observados de Z . Para evitar esta dificultad se propone el estimador

$$\tilde{Z} = K \exp\left(Y^* + \frac{1}{2}\sigma_{KY}^2\right) \quad (4.57)$$

donde K se determina imponiendo que la media aritmética de los valores estimados de \tilde{Z} coincida con el valor esperado m_Z .

CAPITULO 5: GEOSTADÍSTICA MULTIVARIADA

5.1 Introducción.

La estimación conjunta de variables aleatorias regionalizadas, más comúnmente conocida como Cokriging (Kriging Conjunto) es el análogo natural del Kriging de una función aleatoria. Mientras que el Kriging utiliza la correlación espacial para determinar los coeficientes en el estimador lineal, el Cokriging utiliza la correlación espacial y la correlación entre funciones aleatorias al mismo tiempo.

Las aplicaciones que han recibido una mayor atención en la geoestadística de minas son los casos donde dos o más variables están muestreadas, pero una está menos muestreada que las otras o existe la presencia de errores de muestreo.

Existe un número de dificultades prácticas, la más importante de todas es la ausencia de modelos estándar para las covarianzas cruzadas o covariogramas. Uno de los modelos más simples, el modelo lineal "estricto", no produce una varianza de estimación menor que el Kriging separado excepto en los dos casos mencionados anteriormente.

5.2 Definiciones de los momentos cruzados de segundo orden

En el análisis de varias variables correlacionadas entre sí existen varios momentos de segundo orden que miden el grado de correlación de las mismas.

Bajo la hipótesis de estacionaridad de segundo orden, se puede definir para cada par de variables $Z_i(\underline{x})$ y $Z_j(\underline{x})$ la covarianza cruzada como:

$$C_{ij}(\underline{h}) = E \left\{ [Z_i(\underline{x} + \underline{h}) - m_i] [Z_j(\underline{x}) - m_j] \right\} \quad (5.1)$$

y correspondientemente el semivariograma cruzado se define como

$$\gamma_{ij}(\underline{h}) = \frac{1}{2} E \left\{ [Z_i(\underline{x} + \underline{h}) - Z_i(\underline{x})] [Z_j(\underline{x} + \underline{h}) - Z_j(\underline{x})] \right\} \quad (5.2)$$

donde $m_i = E[Z_i(\underline{x})]$ y $m_j = E[Z_j(\underline{x})]$ son los valores esperados de las variables $Z_i(\underline{x})$ y $Z_j(\underline{x})$ respectivamente.

En el caso particular cuando $i = j$, los momentos cruzados se convierten en la covarianza y en la semivarianza de la variable $Z_i(\underline{x})$.

El semivariograma cruzado $\gamma_{ij}(\underline{h})$, a diferencia del semivariograma de una variable que siempre es positivo, puede tomar valores negativos. Los valores negativos se asocian a la situación en que el incremento de una de las variables implica en promedio el decremento de la otra

5.3 Propiedades de los momentos cruzados de segundo orden

De su definición se infiere que el semivariograma cruzado $\gamma_{ij}(\underline{h})$ es simétrico con respecto al intervalo o vector de separación \underline{h} y con respecto a los índices (i, j) de las variables, de modo que se cumple que:

$$\gamma_{ij}(\underline{h}) = \gamma_{ij}(-\underline{h}) \quad (5.3)$$

y

$$\gamma_{ij}(\underline{h}) = \gamma_{ji}(\underline{h}) \quad (5.4)$$

Sin embargo, la covarianza cruzada $C_{ij}(\underline{h})$ no es simétrica en general. Se cumple que:

$$C_{ij}(\underline{h}) = C_{ji}(-\underline{h}) \quad (5.5)$$

En el caso en que la covarianza cruzada $C_{ij}(\underline{h})$ sea simétrica, entonces resulta ésta equivalente al semivariograma cruzado.

La expresión general que relaciona al semivariograma cruzado con la covarianza cruzada se escribe como:

$$2\gamma_{ij}(\underline{h}) = 2C_{ij}(\underline{0}) - C_{ij}(\underline{h}) - C_{ij}(-\underline{h}) \quad (5.6)$$

Por lo que si consideramos el caso simétrico esta expresión se reduce a:

$$\gamma_{ij}(\underline{h}) = C_{ij}(\underline{0}) - C_{ij}(\underline{h}) \quad (5.7)$$

que es análoga a la existente entre semivariogramas y covarianzas simples.

El coeficiente de correlación entre las variables $Z_i(\underline{x})$ y $Z_j(\underline{x})$ en un mismo punto \underline{x} se define como:

$$\rho_{ij} = \frac{C_{ij}(\underline{0})}{\sqrt{C_{ii}(\underline{0})C_{jj}(\underline{0})}} \quad (5.8)$$

5.4 Cokriging en el caso estacionario.

Sean $Z_1(\underline{x}), Z_2(\underline{x}), \dots, Z_n(\underline{x})$ funciones aleatorias tales que:

$$E[Z_i(\underline{x})] = m_i, \quad i = 1, \dots, n \quad (5.9)$$

$$C_{ij}(\underline{x} - \underline{y}) = E[Z_i(\underline{x})Z_j(\underline{y})], \quad i, j = 1, \dots, n \quad (5.10)$$

existe para todo i, j y la covarianza de $Z_i(\underline{x})$ y $Z_j(\underline{x})$ depende solamente de la diferencia $\underline{x} - \underline{y}$.

Esto quiere decir que $\{Z_i\}_{i=1}^n$ son funciones estacionarias de segundo orden conjuntamente.

Dados los puntos muestreados $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_m$ y un punto no muestreado \underline{x} , el objetivo es estimar conjuntamente cada variable aleatoria $Z_i(\underline{x})$ mediante combinaciones lineales de observaciones de todas las variables aleatorias $Z_j(\underline{x}_k)$, $k = 1, \dots, m$ $j = 1, \dots, n$

$$Z_i^*(\underline{x}) = \sum_{k=1}^m \sum_{j=1}^n \lambda_{ij}^k Z_j(\underline{x}_k) \quad (5.11)$$

En forma matricial, si $\underline{Z}(\underline{x}) = [Z_1(\underline{x}), Z_2(\underline{x}), \dots, Z_n(\underline{x})]^T$ entonces la expresión anterior se convierte en:

$$\underline{Z}^*(\underline{x}) = \sum_{k=1}^m \underline{\Gamma}_k \underline{Z}(\underline{x}_k); \quad \text{donde} \quad \underline{\Gamma}_k = [\lambda_{ij}^k] \quad (5.12)$$

Para que $\underline{Z}^*(\underline{x})$ sea un estimador no sesgado de $\underline{Z}(\underline{x})$ se debe cumplir que:

$$\sum_{k=1}^m \underline{\Gamma}_k = \underline{I} \quad (5.13)$$

donde \underline{I} es la matriz identidad.

$$\sum_{k=1}^m \lambda_{ij}^k = \delta_{ij} = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases} \quad (5.14)$$

Se toma que la varianza de la estimación sea:

$$\sum_{j=1}^n \text{Var}[Z_j^*(\underline{x}) - Z_j(\underline{x})] = E\left\{\left[\underline{Z}^*(\underline{x}) - \underline{Z}(\underline{x})\right]^T \left[\underline{Z}^*(\underline{x}) - \underline{Z}(\underline{x})\right]\right\} \quad (5.15)$$

El sistema puede ser escrito en la forma:

$$\begin{cases} \sum_{j=1}^m \Gamma_{\underline{j}} \underline{C}(\underline{x}_i - \underline{x}_j) + \underline{\mu} = \underline{C}(\underline{x}_i - \underline{x}) \\ \sum_{k=1}^m \Gamma_{\underline{k}} = \underline{I}, \quad i = 1, \dots, m \end{cases} \quad (5.16)$$

donde $\underline{\mu} = [\mu_{ij}]$ - es una matriz de $n \times n$ multiplicadores de Lagrange y

$$\underline{C}(\underline{x}_i - \underline{x}_j) = \begin{bmatrix} C_{11}(\underline{x}_i - \underline{x}_j) & C_{12}(\underline{x}_i - \underline{x}_j) & \dots & C_{1n}(\underline{x}_i - \underline{x}_j) \\ C_{21}(\underline{x}_i - \underline{x}_j) & C_{22}(\underline{x}_i - \underline{x}_j) & \dots & C_{2n}(\underline{x}_i - \underline{x}_j) \\ \dots & \dots & \dots & \dots \\ C_{n1}(\underline{x}_i - \underline{x}_j) & C_{n2}(\underline{x}_i - \underline{x}_j) & \dots & C_{nn}(\underline{x}_i - \underline{x}_j) \end{bmatrix} \quad (5.17)$$

Entonces si escribimos el sistema del Cokriging en forma matricial tenemos que:

$$\underline{K} \underline{\Lambda} = \underline{D} \quad (5.18)$$

Donde

$$\underline{K} = \begin{bmatrix} \underline{C}(\underline{x}_1 - \underline{x}_1) & \dots & \underline{C}(\underline{x}_1 - \underline{x}_m) & \underline{I} \\ \dots & \dots & \dots & \dots \\ \underline{C}(\underline{x}_m - \underline{x}_1) & \dots & \underline{C}(\underline{x}_m - \underline{x}_m) & \underline{I} \\ \underline{I} & \dots & \underline{I} & \underline{0} \end{bmatrix} \quad (5.19)$$

$$\underline{\Lambda} = \begin{bmatrix} \underline{\Gamma}_1 \\ \dots \\ \underline{\Gamma}_m \\ \underline{\mu} \end{bmatrix}, \quad \underline{D} = \begin{bmatrix} \underline{C}(\underline{x}_1 - \underline{x}) \\ \dots \\ \underline{C}(\underline{x}_m - \underline{x}) \\ \underline{I} \end{bmatrix} \quad (5.20)$$

y la varianza total de la estimación puede ser escrita como:

$$\sigma_{CK}^2 = \text{Tr} \underline{C}(\underline{0}) - \text{Tr} \sum_{j=1}^m \Gamma_{\underline{j}} \underline{C}(\underline{x} - \underline{x}_j) - \text{Tr} \underline{\mu} \quad (5.21)$$

la cual representa una varianza acumulada, mientras que la varianza de la estimación de la variable $Z_j(\underline{x})$ está dada por

$$\sigma_{CKj}^2 = C_{jj}(\underline{0}) - \sum_{i=1}^n \sum_{k=1}^m \lambda_{ij}^k C_{ij}(\underline{x} - \underline{x}_k) - \mu_{jj} \quad (5.22)$$

Observaciones prácticas:

- A pesar de la asimetría de $\underline{\underline{C}}(\underline{x}-\underline{y})$, la matriz de coeficientes $\underline{\underline{K}}$ es simétrica.
- Todas las entradas son invertibles (excepto 0) de forma tal que puede ser reducida a una matriz triangular mediante la operación con matrices de menor dimensión y así simplificar el cómputo.
- Si $\underline{\underline{C}}(\underline{x}-\underline{y})$ tiene forma diagonal, es decir, las componentes no están correlacionadas entonces el sistema se convierte en n sistemas de ecuaciones independientes separadas.

5.5 Cokriging en el caso de funciones aleatorias intrínsecas.

Como con el Kriging de una variable es más simple usar covariogramas que covarianzas cruzadas, pero esto no siempre es posible; la hipótesis intrínseca análoga es:

$$i) E[Z_i(\underline{x}+\underline{h})-Z_i(\underline{x})]=0, \quad \forall i=1,\dots,n$$

$$ii) Cov[Z_i(\underline{x}+\underline{h})-Z_i(\underline{x}), Z_j(\underline{x}+\underline{h})-Z_j(\underline{x})]=2\gamma_{ij}(\underline{h}), \quad \forall i,j=1,\dots,n$$

existe y depende sólo de \underline{h} .

Si $i)$ se cumple entonces la matriz de semivarianzas cruzadas puede ser escrita como:

$$\underline{\underline{\gamma}}(\underline{h})=[\gamma_{ij}(\underline{h})]=\frac{1}{2}E\left\{[\underline{Z}(\underline{x}+\underline{h})-\underline{Z}(\underline{x})][\underline{Z}(\underline{x}+\underline{h})-\underline{Z}(\underline{x})]^T\right\} \quad (5.23)$$

donde $\underline{\underline{\gamma}}(\underline{h})$ es simétrica en contraste con $\underline{\underline{C}}(\underline{x}-\underline{y})$.

Entonces al sustituir en el sistema de ecuaciones del Kriging a las funciones de covarianzas cruzadas por los co-semivariogramas o variogramas cruzados, se obtiene

$$\begin{bmatrix} \underline{\underline{\gamma}}(\underline{x}_1-\underline{x}_1) & \dots & \underline{\underline{\gamma}}(\underline{x}_1-\underline{x}_m) & \underline{\underline{I}} \\ \dots & \dots & \dots & \dots \\ \underline{\underline{\gamma}}(\underline{x}_m-\underline{x}_1) & \dots & \underline{\underline{\gamma}}(\underline{x}_m-\underline{x}_m) & \underline{\underline{I}} \\ \underline{\underline{I}} & \dots & \underline{\underline{I}} & \underline{\underline{0}} \end{bmatrix} \begin{bmatrix} \underline{\underline{\Gamma}}_1 \\ \dots \\ \underline{\underline{\Gamma}}_m \\ \underline{\underline{\mu}} \end{bmatrix} = \begin{bmatrix} \underline{\underline{\gamma}}(\underline{x}-\underline{x}_1) \\ \dots \\ \underline{\underline{\gamma}}(\underline{x}-\underline{x}_m) \\ \underline{\underline{I}} \end{bmatrix} \quad (5.24)$$

La varianza de la estimación se expresa como:

$$\sigma_{CK}^2 = \text{Tr} \begin{bmatrix} \underline{\Gamma}_1 & \dots & \underline{\Gamma}_m & \underline{\mu} \end{bmatrix} \begin{bmatrix} \underline{\gamma}(\underline{x} - \underline{x}_1) \\ \dots \\ \underline{\gamma}(\underline{x} - \underline{x}_m) \\ \underline{I} \end{bmatrix} \quad (5.25)$$

5.6 Cokriging en el caso no estacionario.

Si se supone que las componentes de $\underline{Z}(\underline{x})$ no son estacionarias, es decir, los valores esperados $\underline{m}(\underline{x})$ están representados localmente mediante combinaciones lineales de funciones conocidas, entonces los resultados de la sección anterior pueden ser extendidos de forma similar al Kriging Universal para una variable.

Sea $\underline{F}(\underline{x}) = [f_1(\underline{x}), \dots, f_p(\underline{x})]$ un vector de funciones conocidas, linealmente independientes.

Entonces se asume que:

$$E[\underline{Z}(\underline{x})] = \underline{B}\underline{F}(\underline{x}) \quad (5.26)$$

donde $\underline{B} = [b_{ij}]$, una matriz de $n \times p$.

Para que $\underline{Z}(\underline{x})$ sea insesgado es suficiente que:

$$\underline{B}\underline{F}(\underline{x}) = \sum_{j=1}^m \underline{\Gamma}_j \underline{B}\underline{F}(\underline{x}_j) \quad (5.27)$$

para cualquier elección de \underline{B} .

Considerando todas las matrices de la forma:

$$\underline{B} = [b_{ij}]; \quad \text{donde } b_{ij} = \begin{cases} 1, & i = i_0, j = j_0 \\ 0, & \text{en otro caso} \end{cases}$$

y escribiendo $\underline{F}_l(\underline{x}_j) = f_l(\underline{x}_j)\underline{I}$ entonces la expresión anterior se convierte

$$\sum_{j=1}^m \underline{\Gamma}_j \underline{F}_l(\underline{x}_j) = \underline{F}_l(\underline{x}); \quad l = 1, \dots, p \quad (5.28)$$

El sistema de ecuaciones para el Kriging Universal resulta

$$\begin{cases} \sum_{j=1}^m \underline{\Gamma}_j \underline{C}(\underline{x}_i - \underline{x}_j) + \sum_{l=1}^p \underline{\mu}_l \underline{F}_l(\underline{x}) = \underline{C}(\underline{x}_i - \underline{x}) \\ \sum_{j=1}^m \underline{\Gamma}_j \underline{F}_l(\underline{x}_j) = \underline{F}_l(\underline{x}); \quad l=1, \dots, p \quad i=1, \dots, m \end{cases} \quad (5.29)$$

donde $\underline{\mu}_l$ son matrices $n \times n$ de multiplicadores de Lagrange.

En forma matricial

$$\underline{W} = \begin{bmatrix} \underline{C}(\underline{x}_1 - \underline{x}_1) & \dots & \underline{C}(\underline{x}_1 - \underline{x}_n) & \underline{F}_1(\underline{x}_1) & \dots & \underline{F}_p(\underline{x}_1) \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \underline{C}(\underline{x}_n - \underline{x}_1) & \dots & \underline{C}(\underline{x}_n - \underline{x}_n) & \underline{F}_1(\underline{x}_n) & \dots & \underline{F}_p(\underline{x}_n) \\ \underline{F}_1(\underline{x}_1) & \dots & \underline{F}_1(\underline{x}_n) & \underline{0} & \dots & \underline{0} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \underline{F}_p(\underline{x}_1) & \dots & \underline{F}_p(\underline{x}_n) & \underline{0} & \dots & \underline{0} \end{bmatrix} \quad (5.30)$$

$$\underline{X} = \begin{bmatrix} \underline{\Gamma}_1 \\ \dots \\ \underline{\Gamma}_n \\ \underline{\mu}_1 \\ \dots \\ \underline{\mu}_p \end{bmatrix}, \quad \underline{H} = \begin{bmatrix} \underline{C}(\underline{x}_1 - \underline{x}) \\ \dots \\ \underline{C}(\underline{x}_n - \underline{x}) \\ \underline{F}_1(\underline{x}) \\ \dots \\ \underline{F}_p(\underline{x}) \end{bmatrix} \quad (5.31)$$

$$\underline{W} \underline{X} = \underline{H} \quad (5.32)$$

La varianza de la estimación puede ser escrita como:

$$\sigma_{CKU}^2 = \text{Tr } \underline{C}(\underline{0}) - \text{Tr } \sum_{i=1}^m \underline{\Gamma}_i \underline{C}(\underline{x} - \underline{x}_i) - \text{Tr } \sum_{l=1}^p \underline{F}_l(\underline{x}) \underline{\mu}_l \quad (5.33)$$

5.7 Estimación de los momentos cruzados de segundo orden

El método más usual para estimar el semivariograma cruzado es el siguiente:

$$\gamma_{ij}^*(\underline{h}) = \frac{1}{2N(\underline{h})} \sum_{k=1}^{N(\underline{h})} [Z_i(\underline{x}_k + \underline{h}) - Z_i(\underline{x}_k)] [Z_j(\underline{x}_k + \underline{h}) - Z_j(\underline{x}_k)] \quad (5.34)$$

donde $N(\underline{h})$ es el número de pares $\{Z_i(\underline{x}_k), Z_i(\underline{x}_k + \underline{h})\}$ y $\{Z_j(\underline{x}_k), Z_j(\underline{x}_k + \underline{h})\}$ separados

a una distancia $h = |\underline{h}|$.

El cual es una generalización del estimador del semivariograma simple y por lo tanto adolece de los mismos problemas y limitaciones.

5.8 Modelación de los momentos cruzados de segundo orden

Es una práctica muy común ajustar modelos desarrollados para una variable como son: esférico, exponencial, etc, al semivariograma cruzado, aunque en este caso no se pueda asegurar que estos modelos sean válidos.

Myers (1984), propuso una metodología para el ajuste de los semivariogramas cruzados, la cual consiste en definir para cada par de variables $Z_i(\underline{x})$ y $Z_j(\underline{x})$ una nueva variable como una combinación lineal de éstas, es decir:

$$Z_{ij}^+(\underline{x}) = aZ_i(\underline{x}) + bZ_j(\underline{x}) \quad (5.35)$$

cuyo semivariograma viene dado por:

$$\begin{aligned} \gamma_{ij}^+(\underline{h}) &= \frac{1}{2} E \left\{ \left[aZ_i(\underline{x} + \underline{h}) + bZ_j(\underline{x} + \underline{h}) - aZ_i(\underline{x}) - bZ_j(\underline{x}) \right]^2 \right\} \\ &= a^2\gamma_{ii}(\underline{h}) + b^2\gamma_{jj}(\underline{h}) + 2ab\gamma_{ij}(\underline{h}) \end{aligned} \quad (5.36)$$

entonces

$$\gamma_{ij}(\underline{h}) = \frac{1}{2ab} \left[\gamma_{ij}^+(\underline{h}) - a^2\gamma_{ii}(\underline{h}) - b^2\gamma_{jj}(\underline{h}) \right] \quad (5.37)$$

De manera análoga se puede obtener el semivariograma de la diferencia

$$Z_{ij}^-(\underline{x}) = aZ_i(\underline{x}) - bZ_j(\underline{x}) \quad (5.38)$$

su semivariograma está dado por:

$$\begin{aligned} \gamma_{ij}^-(\underline{h}) &= \frac{1}{2} E \left\{ \left[aZ_i(\underline{x} + \underline{h}) - bZ_j(\underline{x} + \underline{h}) - aZ_i(\underline{x}) + bZ_j(\underline{x}) \right]^2 \right\} \\ &= a^2\gamma_{ii}(\underline{h}) + b^2\gamma_{jj}(\underline{h}) - 2ab\gamma_{ij}(\underline{h}) \end{aligned} \quad (5.39)$$

y por lo tanto

$$\gamma_{ij}(\underline{h}) = -\frac{1}{2ab} \left[\gamma_{ij}^-(\underline{h}) - a^2\gamma_{ii}(\underline{h}) - b^2\gamma_{jj}(\underline{h}) \right] \quad (5.40)$$

Combinando ambos resultados se obtiene una tercera expresión

$$\gamma_{ij}(\underline{h}) = \frac{1}{4ab} \left[\gamma_{ij}^+(\underline{h}) - \gamma_{ij}^-(\underline{h}) \right] \quad (5.41)$$

Entonces el procedimiento más adecuado para modelar el semivariograma cruzado consiste en modelar los 4 semivariogramas simples: $\gamma_{ii}(\underline{h})$, $\gamma_{jj}(\underline{h})$, $\gamma_{ij}^+(\underline{h})$ y $\gamma_{ij}^-(\underline{h})$, comprobar que satisfagan las tres expresiones anteriores y además que se cumpla la desigualdad de Cauchy-Schwartz:

$$\left| \gamma_{ij}^+(\underline{h}) \right|^2 \leq \gamma_{ii}(\underline{h}) \gamma_{jj}(\underline{h}) \quad (5.42)$$

donde

$\gamma_{ij}^+(\underline{h})$ - es el semivariograma de la suma de Z_i y Z_j .

$\gamma_{ii}(\underline{h})$ y $\gamma_{jj}(\underline{h})$ - son los semivariogramas de Z_i y Z_j respectivamente.

En resumen, para modelar el variograma cruzado hay que calcular los variogramas muestrales para Z_i , Z_j , Z_i+Z_j y Z_i-Z_j . Para cada miembro de la cuarteta se selecciona uno de los modelos estándar tales como: esférico, exponencial, gaussiano, etc., entonces el modelo para $\gamma_{ij}(\underline{h})$ se determina como una combinación de éstos usando la expresión anterior, pero con la condición de que satisfaga la desigualdad de Cauchy-Schwartz. En el caso en que no se cumpla esta condición última se deben realizar modificaciones a cada uno de los modelos ajustados a $\gamma_{ii}(\underline{h})$, $\gamma_{jj}(\underline{h})$, $\gamma_{ij}^+(\underline{h})$ y $\gamma_{ij}^-(\underline{h})$, de manera tal que ésta sea satisfecha.

Para modelar el variograma cruzado de n variables aleatorias regionalizadas, se deben realizar solamente $(n^2 + n)/2$ ajustes de variogramas simples.

Pero en realidad este procedimiento no garantiza que se obtenga un modelo válido de semivariogramas cruzados, ya que la desigualdad de Cauchy-Schwartz es una condición necesaria pero no suficiente.

La condición que se requiere para que el modelo sea válido es la análoga al caso univariado, es decir

$$\sum_i \sum_j \Gamma_{ii}^T \underline{C}(\underline{h}) \Gamma_{jj} \geq 0 \quad (5.43)$$

que la matriz $\underline{C}(\underline{h})$ sea positiva semidefinida o equivalentemente en términos de las semivarianzas

$$-\sum_i \sum_j \Gamma_{ii}^T \gamma(\underline{h}) \Gamma_{jj} \geq 0, \quad \text{si} \quad \sum_i \Gamma_{ii} = 0 \quad (5.44)$$

que significa que la matriz $\gamma(\underline{h})$ sea condicionalmente negativa semidefinida.

Estas condiciones no son sencillas de comprobar y en el caso univariado se resuelve usando combinaciones lineales de modelos autorizados, es decir:

$$\sum_i a_i \gamma_i(\underline{h}) \quad (5.45)$$

es un modelo válido, donde a_1, \dots, a_m son coeficientes positivos y $\gamma_1(\underline{h}), \dots, \gamma_m(\underline{h})$ son modelos simples de variogramas válidos.

Por lo que la alternativa análoga al caso univariado es:

$$\sum_i \underline{A}_i \gamma_i(\underline{h}) \quad (5.46)$$

es un modelo válido, donde $\underline{A}_1, \dots, \underline{A}_m$ son matrices de coeficientes positivas definidas y $\gamma_1(\underline{h}), \dots, \gamma_m(\underline{h})$ son modelos simples de variogramas válidos. Esto se conoce como modelo de correogionalización lineal.

5.9 Modelo de Correogionalización Lineal

Un modelo comúnmente usado para un conjunto de funciones aleatorias es el modelo multivariado de funciones de covarianzas anidadas

$$\underline{C}(\underline{h}) = \sum_{k=0}^S \underline{V}_{\underline{k}} \rho_k(\underline{h}) \quad (5.47)$$

donde k es el índice de $S+1$ estructuras anidadas a diferentes escalas con función de correlación $\rho_k(\underline{h})$ y las matrices de correogionalización $\underline{V}_{\underline{k}}$ son las matrices de covarianzas que describen la correlación multivariada a la escala k .

El modelo de funciones aleatorias asociado es conocido como modelo de correogionalización lineal (MCL) y se expresa como:

$$\underline{Z}(\underline{x}) = \sum_{k=0}^S \underline{A}_{\underline{k}} \underline{Y}_{\underline{k}}(\underline{x}) \quad (5.48)$$

$$Z_i(\underline{x}) = \sum_{k=0}^S \sum_{j=1}^p a_{ij}^k Y_{jk}(\underline{x}); \quad i = 1, \dots, n \quad (5.49)$$

donde $\underline{Y}_{\underline{k}}(\underline{x}) = [Y_{1k}(\underline{x}), Y_{2k}(\underline{x}), \dots, Y_{pk}(\underline{x})]^T$ es el vector de p funciones aleatorias estacionarias de segundo orden no correlacionadas (independientes entre sí) definidas en la escala espacial k .

Una vía para determinar los coeficientes del modelo de correogionalización lineal (MCL), a partir de un modelo multivariado de funciones de covarianzas anidadas es mediante la aplicación del método de componentes principales basado en la descomposición en valores propios de las matrices de correogionalización $\underline{V}_{\underline{k}}$.

5.10 Análisis Estructural Multivariado

El análisis estructural multivariado que se requiere para el Cokriging es mucho más complejo y sofisticado que el que demanda el Kriging ya que para modelar los variogramas cruzados de n variables aleatorias regionalizadas, se deben modelar (ajustar) un total de $n(n+1)/2$ variogramas simples. Mientras que el uso de modelos de variogramas autorizados o combinaciones de éstos no garantiza, es decir no es una condición suficiente, como lo es en el caso univariado para que la matriz de covarianzas $\underline{\underline{C}}(\underline{h})$ sea positiva definida.

La manera más comúnmente aceptada en la actualidad para realizar un análisis estructural multivariado es a través de un modelo de correogionalización lineal (Goovaerts, 1997). No obstante existen otras metodologías menos difundidas que usan métodos espectrales y están basadas en el teorema de Bochner (Christakos, 1992; Wackernagel, 1995).

El punto medular para poder establecer un modelo de correogionalización lineal consiste en probar que las matrices de coeficientes $\underline{\underline{V}}_k$ son positivas semidefinidas.

Por definición, una matriz $\underline{\underline{V}}_k$ es positiva semidefinida (Golub y Van Loan, 1989) si

$$\underline{b}^T \underline{\underline{V}}_k \underline{b} \geq 0, \quad \forall \underline{b} \quad (5.50)$$

donde \underline{b} es un vector cualquiera. Cuando una matriz es positiva semidefinida sus valores propios y los determinantes de ella y de todos sus menores principales son no negativos.

Un modelo de correogionalización lineal está dado por

$$\underline{\underline{C}}(\underline{h}) = \sum_{k=0}^S \underline{\underline{V}}_k \rho_k(\underline{h}) \Leftrightarrow C_{ij}(\underline{h}) = \sum_{k=0}^S \sigma_{ij}^k \rho_k(\underline{h}) \quad (5.51)$$

en términos de las covarianzas o equivalentemente

$$\underline{\underline{\gamma}}(\underline{h}) = \sum_{k=0}^S \underline{\underline{V}}_k \gamma_k(\underline{h}) \Leftrightarrow \gamma_{ij}(\underline{h}) = \sum_{k=0}^S \sigma_{ij}^k \gamma_k(\underline{h}) \quad (5.52)$$

en términos de las semivarianzas. Y se interpreta como $S+1$ estructuras anidadas a diferentes escalas y las matrices de correogionalización $\underline{\underline{V}}_k$ son las matrices de covarianzas que describen la correlación multivariada a la escala k . Note que a cada escala k le corresponde una estructura elemental o básica con mesetas igual a la unidad ($\rho_k(\underline{h})$ o $\gamma_k(\underline{h})$). Si determinada estructura básica no está presente, se le hace corresponder un coeficiente cero en la matriz $\underline{\underline{V}}_k$.

El nombre de modelo lineal de correogionalización se origina del hecho de que las matrices de covarianzas o semivarianzas de las Ecs. (5.51) y (5.52) se obtienen considerando la transformación lineal de funciones aleatorias Ec.(5.48).

Una condición suficiente para que el modelo de correogionalización lineal sea válido consiste en que todas las matrices de correogionalización $V_{\underline{k}}$ sean positivas semidefinidas.

Si $V_{\underline{k}}$ es una matriz simétrica y positiva semidefinida, entonces se cumple que

$$|\sigma_{ij}^k| \leq \sqrt{\sigma_{ii}^k \sigma_{jj}^k}, \quad \forall i \neq j \quad (5.53)$$

Pero en general la proposición inversa es falsa.

No obstante en el caso de matrices de orden dos si se cumple, es decir es una condición necesaria y suficiente para que la matriz sea positiva semidefinida. Esto resulta muy útil, puesto que es muy común la modelación conjunta de sólo dos funciones aleatorias.

Un ejemplo de modelo de correogionalización lineal en el caso de dos funciones aleatorias $Z_1(\underline{x})$ y $Z_2(\underline{x})$, es

$$\begin{pmatrix} \gamma_{11}(\underline{h}) & \gamma_{12}(\underline{h}) \\ \gamma_{21}(\underline{h}) & \gamma_{22}(\underline{h}) \end{pmatrix} = \begin{pmatrix} \sigma_{11}^0 & \sigma_{12}^0 \\ \sigma_{21}^0 & \sigma_{22}^0 \end{pmatrix} \gamma_0(\underline{h}) + \dots + \begin{pmatrix} \sigma_{11}^S & \sigma_{12}^S \\ \sigma_{21}^S & \sigma_{22}^S \end{pmatrix} \gamma_S(\underline{h}) \quad (5.54)$$

Para asegurar de que el modelo sea válido es suficiente probar que

$$\begin{aligned} \sigma_{11}^k > 0 \text{ y } \sigma_{22}^k > 0, \quad \forall k = 0, \dots, S \\ |\sigma_{12}^k| \leq \sqrt{\sigma_{11}^k \sigma_{22}^k}, \quad \forall k = 0, \dots, S \end{aligned} \quad (5.55)$$

Procedimiento general para la obtención de un modelo de correogionalización lineal

El procedimiento general para ajustar un modelo de correogionalización lineal consiste en postular el número de estructuras y sus modelos elementales correspondientes para los cuales están definidos los rangos o alcances, y luego intentar el ajuste de las mesetas mediante prueba o error, o aplicando algún método de optimización.

El esquema general usando el método de prueba y error es el siguiente:

1. Modelar cada semivariograma simple $\gamma_{ii}(\underline{h})$, $i = 1, \dots, n$ y semivariograma cruzado $\gamma_{ij}(\underline{h})$, $i \neq j$, y $i, j = 1, \dots, n$ individualmente según el procedimiento expuesto en el capítulo 3 del análisis estructural de una función aleatoria.
2. Determinar el número de estructuras anidadas $S+1$ de manera que sea mínimo (es deseable que sea cuanto más tres), según las consideraciones siguientes:
 - a) Si $\sigma_{ij}^k > 0$ entonces $\sigma_{ii}^k > 0$ y $\sigma_{jj}^k > 0$. Es decir, si una estructura $\gamma_k(\underline{h})$ hace contribución al modelo anidado del variograma cruzado $\gamma_{ij}(\underline{h})$, entonces debe

- contribuir también en el modelo de los variogramas simples $\gamma_{ii}(\underline{h})$ y $\gamma_{jj}(\underline{h})$. Lo contrario es falso
- b) Si $\sigma_{ii}^k > 0$ y $\sigma_{jj}^k > 0$ no implica nada sobre σ_{ij}^k . Es decir, si una estructura $\gamma_k(\underline{h})$ hace contribución a los modelos anidados de los variogramas simples $\gamma_{ii}(\underline{h})$ y $\gamma_{jj}(\underline{h})$, dicha estructura puede contribuir o no en el modelo anidado del variograma cruzado $\gamma_{ij}(\underline{h})$.
 - c) Si $\sigma_{ii}^k = 0$, entonces $\sigma_{ij}^k = 0$, $\forall j = 1, \dots, n$. Es decir, si una estructura $\gamma_k(\underline{h})$ no contribuye en el modelo anidado del variograma simple $\gamma_{ii}(\underline{h})$, entonces dicha estructura no puede contribuir en ninguno de los modelos anidados de los variogramas cruzados $\gamma_{ij}(\underline{h})$ que involucran a la componente i .
3. Comprobar que todos los determinantes de los menores de orden dos son no negativos.
 4. Verificar que todas las matrices de correogionalización $V_{\underline{k}}$ sean positivas semidefinidas, en caso contrario hacer los cambios necesarios hasta satisfacer la condición o volver al paso 2.

Goulard (1989) y Goulard y Voltz (1992), propusieron un algoritmo iterativo para el ajuste de los coeficientes mediante mínimos cuadrados bajo restricciones de positividad.

5.11 Validación del modelo de semivariograma cruzado

El método convencional de validación del semivariograma cruzado $\gamma_{ij}(\underline{h})$ en el caso de dos variables consiste en estimar por Cokriging los valores de $Z_i(\underline{x})$ y $Z_j(\underline{x})$ en los puntos muestrales usando el procedimiento de *leave one out*. Con los valores estimados obtenidos $Z_i^*(\underline{x})$, $Z_j^*(\underline{x})$ y sus correspondientes varianzas de la estimación σ_i^2 , σ_j^2 se calculan los criterios convencionales de la validación cruzada para una variable (error medio, error cuadrático medio, etc,...)

El método anterior no es óptimo, ya que la validación de $\gamma_{ij}(\underline{h})$ debe ligarse con la validación de los semivariogramas $\gamma_{ii}(\underline{h}), \gamma_{jj}(\underline{h})$, es decir se debería validar primero por separado $\gamma_{ii}(\underline{h}), \gamma_{jj}(\underline{h})$ y luego $\gamma_{ij}(\underline{h})$ de manera conjunta.

5.12 Ejemplos de Aplicaciones del Cokriging

5.12.1 El caso de dos variables correlacionadas

Supongamos que tenemos dos variables que son estacionarias de segundo orden, entonces se cumple que:

$$E[Z_1(\underline{x})] = m_1, \quad E[Z_2(\underline{x})] = m_2 \quad (5.56)$$

$$C_{12}(\underline{x} - \underline{y}) = E[Z_1(\underline{x}) Z_2(\underline{y})] \quad (5.57)$$

El estimador sería de la siguiente forma:

$$\underline{Z}^*(\underline{x}) = \sum_{k=1}^n \underline{\Gamma}_{\underline{k}} \underline{Z}(\underline{x}_k) \quad (5.58)$$

$$Z_i^*(\underline{x}) = \sum_{k=1}^n \sum_{j=1}^2 \lambda_{ij}^k Z_j(\underline{x}_k), \quad i = 1, 2 \quad (5.59)$$

donde $\underline{\Gamma}_{\underline{k}} = \begin{bmatrix} \lambda_{11}^k & \lambda_{12}^k \\ \lambda_{21}^k & \lambda_{22}^k \end{bmatrix}$, $\sum_{k=1}^n \underline{\Gamma}_{\underline{k}} = \underline{I}$, \underline{I} - matriz identidad.

$$\sum_{k=1}^n \lambda_{ij}^k = \delta_{ij} = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases} \quad (5.60)$$

$$\underline{\underline{C}}(\underline{x} - \underline{y}) = \begin{bmatrix} C_{11}(\underline{x} - \underline{y}) & C_{12}(\underline{x} - \underline{y}) \\ C_{21}(\underline{x} - \underline{y}) & C_{22}(\underline{x} - \underline{y}) \end{bmatrix} \quad (5.61)$$

$$\underline{\underline{K}} = \begin{bmatrix} \begin{bmatrix} C_{11}(\underline{x}_1 - \underline{x}_1) & C_{12}(\underline{x}_1 - \underline{x}_1) \\ C_{21}(\underline{x}_1 - \underline{x}_1) & C_{22}(\underline{x}_1 - \underline{x}_1) \end{bmatrix} & \cdots & \begin{bmatrix} C_{11}(\underline{x}_1 - \underline{x}_n) & C_{12}(\underline{x}_1 - \underline{x}_n) \\ C_{21}(\underline{x}_1 - \underline{x}_n) & C_{22}(\underline{x}_1 - \underline{x}_n) \end{bmatrix} & \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ \cdots & \cdots & \cdots & \cdots \\ \begin{bmatrix} C_{11}(\underline{x}_n - \underline{x}_1) & C_{12}(\underline{x}_n - \underline{x}_1) \\ C_{21}(\underline{x}_n - \underline{x}_1) & C_{22}(\underline{x}_n - \underline{x}_1) \end{bmatrix} & \cdots & \begin{bmatrix} C_{11}(\underline{x}_n - \underline{x}_n) & C_{12}(\underline{x}_n - \underline{x}_n) \\ C_{21}(\underline{x}_n - \underline{x}_n) & C_{22}(\underline{x}_n - \underline{x}_n) \end{bmatrix} & \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} & \cdots & \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} & \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \end{bmatrix} \quad (5.62)$$

$$\underline{\underline{\Lambda}} = \begin{bmatrix} \begin{bmatrix} \lambda_{11}^1 & \lambda_{12}^1 \\ \lambda_{21}^1 & \lambda_{22}^1 \end{bmatrix} & & \\ & \dots & \\ \begin{bmatrix} \lambda_{11}^n & \lambda_{12}^n \\ \lambda_{21}^n & \lambda_{22}^n \end{bmatrix} & & \\ & & \begin{bmatrix} \mu_{11} & \mu_{12} \\ \mu_{21} & \mu_{22} \end{bmatrix} \end{bmatrix}, \quad \underline{\underline{D}} = \begin{bmatrix} \begin{bmatrix} C_{11}(\underline{x}_1 - \underline{x}) & C_{12}(\underline{x}_1 - \underline{x}) \\ C_{21}(\underline{x}_1 - \underline{x}) & C_{22}(\underline{x}_1 - \underline{x}) \end{bmatrix} & & \\ & \dots & \\ \begin{bmatrix} C_{11}(\underline{x}_n - \underline{x}) & C_{12}(\underline{x}_n - \underline{x}) \\ C_{21}(\underline{x}_n - \underline{x}) & C_{22}(\underline{x}_n - \underline{x}) \end{bmatrix} & & \\ & & \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \end{bmatrix} \quad (5.63)$$

$$\underline{\underline{K}} \underline{\underline{\Lambda}} = \underline{\underline{D}} \quad (5.64)$$

$$\sigma_{CKj}^2 = C_{jj}(\underline{0}) - \sum_{i=1}^2 \sum_{k=1}^n C_{ij}(\underline{x} - \underline{x}_k) \lambda_{ij}^k - \mu_{jj} \quad (5.65)$$

5.12.2 Estimación de combinaciones lineales de variables corregeionalizadas

Existe un número de aplicaciones donde es deseado estimar una combinación lineal de variables corregeionalizadas. Por ejemplo si se necesita identificar anomalías entonces una combinación lineal de algunas variables puede ser relevante y de este modo las variables que son indicativas de la presencia de un elemento particular pueden ser usadas. Existen varios enfoques posibles. El más directo es tomar un conjunto de datos multivariados y formar una combinación lineal para obtener un nuevo conjunto de datos para la variable construida, entonces es calculado el variograma muestral, luego modelado y finalmente se le aplica el Kriging.

Otro enfoque consiste en estimar cada variable y luego construir la combinación lineal. Este puede ser llevado a cabo mediante la estimación de cada variable por separado o mediante la estimación conjunta por Cokriging. La diferencia de estos métodos se refleja en las varianzas de la estimación.

Al primer enfoque lo nombraremos como la estimación directa y al segundo como la estimación conjunta.

a) Estimación directa.

Dados: $\underline{Z}(\underline{x}) = [Z_1(\underline{x}) \quad \dots \quad Z_m(\underline{x})]^T$ y $\underline{A}^T = [a_1 \quad \dots \quad a_m]$, donde $a_i \in \mathbb{R}$

El objetivo es estimar la combinación lineal $W(\underline{x}) = \underline{A}^T \underline{Z}(\underline{x})$ considerándola como una nueva función aleatoria.

Aunque en la práctica un modelo de variograma $\gamma_w(\underline{h})$ será estimado y ajustado para la función aleatoria $W(\underline{x})$ y no a las componentes por separado, resulta útil relacionarlo con los variogramas separados.

$$\begin{aligned}\gamma_w(\underline{h}) &= \frac{1}{2} E \left\{ \left[\underline{A}^T \underline{Z}(\underline{x} + \underline{h}) - \underline{A}^T \underline{Z}(\underline{x}) \right]^2 \right\} \\ &= \frac{1}{2} \underline{A} E \left\{ \left[\underline{Z}(\underline{x} + \underline{h}) - \underline{Z}(\underline{x}) \right] \left[\underline{Z}(\underline{x} + \underline{h}) - \underline{Z}(\underline{x}) \right]^T \right\} \underline{A}^T \\ &= \frac{1}{2} \underline{A} \underline{\gamma}(\underline{h}) \underline{A}^T\end{aligned}\quad (5.66)$$

donde $\underline{\gamma}(\underline{h})$ es la matriz de semivarianzas cruzadas.

El estimador puede ser expresado como:

$$W^*(\underline{x}_0) = \sum_{j=1}^n \lambda_j W(\underline{x}_j), \quad \sum_{j=1}^n \lambda_j = 1 \quad (5.67)$$

La varianza de la estimación para $W^*(\underline{x}_0)$ es :

$$\sigma_{KD}^2 = \sum_{j=1}^n \gamma_w(\underline{x}_0 - \underline{x}_j) \lambda_j - \mu \quad (5.68)$$

o puede ser expresada equivalentemente como

$$\sigma_{KD}^2 = \frac{1}{2} \sum_{j=1}^n \underline{A} \underline{\gamma}(\underline{x}_0 - \underline{x}_j) \lambda_j \underline{A}^T - \mu \quad (5.69)$$

b) Estimación conjunta.

En lugar de la estimación directa de $W(\underline{x}_0)$ se estima $\underline{Z}(\underline{x}) = [Z_1(\underline{x}) \dots Z_m(\underline{x})]^T$ por cokriging y entonces se forma $W^{**}(\underline{x}_0) = \underline{A}^T \underline{Z}^*(\underline{x}_0)$. El estimador resultante se puede expresar como:

$$W^{**}(\underline{x}_0) = \underline{A}^T \underline{Z}^*(\underline{x}_0) = \sum_{j=1}^n \underline{A}^T \underline{\Gamma}_{\underline{j}} \underline{Z}(\underline{x}_j) \quad (5.70)$$

Una condición suficiente para que coincidan ambos estimadores es que se cumpla:

$$\lambda_j \underline{A}^T = \underline{\Gamma}_{\underline{j}} \underline{A}^T, \quad \forall j \quad (5.71)$$

La varianza de la estimación está dada por

$$\sigma_{KC}^2 = \sum_{j=1}^n \underline{A} \underline{\gamma}(\underline{x}_0 - \underline{x}_j) \underline{\Gamma}_{\underline{j}} \underline{A}^T - \underline{A} \underline{\mu} \underline{A}^T \quad (5.72)$$

Como vemos la condición anterior es casi suficiente para que $\sigma_{KD}^2 = \sigma_{KC}^2$ pero sería necesario también que se cumpliera que $\underline{\mu} = A\underline{\mu}A^T$. No obstante en general $W^*(\underline{x}_0) \neq W^{**}(\underline{x}_0)$ y $\sigma_{KD}^2 \neq \sigma_{KC}^2$.

El caso particular de estimar las componentes de $\underline{Z}(\underline{x}) = [Z_1(\underline{x}) \dots Z_m(\underline{x})]^T$ por separado en lugar de la estimación conjunta produce una estimación no óptima y necesita un factor de ajuste.

c) Comentarios generales

El método directo tiene un número de ventajas obvias:

- Sólo un variograma debe ser modelado.
- Requiere menor tiempo de cómputo debido a que las componentes de la matriz de coeficientes son escalares en lugar de matrices.

No obstante cualquier cambio en las componentes de la matriz A implica una nueva modelación del variograma y la resolución de un sistema de ecuaciones para Kriging. Si los coeficientes de A cambian constantemente es preferible utilizar la estimación conjunta ya que esto no implica ningún cambio, solamente multiplicar por la nueva matriz A .

El uso de una variable compuesta podría obscurecer las diferencias en anisotropía o la presencia de tendencia en algunas variables.

5.12.3 El problema de variables pobremente muestreadas

En contraste con los problemas donde el interés es estimar varias funciones aleatorias simultáneamente mediante el uso del estimador Cokriging para todas las variables en todas las ubicaciones muestrales, pocas veces los datos muestrales en otras variables es usado para mejorar la estimación de la variable primaria o más comúnmente para compensar muestras perdidas de la variable primaria.

Sean $\{\underline{x}_1, \dots, \underline{x}_n\}$ las posiciones de la red de muestreo y $\underline{Z} = [Z_1, \dots, Z_m]^T$ el vector de las funciones aleatorias. Supongamos que los datos en los puntos $\underline{x}_2, \dots, \underline{x}_n$ son obtenidos para todas las variables pero en \underline{x}_1 sólo es obtenido para Z_2, \dots, Z_m .

Deseamos estimar $Z_1(\underline{x}_1)$ usando todos los datos disponibles suponiendo que $\underline{\gamma}(\underline{h})$ es conocido.

El estimador lineal del cokriging $Z_1^*(\underline{x}_1)$ puede ser escrito en la forma usual:

$$\underline{Z}_1^*(\underline{x}_1) = \sum_{k=1}^n \Gamma_k^1 \underline{Z}(\underline{x}_k) \quad (5.73)$$

donde $\underline{\Gamma}_k^1 = \begin{bmatrix} \lambda_{11}^k \\ \dots \\ \lambda_{1m}^k \end{bmatrix}$ y se requiere que $\lambda_{11}^1 = 0$. La condición de universalidad está dada por:

$$\sum_{k=1}^n \underline{\Gamma}_k^1 = \begin{bmatrix} 1 \\ 0 \\ \dots \\ 0 \end{bmatrix} \quad (5.74)$$

No hay pérdida de generalidad si se considera que Z_1 es la variable primaria y si los datos están perdidos para Z_1 en otras ubicaciones k_0 o para otras variables j_0 en otras posiciones se deben imponer condiciones adicionales de la forma $\lambda_{1j_0}^{k_0} = 0$, ya que $\lambda_{1j_0}^{k_0}$ determina la contribución de $Z_{j_0}(\underline{x}_{k_0})$ a la estimación de Z_1^* .

5.13 Modelo de correlación intrínseco

Dado un vector de n funciones aleatorias $\underline{Z}(\underline{x}) = [Z_1(\underline{x}), Z_2(\underline{x}), \dots, Z_n(\underline{x})]^T$, el modelo más simple para su matriz de covarianzas conjuntas (cruzadas) es el siguiente:

$$\underline{C}(\underline{h}) = \underline{V} \underline{\rho}(\underline{h}) \quad (5.75)$$

donde \underline{V} es la matriz de covarianzas cruzadas en $\underline{h} = \underline{0}$, es decir

$$\sigma_{ij} = C_{ij}(\underline{0}) = \text{Cov}\{Z_i(\underline{x}), Z_j(\underline{x})\} = E[(Z_i(\underline{x}) - m_i)(Z_j(\underline{x}) - m_j)] \quad (5.76)$$

y $\underline{\rho}(\underline{h})$ es la función de correlación directa.

Recordando que la función de correlación cruzada por definición se expresa como:

$$\rho_{ij}(\underline{h}) = \frac{C_{ij}(\underline{h})}{\sqrt{C_{ii}(\underline{h})C_{jj}(\underline{h})}} \quad (5.77)$$

y empleando el modelo de covarianzas cruzadas introducido en la Ec. (5.75), es decir,

$$C_{ij}(\underline{h}) = \sigma_{ij} \underline{\rho}(\underline{h}) \quad (5.78)$$

entonces resulta que el coeficiente de correlación entre dos funciones aleatorias $Z_i(\underline{x})$ y $Z_j(\underline{x})$ está dado por

$$r_{ij} = \frac{\sigma_{ij} \underline{\rho}(\underline{h})}{\sqrt{\sigma_{ii} \underline{\rho}(\underline{h}) \sigma_{jj} \underline{\rho}(\underline{h})}} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii} \sigma_{jj}}} \quad (5.79)$$

y no depende de la escala espacial. Por este motivo, el modelo propuesto es conocido como *modelo de correlación intrínseco*.

Un modo de verificar la existencia de correlación intrínseca se basa precisamente en la propiedad anterior. Para lo cual se calcula y grafica el *coeficiente de codispersión* el cual se define como:

$$cc_{ij}(\underline{h}) = \frac{\gamma_{ij}(\underline{h})}{\sqrt{\gamma_{ii}(\underline{h})\gamma_{jj}(\underline{h})}} \quad (5.80)$$

Si el coeficiente de codispersión es constante para todo $i \neq j$, esto es un indicador de que la correlación no depende de la escala espacial y por lo tanto se puede considerar un modelo de correlación intrínseca.

El modelo lineal asociado al modelo de correlación intrínseco se escribe como:

$$\underline{Z}(\underline{x}) = \underline{A}\underline{Y}(\underline{x}) \quad (5.81)$$

o equivalentemente

$$Z_i(\underline{x}) = \sum_j^p a_{ij} Y_j(\underline{x}) \quad (5.82)$$

donde \underline{A} es la matriz $n \times p$ de los coeficientes de la transformación lineal $[a_{ij}]$, $\underline{Y}(\underline{x}) = [Y_1(\underline{x}), Y_2(\underline{x}), \dots, Y_p(\underline{x})]^T$ es un vector de p funciones aleatorias estacionarias de segundo orden no correlacionadas (independientes entre sí), cuya función directa de correlación $\rho(\underline{h})$ no depende del índice j .

Para un modelo de correlación intrínseco dado existen muchas matrices \underline{A} que satisfacen $\underline{V} = \underline{A}\underline{A}^T$, pero el modo natural para obtener los coeficientes de la transformación Ec. (5.82) lo ofrece el Análisis de Componentes Principales (ACP), el cual está basado en la descomposición en valores propios de la matriz de covarianzas \underline{V} y los factores $Y_j(\underline{x})$ pueden ser interpretados como componentes principales. Estas componentes se obtienen imponiendo la condición de que expliquen en una proporción elevada (usualmente 80% o más) la varianza total $\sigma_{\underline{Z}}^2$ del vector de funciones aleatorias $\underline{Z}(\underline{x})$.

El objetivo que persigue el análisis de componentes principales es transformar el vector de funciones aleatorias correlacionadas $\underline{Z}(\underline{x})$ de dimensión n a un vector de funciones aleatorias (componentes principales) no correlacionadas $\underline{Y}(\underline{x})$ de dimensión p , donde se espera que $p \ll n$.

Sean $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n > 0$ los valores propios y $\underline{q}_1, \underline{q}_2, \dots, \underline{q}_n$ los correspondientes vectores propios asociados de la matriz $\underline{\underline{V}}$, donde se cumple que $\underline{q}_i^T \underline{q}_j = \delta_{ij}$. Las mayores varianzas están directamente relacionadas con los mayores valores propios λ_i de la matriz de covarianzas de los datos $\underline{\underline{V}}$,

$$\sigma_{\underline{\underline{Z}}}^2 = \text{Tr} \underline{\underline{V}} = \sum_{i=1}^n \sigma_{ii} = \sum_{j=1}^p \lambda_j \quad (5.83)$$

por los que se escogen las primeras p componentes ordenadas de mayor a menor en orden decreciente del por ciento de la varianza total que explican.

Entonces, aplicando la descomposición en valores propios de la matriz de covarianzas $\underline{\underline{V}}$, ésta se puede expresar como:

$$\underline{\underline{V}} \cong \underline{\underline{A}} \underline{\underline{A}}^T \quad (5.84)$$

con $\underline{\underline{A}} = \underline{\underline{\tilde{Q}}} \sqrt{\underline{\underline{\Lambda}}}$ y $\underline{\underline{\tilde{Q}}} \underline{\underline{\tilde{Q}}}^T = \underline{\underline{I}}$, donde $\underline{\underline{\tilde{Q}}}$ y $\underline{\underline{\Lambda}}$ son las matrices de los p vectores propios y de los p valores propios correspondientemente.

Entonces los coeficientes a_{ij} de la matriz $\underline{\underline{A}}$ están dados por:

$$a_{ij} = \sqrt{\lambda_j} q_{ij} \quad (5.85)$$

y por lo tanto la Ec.(5.82) se transforma en

$$Z_i(\underline{x}) = \sum_j^p \sqrt{\lambda_j} q_{ij} Y_j(\underline{x}) \quad (5.86)$$

donde q_{ij} es la componente i del vector propio \underline{q}_j y λ_j es el valor propio asociado a éste.

Las componentes se relacionan con las variables originales mediante

$$\underline{Y}(\underline{x}) = \underline{\underline{\tilde{Q}}}^T \underline{Z}(\underline{x}) \quad (5.87)$$

o puede escribirse escalado con varianza unitaria

$$Y_j(\underline{x}) = \frac{1}{\sqrt{\lambda_j}} \underline{q}_j^T \underline{Z}(\underline{x}) = \frac{1}{\sqrt{\lambda_j}} \sum_{i=1}^n q_{ij} Z_i(\underline{x}); \quad j = 1, \dots, p \quad (5.88)$$

Lo cual se verifica si suponemos que $\underline{Y}(\underline{x}) = \underline{\underline{B}} \underline{Z}(\underline{x})$, donde $\underline{\underline{B}}$ es una matriz $p \times n$ que cumple que $\underline{\underline{B}}^T \underline{\underline{B}} = \underline{\underline{I}}$, entonces

$$\underline{Y} \underline{Y}^T = \underline{\underline{B}} \underline{Z} \underline{Y}^T = \underline{\underline{B}} \underline{Z} (\underline{\underline{B}} \underline{Z})^T = \underline{\underline{B}} \underline{Z} \underline{Z}^T \underline{\underline{B}}^T \quad (5.89)$$

aplicando el operador valor esperado a ambos miembros se obtiene

$$E[\underline{Y} \underline{Y}^T] = \underline{\underline{B}} E[\underline{Z} \underline{Z}^T] \underline{\underline{B}}^T \quad (5.90)$$

$$\underline{\underline{D}} = \underline{\underline{B}} \underline{\underline{V}} \underline{\underline{B}}^T \quad (5.91)$$

multiplicando ambos miembros por $\underline{\underline{B}}$ resulta

$$\underline{\underline{DB}} = \underline{\underline{BV}} \quad (5.92)$$

donde inmediatamente se puede ver que la matriz $\underline{\underline{Q}}$ ofrece la solución ya que satisface:

$$\underline{\underline{\Lambda \tilde{Q}^T}} = \underline{\underline{\tilde{Q}^T V}} \quad (5.93)$$

Una vez conocidos los valores de las componentes $Y_j(\underline{x})$ en los puntos de medición usando Ec.(5.88), se puede estimar sus valores en cualquier otro punto usando Kriging ordinario y a partir de éstos se obtienen los valores estimados de las variables originales $Z_i(\underline{x})$ empleando la Ec. (5.82).

El modelo de correlación intrínseco, también conocido como *corregionalización isotópica*, es un caso importante de referencia cuando las variables están todas muestreadas en los mismos puntos de observación, ya que la estimación conjunta se simplifica, puesto que

- 1) Se reducen las dimensiones del problema original a través del análisis de componentes principales, con las consecuentes ventajas de reducción en el almacenamiento de la información.
- 2) Se usa el Kriging Ordinario para p componentes en lugar de Cokriging para n variables, donde usualmente $p \ll n$.

5.14 Análisis de Componentes Principales basado en el Modelo Lineal de Corregionalización

El modelo de Análisis de Componentes Principales estándar es sólo adecuado para datos que puedan ser considerados como una realización de un modelo de correlación intrínseco. El MCL cubre una clase más general de fenómenos que los que pueden ser descritos con un variograma multivariado anidado. En este modelo tenemos diferentes estructuras de correlación a diferentes escalas del espacio (tiempo). Un Análisis de Componentes Principales Corregionalizado se realiza separadamente para cada una de estas escalas.

El análisis de valores propios es realizado en cada matriz de corregionalización $\underline{\underline{V}}_k$ del variograma multivariado anidado. Los coeficientes de correlación r_{ij}^k entre las variables originales y las componentes principales en una escala dada k pueden ser usados para construir círculos de correlación que ayudan a entender la estructura de correlación en esa escala espacial (temporal).

Los valores estimados de las componentes principales $Y_{jk}^*(\underline{x})$ son obtenidos mediante la aplicación de Cokriging y pueden ser graficados como mapas.

El Análisis de Componentes Principales Corregionalizado (ACPC) ha sido exitosamente aplicado en numerosos casos de estudio.

5.15 Dificultades y Aspectos Prácticos del Cokriging.

El problema de estimar varias variables corregionalizadas simultáneamente usando el Cokriging es el enfoque más riguroso y el que se basa en un menor número de hipótesis. Sin embargo, requiere que se disponga de un número relativamente elevado de puntos muestrales donde estén medidas todas las variables para una adecuada estimación de los semivariogramas cruzados. Cuando no se cumple este requisito el Cokriging puede perder su superioridad sobre otros métodos alternativos.

La mayoría de los problemas encontrados en la práctica del Cokriging son los mismos que los encontrados en la práctica del Kriging pero quizás magnificados por el número de variables incorporadas, por una parte, exigiendo un tiempo de cálculo considerable en la modelación de los semivariogramas cruzados mediante la modelación y validación de múltiples semivariogramas y por otra en la estimación de las variables debido al aumento de la complejidad de los sistemas de ecuaciones a resolver.

5.16 Algunos métodos alternativos al Cokriging

A continuación revisaremos algunas técnicas alternativas al Cokriging.

5.16.1 Kriging combinado con Regresión Lineal

Este método fue propuesto por Delhomme(1974,1976) y consiste en establecer un modelo de regresión lineal entre las variables $Z(\underline{x})$ y $Y(\underline{x})$ de la manera siguiente:

$$Z(\underline{x}) = aY(\underline{x}) + b \quad (5.94)$$

donde a y b son los parámetros de la recta de regresión determinados a partir de l pares de datos $\{Z(\underline{x}_i), Y(\underline{x}_i)\}$ si las variables regionalizadas Z y Y están muestreadas en l puntos simultáneamente, y además consideremos que Z está muestreada en los l puntos mientras que Y en m puntos, donde $m > l$, es decir Y está más densamente muestreada que Z .

Entonces podemos estimar a Z en los $m-l$ puntos restantes a través de Y usando el modelo de regresión lineal anteriormente obtenido como sigue:

$$\hat{Z}(\underline{x}_j) = aY(\underline{x}_j) + b, \quad j = l+1, \dots, m \quad (5.95)$$

y las varianzas de los errores de la predicción por regresión están dadas por

$$\sigma_j^2 = \sigma_e^2 \left[\frac{1}{N} + \frac{(Y_j - \bar{Y})^2}{\sum_{i=1}^N (Y_i - \bar{Y})^2} \right], \quad j = l+1, \dots, m \quad (5.96)$$

donde $\bar{Y} = \frac{1}{N} \sum_{i=1}^N Y_i$, $\sigma_e^2 = \frac{1}{N-2} \sum_{i=1}^N (Z_i - aY_i - b)^2$,

Con los l datos originales de Z se estima el variograma y usando todos los m datos se realiza mediante Kriging su estimación espacial.

5.16.2 Kriging con una Deriva Externa

Este método fue desarrollado por el grupo de Matheron de la Escuela de Minas de Fontainebleau y no ha sido muy usado.

Para su aplicación se requiere que el valor esperado de una variable Z sea una función lineal conocida dependiente de otra variable Y , como sigue:

$$E[Z(\underline{x}_i)|Y(\underline{x}_i)] = c_1 Y(\underline{x}_i) + c_2 \quad (5.97)$$

donde c_1 y c_2 son dos constantes, que no se requieren conocer.

Además se necesita que la segunda variable Y haya sido muestreada en un gran número de puntos y que ofrezca una información bastante buena acerca de la estructura de correlación de la primera variable Z .

Considerando el estimador lineal habitual del Kriging para Z en un punto no muestral \underline{x}_k :

$$Z^*(\underline{x}_k) = \sum_{i=1}^N \lambda_i Z(\underline{x}_i) \quad (5.98)$$

Si aplicamos las condiciones del Kriging:

sesgo nulo

$$E[Z^*(\underline{x}_k) - Z(\underline{x}_k)|Y(\underline{x}_k)] = 0 \quad (5.99)$$

y varianza del error mínima

$$\min E\left\{[Z^*(\underline{x}_k) - Z(\underline{x}_k)]^2 | Y(\underline{x}_k)\right\} = 0 \quad (5.100)$$

obtenemos:

$$\begin{cases} \sum_{j=1}^N \lambda_j \tilde{\gamma}_{ij} + \mu_1 + \mu_2 Y(\underline{x}_i) = \tilde{\gamma}_{ik}, & i = 1, \dots, N \\ \sum_{j=1}^N \lambda_j = 1 \\ \sum_{j=1}^N \lambda_j Y(\underline{x}_j) = Y(\underline{x}_k) \end{cases} \quad (5.101)$$

donde $\tilde{\gamma}_{ij}$ es el semivariograma de Z condicionado a la variable Y .

La varianza condicionada del error de la estimación viene dada por:

$$\tilde{\sigma}_e^2 = Var\left[Z^*(\underline{x}_k) - Z(\underline{x}_k) | Y(\underline{x}_k)\right] = \sum_{i=1}^N \lambda_i \tilde{\gamma}_{ik} + \mu_1 + \mu_2 Y(\underline{x}_k) \quad (5.102)$$

Nótese que hay que conocer a la variable Y en los puntos \underline{x}_k (o estimarla) y el semivariograma $\tilde{\gamma}_{ij}$ coincide con el semivariograma de los residuos $R(\underline{x})$ obtenidos mediante:

$$R(\underline{x}_i) = Z(\underline{x}_i) - c_1 Y(\underline{x}_i) - c_2 \quad (5.103)$$

En la práctica suele reemplazarse el semivariograma $\tilde{\gamma}_{ij}$ por el semivariograma de Z sin condicionamiento. Esto parece no afectar significativamente los valores estimados de Z y tiene el efecto de sobrevalorar la varianza del error.

5.17 Kriging Factorial

Un método tradicionalmente relacionado con el de componentes principales es el análisis factorial., por lo que se puede definir de manera análoga el Kriging factorial que fue desarrollado por Matheron (1982). Este se basa en la descomposición de la estructura de correlación espacial de las variables $Z_i(\underline{x})$ como una combinación lineal de varias subestructuras básicas, es decir,

$$\underline{\underline{C}}(\underline{h}) = \sum_{e=1}^{N_e} \underline{\underline{D}}^e g_e(\underline{h}) \quad (5.104)$$

o equivalentemente

$$C_{ij}(\underline{h}) = \sum_{e=1}^{N_e} D_{ij}^e g_e(\underline{h}) \quad (5.105)$$

donde $g_e(\underline{h})$ es la subestructura básica e , N_e es el número total de subestructuras y $\underline{\underline{D}}^e$ son matrices $n \times n$ positivas definidas que en principio deben ser conocidas. Esta descomposición obedece a su vez a la descomposición de cada variable $Z_i(\underline{x})$ como la suma de una serie de factores $F_i^e(\underline{x})$, es decir

$$Z_i(\underline{x}) = \sum_{e=1}^{N_e} F_i^e(\underline{x}) + m_i(\underline{x}) \quad (5.106)$$

donde $m_i(\underline{x})$ son los valores medios de $Z_i(\underline{x})$, $E[F_i^e(\underline{x})] = 0$ y su covarianza cruzada está dada por:

$$E[F_i^e(\underline{x} + \underline{h}) F_i^{e'}(\underline{x})] = \begin{cases} D_{ij}^e g_e(\underline{h}), & \text{si } e = e' \\ 0, & \text{si } e \neq e' \end{cases} \quad (5.107)$$

Estos factores se expresan como combinaciones lineales de una serie de variables $Y_j^e(\underline{x})$ independientes entre sí

$$F_i^e(\underline{x}) = \sum_{j=1}^m a_{ij}^e Y_j^e(\underline{x}) \quad (5.108)$$

donde los coeficientes a_{ij}^e se obtienen mediante la descomposición:

$$\underline{\underline{D}}^e = \underline{\underline{A}} \underline{\underline{A}}_e^T, \quad e = 1, \dots, N_e \quad (5.109)$$

que siempre es factible puesto que las matrices $\underline{\underline{D}}^e$ son positivas definidas. Las variables $Z_i(\underline{x})$ se relacionan con las componentes $Y_j^e(\underline{x})$ mediante

$$Z_i(\underline{x}) = \sum_{e=1}^{N_e} \sum_{j=1}^m a_{ij}^e Y_j^e(\underline{x}); \quad i = 1, \dots, n \quad (5.110)$$

El objetivo del Kriging factorial no es estrictamente la estimación conjunta de las variables originales $Z_i(\underline{x})$, si no mas bien se pretende estimar la distribución espacial de la contribución de cada una de las componentes. En ciertos casos es posible interpretar la descomposición Ec.(5.104) como la descomposición de la variabilidad a distintas escalas según Wackernagel, et al. (1989).

Las componentes $Y_k^e(\underline{x})$ pueden estimarse mediante Kriging ordinario a partir de los datos de las variables originales $Z_i(\underline{x})$, es decir

$$Y_k^e(\underline{x}_0) = \sum_{i=1}^n \sum_{\alpha=1}^N \lambda_{i\alpha}^e Z_i(\underline{x}_\alpha); \quad k = 1, \dots, m; \quad e = 1, \dots, N_e \quad (5.111)$$

donde los coeficientes del Kriging $\lambda_{i\alpha}^e$ se obtienen como es habitual imponiendo la condición de sesgo nulo y que sea mínima la varianza del error de la estimación.

Las ecuaciones del Kriging factorial que resultan son:

$$\begin{cases} \sum_{j=1}^n \sum_{\beta=1}^N \lambda_{j\beta}^e C_{ij}(\underline{x}_\alpha - \underline{x}_\beta) + \mu_i^e = a_{ik}^e g_e(\underline{x}_\alpha - \underline{x}_0), & i = 1, \dots, n \\ \sum_{\beta=1}^N \lambda_{i\beta}^e = 0, & \alpha = 1, \dots, N \end{cases} \quad (5.112)$$

para cada estructura $e = 1, \dots, N_e$ y para cada componente $k = 1, \dots, m$.

Y la varianza de la estimación es:

$$\begin{aligned}\sigma_{e,k}^2 &= E\left[\left(Y_k^{e*}(\underline{x}) - Y_k^e(\underline{x})\right)^2\right] \\ &= 1 + \sum_{i=1}^n \sum_{j=1}^n \sum_{\alpha=1}^N \sum_{\beta=1}^N \lambda_{i\alpha}^e \lambda_{j\beta}^e C_{ij}(\underline{x}_\alpha - \underline{x}_\beta) - 2 \sum_{i=1}^n \sum_{\alpha=1}^N \lambda_{i\alpha}^e a_{ik}^e g_e(\underline{x}_\alpha - \underline{x}_0)\end{aligned}\quad (5.113)$$

Aplicación del Kriging Factorial

El Kriging Factorial permite separar un vector de n funciones aleatorias en m componentes, cada una representando una escala de variación diferente (o longitud de onda) del fenómeno que se estudia. Por ejemplo, en las mediciones magnéticas o gravimétricas uno desea distinguir las longitudes de onda largas asociadas con las fuentes profundas de las cortas que reflejan fuentes superficiales. En éste sentido se puede considerar un método de filtrado.

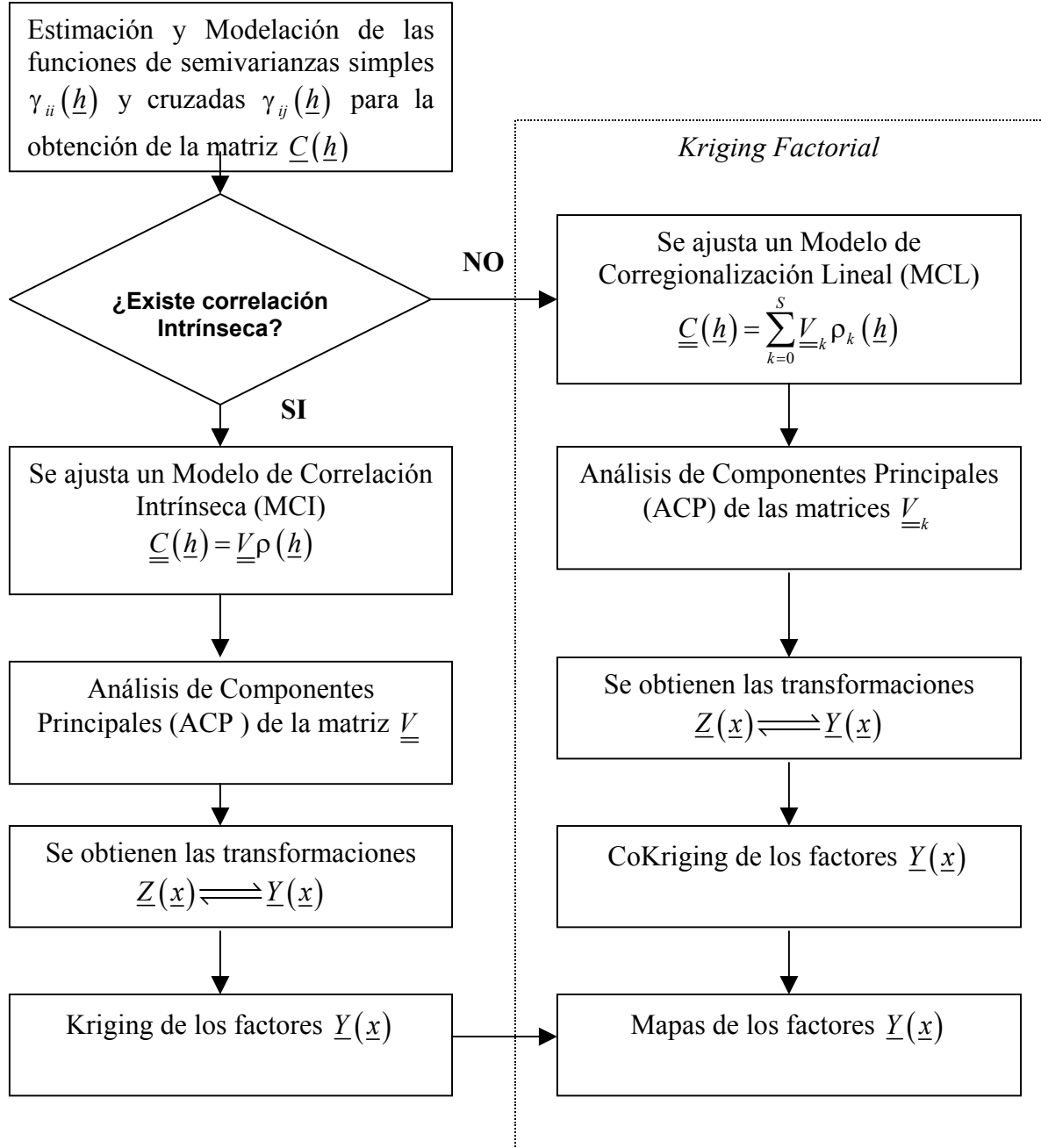
Esta separación o filtrado se realiza tradicionalmente usando análisis espectral en dos o más dimensiones, pero usando el Kriging Factorial es posible e incluso ventajoso obtener la determinación espacial de las componentes que se deseen directamente. Ya que los métodos espectrales tienen ciertas limitaciones como son:

- a) requieren de una malla rectangular completa, que si no se posee, se requiere entonces interpolar o aún peor extrapolar (se usa en ocasiones para esto el Kriging) lo cual altera las características espectrales,
- b) opaca la resolución de las anomalías locales sobre todas las frecuencias.

El Kriging Factorial se ha aplicado con éxito en prospección geoquímica (Sandjivy, 1984), prospección geofísica Galli et al., 1984, Chiles y Guillén, 1984, exploración petrolera (Jaquet, 1989), etc.

5.18 Esquema General para el Análisis de Corregionalización

Dado un vector $\underline{Z}(\underline{x})$ de n funciones aleatorias se procede de la siguiente manera:



5.19 Manejo de procesos espacio-temporales

Existen un gran número de fenómenos cuyo estudio de manera natural requiere de un marco espacio-temporal. Este es el caso, por ejemplo, de las variables meteorológicas o la evolución de un contaminante en un acuífero, etc. Históricamente, se han desarrollado dos vertientes principales usando un enfoque estadístico, una que se encarga de la modelación de fenómenos temporales, que son las series de tiempo o cronológicas, y la otra que estudia los procesos que varían sólo en el espacio, de la cual se ha encargado la estadística espacial y en particular la geoestadística.

Por lo anteriormente planteado resulta conveniente definir procesos estocásticos espacio-temporales que pueden ser caracterizados mediante funciones aleatorias $Z(\underline{x}, t)$ que varían en el espacio, es decir dependen de la posición $\underline{x} \in \mathbb{R}^3$ y dependen del tiempo t . Dicho de otra manera, a cada punto del espacio-tiempo (\underline{x}, t) le hacemos corresponder una variable aleatoria Z .

En la práctica, el problema a resolver consiste en estimar valores en puntos desconocidos (\underline{x}_0, t_0) de la función aleatoria $Z(\underline{x}, t)$, dados un conjunto finito, y con frecuencia reducido, de mediciones en espacio $\{\underline{x}_1, \underline{x}_2, \dots, \underline{x}_M\}$ y en tiempo $\{t_1, t_2, \dots, t_N\}$. Hasta la fecha no existen un método general para la solución de éste problema, no obstante existe una serie de alternativas propuestas para su manejo aplicables bajo ciertas condiciones. A continuación enumeraremos las más comunes:

- 1) Se considera al tiempo como otra coordenada adicional, para esto es necesario definir un vector velocidad de transmisión $\underline{v} = \frac{\underline{x}}{t}$, (Bilonik, 1985).
- 2) Se separa de manera explícita la dependencia espacial de la temporal, considerando que la deriva de $Z(\underline{x}, t)$ está compuesta por una componente espacial $m_{\underline{x}}(\underline{x})$ y otra temporal $m_t(t)$, y la función de covarianzas se modela como la suma de una covarianza espacial y otra temporal, (Rouhani, 1989).
- 3) Se supone que para un tiempo t_i dado, $Z(\underline{x}, t_i)$ es una función aleatoria intrínseca, cuyo semivariograma depende de t_i , a través de uno de sus parámetros, Bastin et al., 1984).
- 4) Se supone que en cada punto \underline{x}_j , $Z(\underline{x}_j, t)$ es una función aleatoria cuya media y covarianza dependen del punto \underline{x}_j . Es decir, se definen M funciones aleatorias $\{Z(\underline{x}_1, t), Z(\underline{x}_2, t), \dots, Z(\underline{x}_M, t)\}$ correlacionadas cuya estimación puede hacerse mediante Cokriging y resulta eficiente cuando $M \ll N$, (Solow y Gorelick, 1986).
- 5) Es lo opuesto al caso anterior, se considera que para un tiempo t_i dado, $Z(\underline{x}, t_i)$ es una función aleatoria, entonces tenemos N funciones aleatorias

espaciales $\{Z(\underline{x}, t_1), Z(\underline{x}, t_2), \dots, Z(\underline{x}, t_N)\}$ correlacionadas entre sí. De manera similar este enfoque se resuelve usando Cokriging y resulta eficiente cuando $M \gg N$, (Myers, 1988).

CAPITULO 6: SIMULACIÓN GEOESTADISTICA

6.1 Introducción

Una variable regionalizada $z(\underline{x})$ es interpretada como una realización de una cierta función aleatoria $Z(\underline{x})$. Esta función aleatoria $Z(\underline{x})$ puede ser caracterizada por una función de distribución de probabilidad y, en dependencia de su grado de estacionaridad, por una función de correlación (función de covarianzas o semivarianzas).

La idea básica detrás de la simulación estadística consiste en obtener nuevas realizaciones “artificiales” $Z_s(\underline{x})$ de la función aleatoria $Z(\underline{x})$ de manera tal que éstas reflejen las mismas propiedades estadísticas que se esperan que posea la función aleatoria $Z(\underline{x})$. Pero como por lo general no conocemos con precisión las propiedades estadísticas de $Z(\underline{x})$ y cuando más lo que podemos hacer es inferirlas a través de una sola realización o muestra de la función aleatoria, entonces lo que se hace es intentar obtener realizaciones simuladas $Z_s(\underline{x})$ que sean estadísticamente equivalentes a la muestra que se posee de la función aleatoria.

La equivalencia estadística en un sentido estricto significa que todas las realizaciones $Z_s(\underline{x})$ tengan la misma distribución de probabilidad de la función aleatoria $Z(\underline{x})$ que se simula, pero en la mayoría de los casos nos tenemos que conformar con que al menos posean los mismos momentos de primer y segundo orden que inferimos a partir de una muestra de $Z(\underline{x})$.

Resulta deseable en muchas aplicaciones quedarse solamente con aquellas realizaciones $Z_s(\underline{x})$ que en los puntos muestrales $\{\underline{x}_i, i = 1, \dots, n\}$ coinciden los valores simulados $Z_s(\underline{x}_i)$ con los valores reales o experimentales $Z_M(\underline{x}_i)$. A estas realizaciones $Z_s(\underline{x})$ de la función aleatoria $Z(\underline{x})$ se les conoce como "*simulaciones condicionales*" $Z_{sc}(\underline{x})$ del fenómeno regionalizado $Z(\underline{x})$.

6.2 Objetivos de la simulación

Las estimaciones espaciales de un fenómeno regionalizado que se pueda describir mediante una función aleatoria $Z(\underline{x})$ son con frecuencia insuficientes debido a múltiples factores, pero sobre todo debido a la carencia de suficiente información (mediciones) acerca de la función aleatoria $Z(\underline{x})$.

Como desafortunadamente no se dispone de un conocimiento exacto de la realidad "*in situ*" y la información disponible en muchos casos está usualmente muy fragmentada y se limita fundamentalmente al conocimiento de unos pocos puntos muestrales, las estimaciones obtenidas a partir de esta información incluso empleando el estimador Kriging, son demasiadas imprecisas para los cálculos exactos de las dispersiones que se requieren en ciertas aplicaciones.

¿Cómo es imposible estimar la realidad "*in situ*" correctamente, con suficiente detalle?

Una idea simple es simular esta realidad en base a un modelo, por lo que la realidad y la simulación son variables diferentes de un mismo fenómeno.

Consideremos el siguiente ejemplo. Tenemos mediciones reales del fenómeno, es decir, una realización $\{Z_M(\underline{x}_i), i = 1, \dots, n\}$ de la función $Z(\underline{x})$ en ciertos puntos \underline{x}_i de la región a estudiar. El enfoque geoestadístico consiste en interpretar la distribución espacial de la variable regionalizada $Z_M(\underline{x})$ como una realización particular de la función aleatoria $Z(\underline{x})$. Esta función aleatoria está caracterizada por su función de distribución de probabilidad o por sus dos primeros momentos, los cuales son estimados a partir de datos experimentales.

Este modelo es entonces adecuado para el problema práctico de la determinación de medidas de la dispersión de los valores observados $Z_M(\underline{x}_i)$, ya que las varianzas de la dispersión de $Z(\underline{x})$ pueden ser expresadas como una función del momento de segundo orden solamente (covarianza o variograma). Una simulación entonces consiste en obtener otra realización $Z_S(\underline{x})$ de esta función aleatoria $Z(\underline{x})$. Las dos realizaciones la real y la simulada difiere una de la otra en determinadas localizaciones pero ambas pertenecen a la misma función aleatoria $Z(\underline{x})$, es decir tienen la misma función de distribución y los mismos momentos de primer y segundo órdenes por lo se dice que son estadísticamente equivalentes.

El fenómeno simulado tiene la ventaja de ser conocido en todos los puntos y no solamente en los puntos experimentales $\{Z_M(\underline{x}_i), i = 1, \dots, n\}$. Con frecuencia al fenómeno simulado se le llama "*modelo numérico*" del fenómeno real.

6.3 Condicionamiento

Existe un número infinito de realizaciones que cumplen con la condición de que sus valores simulados coinciden con los valores experimentales, es decir

$$Z_{SC}(\underline{x}_i) \equiv Z_M(\underline{x}_i); \quad i = 1, \dots, n \quad (6.1)$$

donde $Z_M(\underline{x}_i)$ es el valor muestral de la función aleatoria $Z(\underline{x})$ en el punto \underline{x}_i .

Esta condición le confiere una cierta robustez a la simulación condicionada $Z_{SC}(\underline{x})$ con respecto a las características de los datos reales $\{Z_M(\underline{x}_i), i = 1, \dots, n\}$ los cuales no son modelados explícitamente por la función aleatoria $Z(\underline{x})$. Si por ejemplo un número suficiente de datos muestran una tendencia local entonces la simulación condicional que está basada incluso en un modelo estacionario reflejará la tendencia local en la misma zona.

La simulación condicional puede ser perfeccionada agregándole todo una suerte de información cualitativa disponible del fenómeno real. Como por ejemplo en el caso de un depósito se le puede añadir la geometría de las fallas principales, etc.

6.4 ¿ Simulación o estimación ?

Los fenómenos simulados tienen los mismos valores experimentales en las localizaciones muestrales y tienen las mismas medidas de dispersión (al menos hasta de orden 2) que el depósito real.

Entonces, ¿en qué sentido la simulación condicional difiere de una estimación ?

La diferencia estriba en sus objetivos.

- A. El objetivo de una estimación es ofrecer en cada punto \underline{x} un valor estimado $Z^*(\underline{x})$ el cual es tan cercano como sea posible del valor real desconocido $Z(\underline{x})$.

Los criterios de la calidad de la estimación son:

- i.) que sea no sesgada

$$E[Z^*(\underline{x}) - Z(\underline{x})] = 0 \quad (6.2)$$

- ii.) y que la varianza el error sea mínima

$$\min_{\underline{x}} E\left[\{Z^*(\underline{x}) - Z(\underline{x})\}^2\right] \quad (6.3)$$

El estimador no tiene que reproducir la variabilidad espacial de los valores reales $Z(\underline{x})$. En el caso del Kriging este produce un suavizado de las dispersiones reales, dicho de otra forma, se subestima la variabilidad de los valores reales.

- B. Mientras que el objetivo de la simulación es reproducir las propiedades estadísticas de la función aleatoria $Z(\underline{x})$. Dicho explícitamente, la simulación $Z_s(\underline{x})$ posee los mismos momentos experimentales (media, covarianza o variograma, así como el histograma) que los valores reales $\{Z_M(\underline{x}_i), i=1, \dots, n\}$, los cuales caracterizan las principales propiedades estadísticas de $Z(\underline{x})$. En particular, en las simulaciones condicionadas en cada punto \underline{x} el valor simulado $Z_{sc}(\underline{x})$ no es el mejor estimador posible de $Z(\underline{x})$. Por el contrario, se puede mostrar que la varianza del error de la estimación de $Z(\underline{x})$ mediante simulación condicional $Z_{sc}(\underline{x})$ es exactamente el doble de la varianza del error usando Kriging.

6.5 La teoría de la simulación condicional

El problema consiste en construir una realización de la función aleatoria $Z_{sc}(\underline{x})$ isomorfa a $Z(\underline{x})$, es decir una realización que tenga el mismo valor esperado ($E[Z(\underline{x})]$) y el mismo momento de segundo orden ($C(\underline{h})$ o $\gamma(\underline{h})$). Además la realización $Z_{sc}(\underline{x})$ debe estar condicionada a los datos experimentales o sea que en los puntos experimentales los valores simulados deben coincidir con los valores observados.

$$Z_{sc}(\underline{x}_i) \equiv Z_M(\underline{x}_i); \quad i=1, \dots, n \quad (6.4)$$

Consideremos el valor real $Z(\underline{x})$ y el valor interpolado por Kriging $Z_K^*(\underline{x})$ a partir de los datos disponibles $\{Z_M(\underline{x}_i), i=1, \dots, n\}$. Estos dos valores difieren en un error desconocido

$$Z(\underline{x}) = Z_K^*(\underline{x}) + \{Z(\underline{x}) - Z_K^*(\underline{x})\} \quad (6.5)$$

Una propiedad característica del Kriging es que su error $Z(\underline{x}) - Z_K^*(\underline{x})$ es ortogonal a los valores interpolados

$$E\left[Z_K^*(\underline{y})\{Z(\underline{x}) - Z_K^*(\underline{x})\}\right] = 0; \quad \forall \underline{x}, \underline{y} \quad (6.6)$$

Para obtener la simulación condicional deseada es suficiente reemplazar el error del Kriging $Z(\underline{x}) - Z_K^*(\underline{x})$ por un error isomorfo e independiente $Z_S(\underline{x}) - Z_{SK}^*(\underline{x})$.

Dada una realización $Z_S(\underline{x})$ cuando aplicamos el Kriging a la misma configuración de datos $\{Z_S(\underline{x}_i), i = 1, \dots, n\}$ el error del Kriging resultará $Z_S(\underline{x}) - Z_{SK}^*(\underline{x})$ e independiente de $Z_K^*(\underline{x})$. La simulación condicional requerida es

$$Z_{SC}(\underline{x}) = Z_K^*(\underline{x}) + \{Z_S(\underline{x}) - Z_{SK}^*(\underline{x})\} \quad (6.7)$$

Los requerimientos de una simulación condicional son satisfechos ya que tenemos lo siguiente:

- a) La función aleatoria $Z_S(\underline{x})$ tiene el mismo valor esperado que $Z(\underline{x})$. Esto sale de la condición de insesgadez del Kriging

$$E[Z_K^*(\underline{x})] = E[Z(\underline{x})] \text{ y } E[Z_{SK}^*(\underline{x})] = E[Z_S(\underline{x})] \quad (6.8)$$

entonces

$$E[Z_{SC}(\underline{x})] = E[Z(\underline{x})] \quad (6.9)$$

- b) La función aleatoria $Z_{SC}(\underline{x})$ es isomorfa a $Z(\underline{x})$, esto sale de que el error del kriging de la simulación $Z_S(\underline{x}) - Z_{SK}^*(\underline{x})$ es isomorfo al error verdadero $Z_S(\underline{x}) - Z_S^*(\underline{x})$ e independiente de $Z_K^*(\underline{x})$. En otras palabras los variogramas de estas dos funciones aleatorias son idénticas pero no necesariamente sus covarianzas.

c) La realización simulada $Z_{SC}(\underline{x})$ es condicional a los datos experimentales.

Esto se obtiene del hecho que en los puntos de observación los valores interpolados por Kriging son iguales a los valores medidos

$$Z_K^*(\underline{x}_i) = Z_M(\underline{x}_i) \text{ y } Z_{SK}^*(\underline{x}_i) = Z_S(\underline{x}_i) \quad \forall \underline{x}_i, i = 1, \dots, n \quad (6.10)$$

6.6 Métodos de simulación de funciones aleatorias más conocidos

Las aplicaciones de la simulación de funciones aleatorias en los últimos tiempos han adquirido una importancia cada vez mayor en la industria del petróleo, en geofísica e hidrogeología lo cual ha hecho que esta área de la geoestadística sea la que más activamente se ha desarrollado. Debido a esto se han diversificado los métodos de simulación, por lo que intentar establecer alguna clasificación sistemática de los mismos resulta una tarea compleja y difícil. Aquí intentaremos dar una idea general de los métodos más conocidos y sus características.

Método de Simulación	Condicionado	Gaussiano	Malla Regular
Matricial	Si	Si	No
Espectral	No	Si	No
Bandas Rotantes	No	Si	No
Secuencial Gaussiano	Si	Si	No
Secuencial Indicador	Si	No	No
Gaussiano Truncado	Si	Si	No
Recocido (Annealing) Simulado	Si	No	Si

Una de las características distintivas de los métodos de simulación es si generan directamente simulaciones condicionadas o si se requiere de un postprocesamiento para condicionarlas usando Kriging. Existe una gran cantidad de métodos que generan simulaciones Gaussianas, por lo que éste es otro punto a tener en cuenta. Y por último veremos cuales métodos requieren de mallas regulares.

En la tabla anterior se resumen las tres características antes mencionadas de los métodos de simulación que revisaremos a continuación con cierto detalle.

6.7 Métodos del tipo Gaussiano

En el caso de los métodos del tipo Gaussiano haremos un paréntesis, y nos detendremos brevemente en ellos antes de abordar en detalle cada uno de los métodos en específico ya que requieren de un tratamiento especial.

Estos métodos requieren que la FDP multivariada de la función aleatoria $Z(\underline{x})$ a simular sea Gaussiana.

Como es conocido la mayoría de los fenómenos de ciencias de la tierra no presentan histogramas simétricos y mucho menos gaussianos. Por lo que nos enfrentamos aquí con la primera dificultad a la hora de aplicar esta clase de métodos.

La primera condición necesaria para que una función aleatoria posea una distribución normal multivariada es que su distribución univariada sea normal. Esto nos dice que necesitamos transformar a la función aleatoria $Z(\underline{x})$ de manera que resulte su FDP normal.

El modo de hacerlo es relativamente simple. Si $Y(\underline{x})$ es una función aleatoria con FDP univariada $F_Y(y) = G(y)$ normal estandarizada $N(0,1)$. Entonces, se cumple que

$$F_Y(y_p) = F_Z(z_p) = p; \quad \forall p \in [0,1] \quad (6.11)$$

y por lo tanto la transformación que se requiere sería

$$y = F_Y^{-1}(F_Z(z)) \quad (6.12)$$

En la práctica los n datos de la muestra de $Z(\underline{x})$ son ordenados de modo creciente de sus valores:

$$z^{(1)} \leq z^{(2)} \leq \dots \leq z^{(n)} \quad (6.13)$$

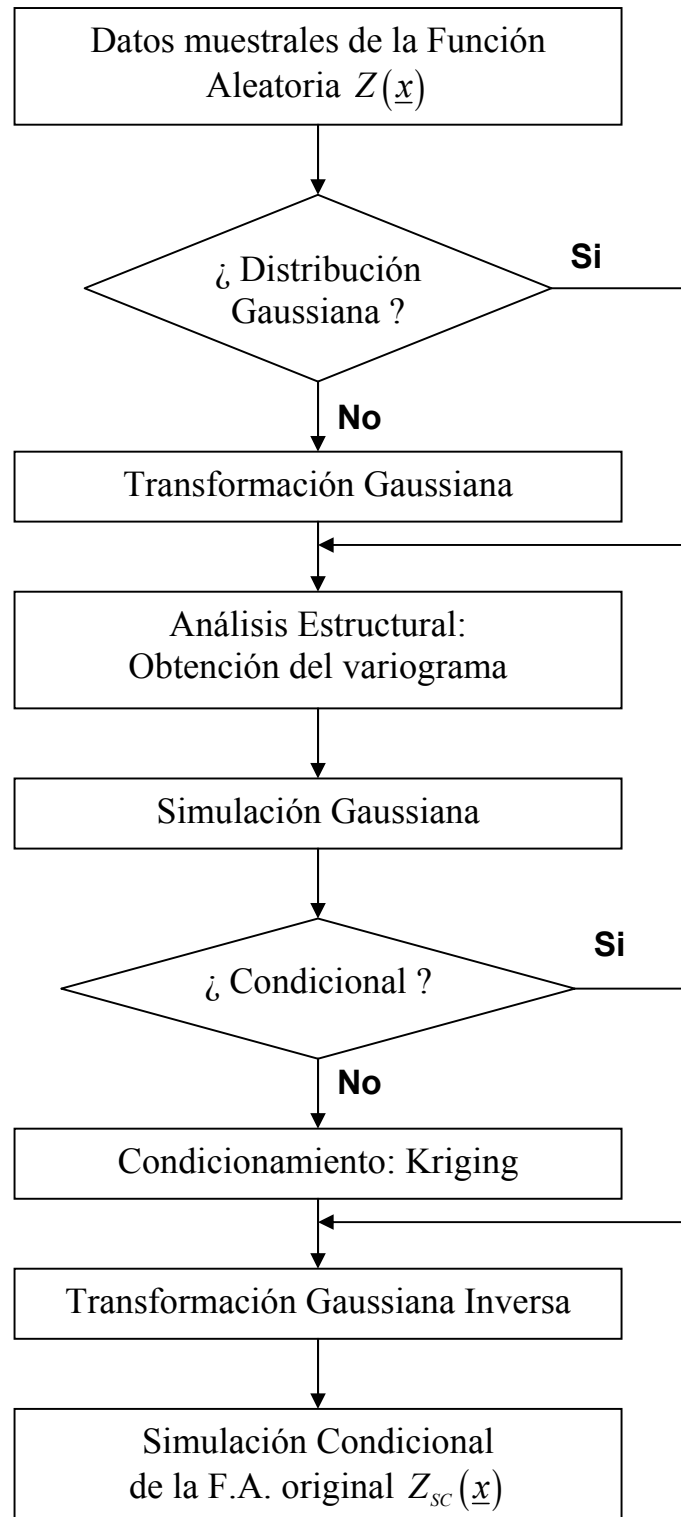


Fig. 6.1: Esquema general de las Simulaciones de tipo Gaussianas

La FDP acumulativa de $Z(\underline{x})$ está dada por $F_Z(z^{(k)}) = k/n$, entonces la transformación correspondiente sería

$$y^{(k)} = G^{-1}(k/n) \quad (6.14)$$

A este tipo de transformación se le conoce como anamorfosis.

Pero esto aún sería insuficiente puesto que se requeriría verificar la normalidad al menos de la distribución bivariada. Aunque en muchos casos para fines prácticos no se lleva tan lejos el análisis y se toma la decisión de considerar a la distribución gaussiana o por el contrario se rechaza esa hipótesis y se elige otro método de simulación no gaussiano.

Luego de realizar la simulación Gaussiana a los datos transformados se requiere hacer la transformación inversa de los valores simulados obtenidos. En la Fig. 6.1 se muestra un esquema general de cómo proceder con las simulaciones de tipo Gaussianas.

6.8 Método matricial

El *método matricial* consiste en generar variables aleatorias cuya matriz de covarianza coincide con la obtenida a partir del kriging. Es sencillo de implementar pero resulta muy costoso al incrementarse la cantidad de puntos a simular ya que el tiempo de procesamiento es superior al cuadrado del número de puntos.

Supongamos que se desea obtener realizaciones de una función aleatoria $Z(\underline{x})$ en n puntos prefijados, es decir, necesitamos obtener realizaciones estadísticamente equivalentes del vector aleatorio $\underline{Z} = \{Z(\underline{x}_1), Z(\underline{x}_2), \dots, Z(\underline{x}_n)\}^T$.

Supongamos además, que son conocidos

- i.) El valor medio esperado de $Z(\underline{x})$

$$m(\underline{x}) = E[Z(\underline{x})] \quad (6.15)$$

ii.) Y la función de covarianzas de cualquier par de puntos $Z(\underline{x}_i)$ y $Z(\underline{x}_j)$ definida como:

$$C(\underline{x}_i, \underline{x}_j) = E\left[\{Z(\underline{x}_i) - m(\underline{x}_i)\}\{Z(\underline{x}_j) - m(\underline{x}_j)\}\right] \quad (6.16)$$

Entonces el vector de los valores esperados de \underline{Z} se puede expresar como:

$$E[\underline{Z}] = \{m(\underline{x}_1), m(\underline{x}_2), \dots, m(\underline{x}_n)\}^T \equiv \underline{m} \quad (6.17)$$

y correspondientemente la matriz de covarianzas

$$Var[\underline{Z}] = \{C(\underline{x}_i, \underline{x}_j)\} \equiv \underline{\underline{C}} \quad (6.18)$$

donde $\underline{\underline{C}}$ es una matriz $n \times n$ simétrica y positiva definida.

Dada una matriz $\underline{\underline{M}}$ que satisfaga

$$\underline{\underline{M}}\underline{\underline{M}}^T = \underline{\underline{C}} \quad (6.19)$$

Entonces el vector aleatorio \underline{Z} puede ser simulado usando la relación:

$$\underline{Z}_S = \underline{m} + \underline{\underline{M}}\underline{\varepsilon} \quad (6.20)$$

donde $\underline{\varepsilon} \equiv \{\varepsilon(\underline{x}_1), \varepsilon(\underline{x}_2), \dots, \varepsilon(\underline{x}_n)\}^T$ es un vector de variables aleatorias no correlacionadas, con valor medio cero y varianza unitaria, es decir:

$$E[\underline{\varepsilon}] = \underline{0} \quad \text{y} \quad Var[\underline{\varepsilon}] = \underline{\underline{I}} \quad (6.21)$$

En particular, para la obtención de la matriz $\underline{\underline{M}}$, la descomposición de Cholesky resulta una buena elección ya que permite expresar eficientemente a $\underline{\underline{C}}$ como:

$$\underline{\underline{C}} = \underline{\underline{L}}\underline{\underline{L}}^T \quad (6.22)$$

donde $\underline{\underline{L}}$ es una matriz $n \times n$ triangular inferior (todos los elementos por encima de la diagonal son ceros).

Es bueno hacer notar que usualmente el vector $\underline{\varepsilon}$ se toma como realizaciones de variables aleatorias con distribución gaussiana y por lo tanto esto implica que \underline{Z}_s es una realización de una distribución gaussiana multivariada, aunque se podría tomar en calidad de $\underline{\varepsilon}$ realizaciones de variables aleatorias no correlacionadas con media cero y varianza uno de cualquier distribución de probabilidad conjunta.

En el caso estacionario los componentes de la matriz de covarianza son:

$$C_{ij} = C(\underline{x}_i, \underline{x}_j) = \sigma_{\underline{Z}}^2 - \gamma_{ij} \quad (6.23)$$

Mientras que para funciones intrínsecas se usa como alternativa equivalente:

$$C_{ij} = \gamma(\underline{x}_0 - \underline{x}_i) + \gamma(\underline{x}_0 - \underline{x}_j) - \gamma_{ij} \quad (6.24)$$

donde \underline{x}_0 es un punto arbitrario pero común a todos los C_{ij} .

Para que la simulación sea condicional basta con tomar al vector de valores esperados \underline{m} igual a la estimación por kriging \underline{Z}^* y a \underline{C} como la matriz de covarianza de dicha estimación, resultando:

$$\underline{Z}_{sc} = \underline{Z}^* + \underline{M}\underline{\varepsilon} \quad (6.25)$$

Este método clásico fue introducido en aplicaciones geoestadísticas por Davis (1987), y es aplicable hasta límite que sea factible la descomposición de Cholesky, es decir, hasta unos cuanto cientos de puntos simulados.

6.9 Método espectral

Los *métodos espectrales* hacen uso de la representación espectral de la función de autocovarianza. Existen diversas variantes. Una se basa en aproximar la función aleatoria mediante serie de cosenos en los que la frecuencia se muestrea aleatoriamente a partir de la función de densidad espectral. Otra se considera como una integración numérica en la que se ponderan las amplitudes y se distribuyen uniformemente las frecuencias. A continuación veremos con algún detalle el primer enfoque.

Supongamos que tenemos una función aleatoria $Z(x)$ en una dimensión y estacionaria de segundo orden con:

i.) valor esperado nulo

$$E[Z(x)] = 0; \quad \forall x \quad (6.26)$$

ii.) función de covarianzas $C(h)$ y por lo tanto la varianza existe, es finita y no depende de x , es decir

$$\sigma^2 = C(0) \quad (6.27)$$

Ha sido demostrado (Bochner, 1955) que cualquier función positiva definida y simétrica $C(h)$ corresponde a la función de covarianzas de una función aleatoria $Z(x)$ estacionaria de segundo orden, más aun $C(h)$ es una función positiva definida si y solo si existe su representación espectral siguiente

$$C(h) = \int_{-\infty}^{+\infty} \cos(\omega h) dF(\omega) \quad (6.28)$$

donde $F(\omega)$ es una función acotada positiva y simétrica; y a $F(\omega)/C(0)$ se le llama función de distribución espectral, ya que $F(\bullet)/C(0)$ es una función de distribución de probabilidad.

Cuando se cumple que $dF(\omega) = f(\omega)d\omega$, lo cual es posible si $\int_{-\infty}^{+\infty} |C(h)| dh < \infty$, entonces $f(\omega)/C(0)$ está definida y se le conoce como función de densidad espectral ya que $\int_{-\infty}^{+\infty} f(\omega)/C(0) d\omega = 1$ por lo que $f(\bullet)/C(0)$ es una función de densidad de probabilidad.

Shinozuka (1971), y Mejia y Rodríguez-Iturbe (1974) propusieron la siguiente fórmula para simular una función aleatoria

$$Z_N(x) \equiv \sigma(2/N)^{1/2} \sum_{i=1}^N \cos(\omega_i x + \phi_i) \quad (6.29)$$

donde ω_i , $i = 1, \dots, N$, son realizaciones independientes de una variable aleatoria con densidad de probabilidad $f(\bullet)/C(0)$, independiente de ϕ_i , $i = 1, \dots, N$ que son realizaciones independientes con distribución uniforme en el intervalo $[-\pi, \pi]$.

Se puede comprobar que $Z_N(x)$ es una función aleatoria, independientemente del valor de N , que tiene valor medio cero

$$E[Z_N(x)] = \frac{(2N)^{1/2}}{2\pi\sigma} \sum_{i=1}^N \int_{-\infty}^{+\infty} \int_{-\pi}^{\pi} \cos(\omega x + \phi) f(\omega) d\phi d\omega = 0 \quad (6.30)$$

y función de covarianzas $C(h)$

$$\begin{aligned} E[Z_N(x+h)Z_N(x)] &= (2\sigma^2 / N) \sum_{k=1}^N \sum_{l=1}^N E[\cos(\omega_k x + \phi_k) \cos(\omega_l (x+h) + \phi_l)] \\ &= 2\sigma^2 E[\cos(\omega_k x + \phi_k) \cos(\omega_k (x+h) + \phi_k)] \\ &= \int_{-\infty}^{+\infty} \cos(\omega h) f(\omega) d\omega = C(h) \end{aligned} \quad (6.31)$$

A medida de que se incrementa el valor de N la función aleatoria $Z_N(x)$ converge débilmente a una función aleatoria Gaussiana $Z(x)$, es decir

$$\lim_{N \rightarrow \infty} Z_N(x) \rightarrow Z(x); \quad \forall x \in [a, b] \quad (6.32)$$

y el proceso se hace ergódico. Por ergodicidad se entiende que cualquier realización de la función aleatoria es representativa de esta, es decir, que las estimaciones muestrales de los primeros momentos de la realización convergen a los momentos de la distribución de probabilidad de la función aleatoria.

Este método se puede generalizar fácilmente para el caso de varias dimensiones cuando $Z(\underline{x})$, para $\underline{x} \in \mathbb{R}^3$ y también puede ser aplicado para simular procesos aleatorios vectoriales.

Como en general se simulan funciones aleatorias $Z(\underline{x})$ con valor esperado no nulo, $E[Z(\underline{x})] = m(\underline{x})$, luego de simular el proceso estacionario con media cero se requiere agregar a los valores simulados el valor medio esperado $m(\underline{x})$ de $Z(\underline{x})$.

El método espectral tiene como ventaja que permite producir simulaciones en cualquier punto dada la función de covarianzas $C(h)$ siempre que se pueda calcular su

transformación de Fourier. Pero su desventaja principal es que requiere de un valor elevado de N (del orden de miles) para que alcance la ergodicidad.

6.10 Método de bandas rotantes

El *método de las bandas rotantes* consiste en generar procesos unidimensionales a lo largo de líneas distribuidas uniformemente en el espacio. La simulación en un punto arbitrario resulta de la suma de las proyecciones de este punto sobre las líneas. Es muy económico puesto que sólo requiere de simulaciones unidimensionales. Su dificultad fundamental estriba en la obtención de la función de covarianza unidimensional especialmente para variables bidimensionales.

Muchos métodos han sido propuestos para la simulación de un proceso estocástico estacionario en una dimensión a partir de la generación de realizaciones de su función aleatoria $Z(x)$ dada su función de covarianzas. Sin embargo cuando se trata de extender estos procedimientos a dos o más dimensiones se tornan en demasiado complicados y en general prohibitivos debido al tiempo de computo que consumen.

La originalidad del método conocido como el de “bandas rotantes” se debe a G. Matheron, el cual consiste en reducir el problema de obtener una simulación en tres dimensiones a varias simulaciones independientes en una sola dimensión a lo largo de ciertas líneas rotadas en el espacio tridimensional \mathbb{R}^3 . Por lo que este método produce simulaciones n -dimensionales a un costo de computo razonable, equivalente de hecho al costo de las simulaciones unidimensionales.

Considere una línea L_1 en el espacio tridimensional \mathbb{R}^3 y una función aleatoria unidimensional, estacionaria de segundo orden $Y(x_{L_1})$ definida sobre dicha línea L_1 , con valor esperado $E[Y(x_{L_1})] = 0$ y función de covarianzas unidimensional $C^{(1)}(h_{L_1})$.

Sea x_{L_1} la proyección de cualquier punto $\underline{x} \in \mathbb{R}^3$ sobre la línea L_1 , entonces se define la siguiente función aleatoria $Z_1(\underline{x})$ en tres dimensiones como:

$$Z_1(\underline{x}) \equiv Y(x_{L_1}); \quad \forall \underline{x} \in \mathbb{R}^3 \quad (6.33)$$

Esta función aleatoria es estacionaria de segundo orden, con valor esperado cero y función de covarianzas tridimensional

$$E[Z_1(\underline{x})Z_1(\underline{x}+\underline{h})] = E[Y(x_{L_1})Y(x_{L_1}+h_{L_1})] = C^{(1)}(h_{L_1}) \quad (6.34)$$

donde h_{L_1} es la proyección del vector \underline{h} sobre la línea L_1 .

En la práctica para producir una realización $z_1(\underline{x})$ de la función aleatoria $Z_1(\underline{x})$, se procede asignando el valor simulado $y(x_{L_1})$ en el punto x_{L_1} de la línea L_1 a todos los puntos que se encuentran dentro de una banda centrada en el plano $x_{L_1} = \text{const.}$, perpendicular a la línea L_1 . El ancho de esta banda es el paso o intervalo de separación Δx_{L_1} entre los puntos simulados sobre la línea L_1 .

Si se consideran N líneas L_1, \dots, L_N distribuidas uniformemente en N direcciones en \mathbb{R}^3 .

Si en cada línea L_i se genera una realización $y(x_{L_i})$ de una función aleatoria unidimensional independiente $Y(x_{L_i})$ e isomorfa a $Y(x_{L_1})$, $\{Y(x_{L_i}), i = 1, \dots, N\}$, entonces se puede definir de manera análoga realizaciones tridimensionales a partir de las simulaciones unidimensionales:

$$z_i(\underline{x}) \equiv y(x_{L_i}); \quad \forall \underline{x} \in \mathbb{R}^3 \quad (6.35)$$

Y finalmente a cada punto del espacio tridimensional $\underline{x} \in \mathbb{R}^3$ se le asigna el siguiente valor a partir de las simulaciones unidimensionales

$$z_s(\underline{x}) = \frac{1}{\sqrt{N}} \sum_{i=1}^N z_i(\underline{x}); \quad \forall \underline{x} \in \mathbb{R}^3 \quad (6.36)$$

Esta realización de la función aleatoria tridimensional es estacionaria de segundo orden con valor medio cero y función de covarianzas:

$$E[Z_s(\underline{x})Z_s(\underline{x}+\underline{h})] = C_s(\underline{h}) \quad (6.37)$$

la cual, cuando N tiende al ∞ , tiende a la función de covarianzas isotrópica $C(\underline{h})$.

Para fines prácticos la función de covarianzas tridimensional $C^{(3)}(\underline{h})$ se considera conocida y la función de covarianzas unidimensional $C^{(1)}(h_{L_1})$ de la función aleatoria a ser simulada en cada una de las N líneas se obtiene a partir de la primera usando la expresión

$$C^{(1)}(r) = \frac{d}{dr} \{r C^{(3)}(r)\} \quad (6.38)$$

Sin embargo no existe una expresión directa para obtener la $C^{(1)}(\underline{h})$ a partir de la función de covarianzas bidimensional $C^{(2)}(\underline{h})$.

Es uno de los métodos mas eficientes para producir simulaciones en varias dimensiones puesto que reduce su complejidad al orden de las simulaciones en una dimensión. Como dificultades prácticas se le debe apuntar que requiere del conocimiento o del cálculo de $C^{(1)}(h)$ y de un algoritmo de simulación en una dimensión que reproduzca esta función de covarianzas. Usualmente se emplea el método espectral.

6.11 Métodos secuenciales

En estos métodos un nuevo valor simulado $Z_s(\underline{x})$ se obtiene a partir de la función de distribución de probabilidad estimada usando los valores observacionales (reales) y los valores previamente simulados en etapas anteriores en una vecindad del punto \underline{x} .

En dependencia de cómo se estime la función distribución de probabilidad, existen dos métodos secuenciales:

- a) Secuencial Indicador:** Usa el Kriging indicador para estimar la función distribución de probabilidad local. Requiere del modelo del semivariograma para cada valor de corte z_c especificado por el usuario o como alternativa mas eficiente pero menos precisa del semivariograma obtenido para el valor corte correspondiente a la mediana z_M .

b) Secuencial Gaussiano: Se obtiene la estimación usando Kriging ordinario aplicado a los datos estandarizados, es decir $Y(\underline{x}_i) = (Z(\underline{x}_i) - m_Z) / \sigma_Z$, $i = 1, \dots, n$. Se requiere del modelo del semivariograma basado en los datos transformados y cada vez que se obtienen los valores estimado son transformados hacia atrás a su escala original. Su dificultad principal estriba en el grado de conocimiento que se posea de la función distribución de probabilidad, lo cual puede ser suplido en parte por la experiencia previa sobre la naturaleza del fenómeno que se simula. Hay programas, como el de GSLIB que permite extrapolar valores de la FDP, usando modelos potencia, hiperbólico y lineal.

6.12 Método Gaussiano Truncado

Consideremos un indicador o una serie de indicadores que se originan aplicando uno o más umbrales a una F.A. Gaussiana $Y(\underline{x})$,

$$I_i(\underline{x}) = 1_{y_{i-1} < Y(\underline{x}) < y_i}, \text{ donde } -\infty = y_0 < y_1 < \dots < y_m = +\infty \quad (6.39)$$

Los umbrales y_i son elegidos de manera que concuerden con las proporciones p_i de los indicadores

$$y_i = G^{-1} \left(\sum_{j=1}^i p_j \right) \quad (6.40)$$

Una vez que el correlograma de la F.A. Gaussiana $Y(\underline{x})$ es conocido, las funciones de covarianzas directa y la cruzada de los indicadores son conocidos. La simulación de los $I_i(\underline{x})$ se reduce a la simulación de $Y(\underline{x})$. En aplicaciones se elige directamente $\rho(h)$ de manera que los variogramas teóricos de los indicadores deducidos a partir de las relaciones (6.41) o (6.42) se ajusten bien a los variogramas muestrales.

$$C(h; z) = \frac{1}{2\pi} \int_0^{\rho(h)} \exp\left(-\frac{z^2}{1+u}\right) \frac{du}{\sqrt{1-u^2}} \quad (6.41)$$

o

$$C_{yy'}(h) = g(y)g(y') \sum_{n=1}^{\infty} \frac{\chi_{n-1}(y)\chi_{n-1}(y')}{n} \rho(h)^n \quad (6.42)$$

donde $\chi_n(y) = H_n(y)/\sqrt{n!}$ es el polinomio de Hermite normalizado el cual forma una base ortonormal con respecto a la distribución gaussiana estándar.

En efecto, estas relaciones no son invertidas fácilmente, excepto si el valor de corte es en $y=0$. Mas aún, mientras cualquier correlograma $\rho(h)$ produce un modelo válido de covarianza indicador mediante la aplicación de (6.41), no es cierto lo inverso, es decir, que la inversión de esta relación sea un modelo válido de correlograma, incluso si partimos de una covarianza indicador, ya que ésta no es necesariamente la covarianza de una F.A. Gaussiana truncada.

Examinemos la construcción de las simulaciones condicionales. Si los datos provienen directamente de $Y(\underline{x})$, estamos en el caso estándar de la sección 6.5, en caso contrario los datos son de los indicadores y se procede en cuatro pasos:

- 1.- Análisis estructural global de los variogramas directos y cruzados de los diferentes indicadores, ajustando un modelo derivado a partir del correlograma $\rho(h)$.
- 2.- Simulación de $Y(\underline{x})$ en los puntos muestrales condicionada en los valores del indicador.
- 3.- Simulación de toda la malla condicionada en estos valores duros simulados.
- 4.- Transformación de la simulación de $Y(\underline{x})$ a la simulación de los indicadores.

Este método puede ser fácilmente adaptado a la regionalización de las proporciones de diferentes indicadores y de los correspondientes umbrales. Es aplicado para la simulación de litofacies. Esto conduce a simulaciones de tipo difusivas, en el sentido que la facies i está rodeada sólo por facies $i-1$ y $i+1$. Pero el método puede ser generalizado usando dos F.A. Simuladas Gaussianas independientes o correlacionadas de manera que las facies no sigan una a la otra en un orden fijo, el cual se conoce como *método de simulaciones plurigaussianas truncadas*.

6.13 Simulaciones Plurigaussianas Truncadas

La definición general de una plurigaussiana truncada está basada en una partición $D_i, i=1, \dots, n$ del espacio p -dimensional, es decir, $D_i \cap D_k = \emptyset$ y $\bigcup_i D_i = \mathbb{R}^p$, y en un vector de p gaussianas:

$$\underline{Y}(\underline{x}) = (Y_1(\underline{x}), Y_2(\underline{x}), \dots, Y_p(\underline{x})) \quad (6.43)$$

donde \underline{x} son las coordenadas de una ubicación en 3D.

6.14 Métodos del tipo Recocido Simulado

El recocido simulado (simulated annealing) es un método heurístico que permite la generación de simulaciones condicionales sujetas a restricciones complejas. Fue introducido por Kirkpatrick (1983) como un método para la solución de problemas combinatorios (discretos) de optimización. Y su nombre proviene por la analogía del método con el proceso de enfriamiento lento del metal fundido de manera que alcance un estado de mínima energía el cual corresponde a una estructura cristalina pura completamente ordenada del metal. Si el enfriamiento no se realiza con la lentitud adecuada, el metal no alcanza dicho estado y degenera en una estructura amorfa o policristalina.

El método computacional inicia con una imagen aleatoria y requiere de los datos observados y del modelo del variograma a reproducir.

Durante su ejecución el algoritmo intenta modificar la imagen inicial de manera tal que reproduzca el variograma y al mismo tiempo mantenga la misma distribución de probabilidad de los datos observados.

Esquema general del algoritmo es:

- 1) Crea imagen inicial asignando a cada nodo de la malla de simulación valores de manera aleatoria generados a partir de la función distribución de probabilidad (FDP) suministrada por el usuario. Cuando el nodo de la malla de simulación coincide con el nodo de una observación o es cercano a éste con cierto error de tolerancia, se le asigna el valor observado a éste. Este imagen inicial así obtenida no tiene que reflejar ningún modelo de variograma específico.
- 2) En una segunda etapa el algoritmo intenta reproducir el variograma modelo intercambiando valores en los nodos de la malla de simulación elegidos de manera aleatoria. Se aceptan como permanentes los intercambios que reduzcan la diferencia entre el variograma suministrado $\gamma(\underline{h})$ y el que refleja los datos simulados $\gamma^*(\underline{h})$. Es decir, se intenta minimizar la siguiente función objetivo

$$FO_1 = \sum_i \left[\frac{\gamma^*(\underline{h}_i) - \gamma(\underline{h}_i)}{\gamma(\underline{h}_i)} \right]^2 \quad (6.44)$$

Para evitar caer en un mínimo local, son aceptados algunos intercambios que incrementan el valor de la función objetivo, es decir la aceptación del intercambio es probabilística. Se espera que para un número elevado de iteraciones la función objetivo tienda al mínimo global y consecuentemente los valores simulados posean una estructura espacial cercana al variograma de los datos observados.

Una alternativa a la función objetivo arriba propuesta puede incluir restricciones y tener la siguiente forma:

$$FO_2 = w_1 \sum_i [\gamma^*(\underline{h}_i) - \gamma(\underline{h}_i)]^2 + w_2 \sum_j [\mu_j^* - \mu_j]^2 \quad (6.45)$$

donde el segundo término es una medida de la discrepancia de cierta propiedad, mientras que w_1 y w_2 le dan un peso relativo a cada uno de los términos.

Desde el punto de vista computacional es un método muy demandante ya que requiere de millones de iteraciones para que la simulación refleje correctamente la estructura de la función de semivarianzas.

ANEXO A: CONCEPTOS BÁSICOS DE ESTADÍSTICA

A.1 ESTADÍSTICA UNIVARIADA.

- **Variable Aleatoria (V.A.):** Es una variable Z (propiedad) que puede tomar una serie de valores o realizaciones $(z_i, i=1,...,N)$ con un conjunto dado de probabilidad de ocurrencia $(p_i, i=1,...,N)$
- **Variable Aleatoria Discreta:** cuando el número de ocurrencias es finito, se conoce como variable aleatoria discreta. Las N probabilidades de las ocurrencias deben satisfacer las condiciones siguientes:
 - 1.- $p_i \geq 0, \quad \forall i = 1, \dots, N$
 - 2.- $\sum_{i=1}^N p_i = 1$
- **Variable Aleatoria Continua:** Si el número de ocurrencias posibles es infinito, por ejemplo el valor de la porosidad del medio se encuentra en el intervalo $[0,100\%]$, entonces se dice que la **VA** Z es continua.
- **Función de Distribución de Probabilidad (FDP)**

- 1.- **VA** discreta, se conoce como histograma acumulativo

$$F(z_i) = \sum_{z_j \leq z_i} p_j \in [0,1] \quad (\text{A.1})$$

- 2.- **VA** continua, se conoce como función de distribución de probabilidad (acumulativa) y se define como:

$$F(z) = \Pr\{Z \leq z\} \in [0,1] \quad (\text{A.2})$$

La **FDP** caracteriza completamente a la **VA**.

Propiedades de una FDP:

- 1.- $F(-\infty) = 0$

$$2.- F(+\infty) = 1$$

$$3.- \Pr\{Z > z\} = 1 - F(z)$$

$$4.- F(z_1) \geq F(z_2), \quad \forall z_1 \geq z_2, \text{ es decir es una función no decreciente.}$$

- **Función de Densidad de Probabilidad (fdp).**

1.- **VA** continua, se define como:

$$f(z) = \frac{dF(z)}{dz} \quad (\text{A.3})$$

cuando esta existe.

2.- **VA** discreta es el histograma.

Propiedades de una fdp:

$$1.- f(-\infty) = f(+\infty) = 0$$

$$2.- \int_{-\infty}^{+\infty} f(z) dz = 1$$

$$3.- F(z_1) = \int_{-\infty}^{z_1} f(z) dz$$

$$4.- \Pr\{a \leq Z \leq b\} = \int_a^b f(z) dz$$

$$5.- f(z) \geq 0$$

- **Percentiles o cuantiles de una distribución $F(z)$.**

El percentil de una distribución $F(z)$ es el valor z_p de la **V.A.** tal que:

$$F(z_p) = p \in [0, 1] \quad (\text{A.4})$$

Si existe la función inversa se puede expresar como:

$$z_p = F^{-1}(p) \quad (\text{A.5})$$

donde p es expresada en por ciento.

Algunos cuantiles de interés:

Mediana, $p = 0.5$ (50 %)

$$M = F^{-1}(0.5) \quad (\text{A.6})$$

Cuartiles

$p=0.25$ (primer cuartil o inferior)

$$z_{0.25} = F^{-1}(0.25) \quad (\text{A.7})$$

$p=0.75$ (tercer cuartil o superior)

$$z_{0.75} = F^{-1}(0.75) \quad (\text{A.8})$$

Rango o intervalo intercuartil (IR)

$$[z_{0.25}, z_{0.75}] \quad (\text{A.9})$$

- Valor Esperado o esperanza matemática de una VA.

Es el valor más probable que puede tomar la VA. Se conoce también como valor medio o media.

1.- VA discreta

$$m = E[Z] = \sum_{i=1}^N p_i z_i \quad (\text{A.10})$$

2.- VA continua

$$m = E[Z] = \int_{-\infty}^{+\infty} z dF(z) = \int_{-\infty}^{+\infty} z f(z) dz \quad (\text{A.11})$$

De manera análoga se puede definir el valor esperado de cualquier función de una VA. Por ejemplo si tenemos la función $\varphi(Z)$ de la VA Z , entonces su valor esperado se define como:

1.- VA discreta

$$E[\varphi(Z)] = \sum_{i=1}^N p_i \varphi(z_i) \quad (\text{A.12})$$

2.- VA continua

$$E[\varphi(Z)] = \int_{-\infty}^{+\infty} \varphi(z) dF(z) = \int_{-\infty}^{+\infty} \varphi(z) f(z) dz \quad (\text{A.13})$$

Propiedades lineales del valor esperado:

Es un operador lineal, que cumple:

1.- Para cualquier a_k – constantes y Z_k – VAs

$$E\left[\sum_k a_k Z_k\right] = \sum_k a_k E\{Z_k\} \quad (\text{A.14})$$

2.- Para cualquier par de funciones $\varphi_1(Z)$ y $\varphi_2(Z)$ de la VA Z

$$E[\varphi_1(Z) + \varphi_2(Z)] = E[\varphi_1(Z)] + E[\varphi_2(Z)] \quad (\text{A.15})$$

- Momento de orden r de una Función de Distribución de Probabilidad

$$m_r = E[Z^r] = \int_{-\infty}^{+\infty} z^r dF(z) = \int_{-\infty}^{+\infty} z^r f(z) dz \quad (\text{A.16})$$

- Momento central de orden r de una Función de Distribución de Probabilidad

$$\mu_r = E[(Z - m)^r] = \int_{-\infty}^{+\infty} (z - m)^r dF(z) = \int_{-\infty}^{+\infty} (z - m)^r f(z) dz \quad (\text{A.17})$$

- Varianza σ^2 de una VA (2^{do} momento central)

Se define como el valor esperado de la desviación cuadrática de la VA Z respecto a su valor medio m . Y caracteriza la dispersión de la distribución alrededor de la media.

$$\sigma^2 = \text{Var}[Z] = E[(Z - m)^2] \geq 0 \quad (\text{A.18})$$

$$\begin{aligned} \sigma^2 &= E[(Z - m)^2] = E[Z^2 - 2mZ + m^2] \\ &= E[Z^2] - 2mE[Z] + m^2 = E[Z^2] - m^2 \end{aligned} \quad (\text{A.19})$$

1.- VA discreta

$$\text{Var}[Z] = \sum_{i=1}^N p_i (z_i - m)^2 \quad (\text{A.20})$$

2.- VA continua

$$\text{Var}[Z] = \int_{-\infty}^{+\infty} (z - m)^2 dF(z) = \int_{-\infty}^{+\infty} (z - m)^2 f(z) dz \quad (\text{A.21})$$

- **Desviación Estándar σ**

$$\sigma = \sqrt{\text{Var}[Z]} \quad (\text{A.22})$$

- **Coefficiente de variación (dispersión relativa)**

$$CV = \sigma / m \quad (\text{A.23})$$

- **Desviación absoluta media alrededor de la media**

$$mAD = E[|Z - m|] \quad (\text{A.24})$$

- **Desviación absoluta media alrededor de la mediana**

$$MAD = E\{|Z - M|\} \quad (\text{A.25})$$

- **Signo de Simetría**

$$\text{Sign}(m - M) = \begin{cases} \text{cuando } > 0, \Rightarrow m > M \text{ y se dice que hay asimetría positiva} \\ \text{cuando } < 0, \Rightarrow m < M \text{ y se dice que hay asimetría negativa} \end{cases} \quad (\text{A.26})$$

Ver figura 1.

- **Coefficiente de simetría (medida de la simetría)**

$$\alpha_1 = \frac{\mu_3}{\mu_2^{3/2}} \quad (\text{A.27})$$

- **Coefficiente de curtosis (medida del achatamiento)**

$$\alpha_2 = \frac{\mu_4}{\mu_2^2} - 3 \quad (\text{A.28})$$

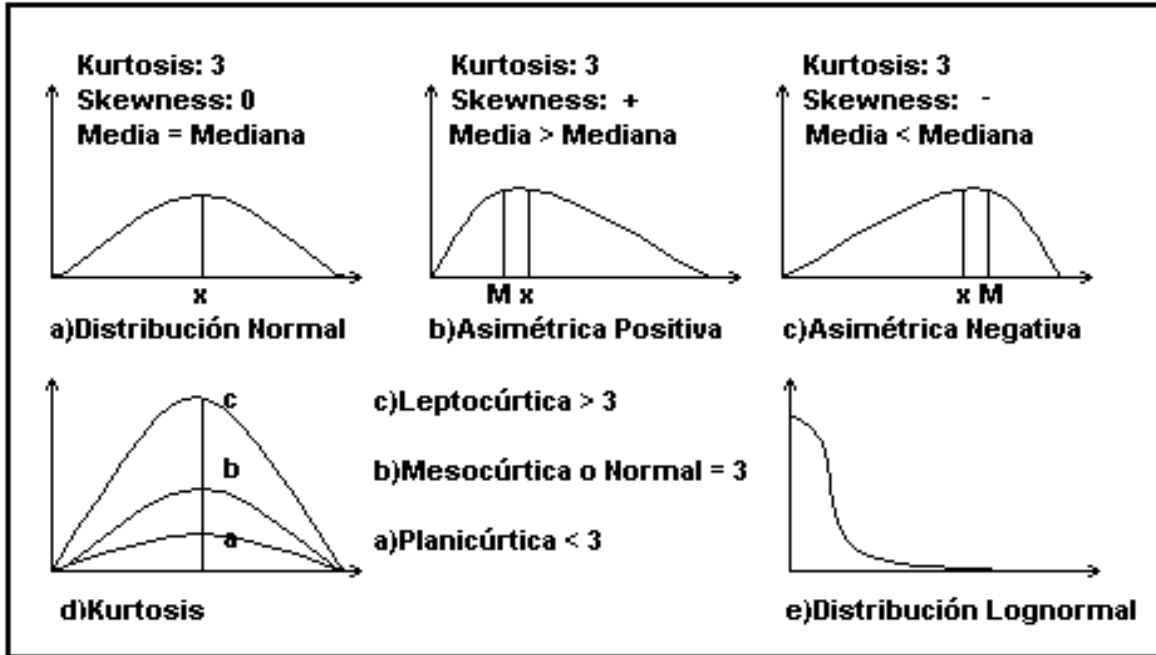


Figura 1. -Diferentes tipos de distribuciones

- Estimadores muestrales

- Valor medio(media)

El promedio es un estimador insesgado del valor medio de la población

$$m^* = \bar{z} = \frac{1}{N} \sum_{i=1}^N z_i \quad (\text{A.29})$$

- Varianza

$$(\sigma^2)^* = S^2 = \frac{1}{N-1} \sum_{i=1}^N (z_i - \bar{z})^2 \quad (\text{A.30})$$

-Distribución Normal o Gaussiana. Características.

Esta distribución está completamente caracterizada por sus dos parámetros: media m y varianza σ^2 y se designa mediante $N(m, \sigma^2)$.

La *fdp* normal o Gaussina está dada por

$$g(z) = \frac{1}{\sigma \sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{z-m}{\sigma} \right)^2 \right] \quad (\text{A.31})$$

Para el caso de la *fdp* estándar los parámetros son: media $m = 0$ y varianza $\sigma^2 = 1$, es decir $N(0,1)$, entonces

$$g_0(z) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{z^2}{2}\right] \quad (\text{A.32})$$

La correspondiente FDP no tiene una expresión analítica, pero está bien tabulada:

$$G_0(z) = \int_{-\infty}^z g_0(u) du \quad (\text{A.33})$$

y correspondientemente

$$G(z) = \int_{-\infty}^z g(u) du = G_0\left(\frac{z-m}{\sigma}\right) \quad (\text{A.34})$$

La distribución gaussiana es simétrica respecto a la media, es decir se cumple que:

$$m = M; \quad g(m+z) = g(m-z); \quad G(m-z) = 1 - G(m+z); \quad (\text{A.35})$$

$$z_{1-p} = 2m - z_p; \quad p \in [0, 0.5] \quad (\text{A.36})$$

Una de las razones de la popularidad de la distribución gaussiana se debe a los teoremas del límite central que dice:

Si n VAs Z_i independientes tienen la misma FDP y media cero, entonces la variable aleatoria

$$Y = \frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i \text{ tiende a tener una distribución Normal cuando } n \rightarrow \infty \quad (\text{A.37})$$

-Modelo LogNormal. Características.

Muchas distribuciones experimentales tienden a ser asimétricas, con media y medianas diferentes, y la mayoría de las variables en ciencias de la tierra toman valores no negativos. Varias transformaciones del modelo normal han sido definidas para acomodarse a las características de las distribuciones de los datos en ciencias de la tierra.

Una VA positiva Y se dice que tiene una distribución lognormal si su logaritmo $X = \ln(Y)$ está normalmente distribuido.

$$Y > 0 \rightarrow \log N(m, \sigma^2), \text{ si } X = \ln Y \rightarrow N(\alpha, \beta^2) \quad (\text{A.38})$$

También la distribución Lognormal está completamente caracterizada por dos parámetros:

- 1.- media y varianza (m, σ^2) parámetros aritméticos
- 2.- media y varianza (α, β^2) de la transformación logarítmica $X = \ln(Y)$, parámetros logarítmicos.

La FDP se expresa más fácilmente a través de los parámetros logarítmicos (α, β^2) .

$$\Pr\{Y \leq y\} = F_Y(y) = G_0\left(\frac{\ln y - \alpha}{\beta}\right) \quad (\text{A.39})$$

Y su correspondiente fdp es

$$f_Y(y) = F'_Y(y) = \frac{1}{\beta y} g_0\left(\frac{\ln y - \alpha}{\beta}\right) \quad (\text{A.40})$$

Relación entre los parámetros aritméticos y logarítmicos:

$$\left\{ \begin{array}{l} m = e^{\alpha + \beta^2/2} \\ \sigma^2 = m^2 [e^{\beta^2} - 1] \end{array} \right\} \leftrightarrow \left\{ \begin{array}{l} \alpha = \ln m - \frac{\beta^2}{2} \\ \beta^2 = \ln \left(1 + \frac{\sigma^2}{m^2} \right) \end{array} \right\} \quad (\text{A.41})$$

La fdp lognormal tiene asimetría positiva.

Un corolario del teorema del límite central dice:

El producto de VA independientes con igual fdp $Y = \prod_{i=1}^n Y_i$ tiende a estar lognormalmente

distribuidas cuando n es grande. Usando el teorema del límite central, tenemos que

$$Y = \prod_{i=1}^n Y_i \Rightarrow \ln Y = \sum_{i=1}^n \ln Y_i \rightarrow \text{Normal cuando } n \rightarrow \infty \quad (\text{A.42})$$

entonces $Y \rightarrow \text{Lognormal}$, cuando $n \rightarrow \infty$.

Métodos Clásicos de Estimación

Un estimador de un parámetro de una población puede ser dado por

- **Estimador Puntual:** Un estimador puntual θ^* de un parámetro θ de una población es un valor simple o realización de una variable aleatoria Θ^* , es decir los valores que pueden tomar el estimador define una variable aleatoria con cierta distribución de probabilidad asociada. Ej: El valor \bar{z} del estadígrafo promedio \bar{Z} calculado para n valores muestrales, es un estimador puntual del parámetro valor medio m de una población.
- **Intervalo de Estimación:** Un intervalo de estimación de un parámetro θ es un intervalo de la forma $a < \theta < b$, donde a y b dependen del estimador puntual θ^* para una muestra dada y también de la distribución muestral del estadígrafo $\hat{\Theta}$.

Def: Un estadígrafo $\hat{\Theta}$ se dice que es un estimador insesgado de un parámetro θ de una población si se cumple que

$$m_{\hat{\Theta}} = E[\hat{\Theta}] = E[\theta] = m_{\theta} \quad (\text{A.43})$$

o dicho de otra manera si el valor esperado del error de la estimación es cero

$$E[e_{\hat{\Theta}}] = E[\hat{\Theta} - \theta] = E[\hat{\Theta}] - E[\theta] = 0 \quad (\text{A.44})$$

Def: Se considera el estimador más eficiente a aquel que tiene la menor varianza de todos los estimadores insesgados posibles de un parámetro θ de una población

Ejemplo: Estimación del valor medio (m)

Estimador puntual:

$$\bar{z} = \frac{1}{N} \sum_{i=1}^N z_i \quad (\text{A.45})$$

$$E[\bar{Z}] = m \quad (\text{A.46})$$

$$\text{Var}[\bar{Z}] = \sigma^2 / N \quad (\text{A.47})$$

Intervalo de estimación:

Cuando N es suficientemente grande (>30) se puede considerar la distribución muestral del estadígrafo \bar{Z} aproximadamente como normal con media $m_{\bar{Z}}$ y desviación estándar $\sigma_{\bar{Z}} = \sigma_Z / \sqrt{N}$.

Sabemos que una VA Z_s con una distribución normal estándar se encuentra en el intervalo $[-1.96, 1.96]$ con una probabilidad de 0.95,

$$\Pr\{-1.96 < Z_s < 1.96\} = 0.95 \quad (\text{A.48})$$

$$\text{donde } Z_s = \frac{\bar{Z} - m_Z}{\sigma_Z / \sqrt{N}}$$

$$\Pr\left\{-1.96 < \frac{\bar{Z} - m_Z}{\sigma_Z / \sqrt{N}} < 1.96\right\} = 0.95 \quad (\text{A.49})$$

$$\Pr\left\{\bar{Z} - \frac{1.96\sigma_Z}{\sqrt{N}} < m_Z < \bar{Z} + \frac{1.96\sigma_Z}{\sqrt{N}}\right\} = 0.95 \quad (\text{A.50})$$

donde $\bar{Z} - \frac{1.96\sigma_Z}{\sqrt{N}} < m_Z < \bar{Z} + \frac{1.96\sigma_Z}{\sqrt{N}}$ intervalo de confianza para 95 %.

Generalizando, tenemos que un intervalo de confianza para la media con una probabilidad de $(1-\alpha)100\%$ para $N > 30$ esta dado por

$$\bar{z} - \frac{z_{\alpha/2}\sigma_Z}{\sqrt{N}} < m_Z < \bar{z} + \frac{z_{\alpha/2}\sigma_Z}{\sqrt{N}} \quad (\text{A.51})$$

donde $\Pr\{-z_{\alpha/2} < Z < z_{\alpha/2}\} = 1 - \alpha$.

En la práctica se sustituyen \bar{z} y σ_z por sus valores estimados m_z^* y $(\sigma_z)^*$ correspondientes, resultando

$$m_z^* - \frac{z_{\alpha/2} (\sigma_z)^*}{\sqrt{N}} < m_z < m_z^* + \frac{z_{\alpha/2} (\sigma_z)^*}{\sqrt{N}} \quad (\text{A.52})$$

A.2 ESTADÍSTICA MULTIVARIADA.

- Función de distribución de probabilidad (FDP) n -variada.

Dadas n variables aleatorias Z_1, Z_2, \dots, Z_n se puede definir la **función de distribución de probabilidad n -variada o conjunta** de manera análoga al caso unidimensional como:

$$F_{Z_1, Z_2, \dots, Z_n} (z_1, z_2, \dots, z_n) = \Pr [Z_1 \leq z_1, Z_2 \leq z_2, \dots, Z_n \leq z_n] \quad (\text{A.53})$$

- Función de distribución de probabilidad marginal.

La distribución de probabilidad marginal de una variable aleatoria Z_i , se define como:

$$F_{Z_i} (z_i) = \lim_{z_1, z_2, \dots, z_{i-1}, z_{i+1}, \dots, z_n \rightarrow \infty} F_{Z_1, Z_2, \dots, Z_n} (z_1, z_2, \dots, z_n) \quad (\text{A.54})$$

Si la función de distribución de probabilidad n -variada se puede expresar como el producto de las funciones de distribución de probabilidad marginales correspondientes, es decir, $F_{Z_1, Z_2, \dots, Z_n} (z_1, z_2, \dots, z_n) = F_{Z_1} (z_1) \cdot F_{Z_2} (z_2) \cdots F_{Z_n} (z_n)$, entonces se dice que las variables aleatorias son independientes.

- Función de densidad de probabilidad (fdp) n -variada.

Cuando existe una función no negativa tal que

$$F_{Z_1, Z_2, \dots, Z_n}(z_1, z_2, \dots, z_n) = \int_{-\infty}^{z_1} \int_{-\infty}^{z_2} \dots \int_{-\infty}^{z_n} f_{Z_1, Z_2, \dots, Z_n}(z_1, z_2, \dots, z_n) dz_1 dz_2 \dots dz_n \quad (\text{A.55})$$

se dice entonces que $f_{Z_1, Z_2, \dots, Z_n}(z_1, z_2, \dots, z_n)$ es la función de densidad de probabilidad n -variada o conjunta.

Cuando las n variables aleatorias Z_1, Z_2, \dots, Z_n son independientes entonces la función de densidad de probabilidad conjunta se puede expresar como el producto $f_{Z_1, Z_2, \dots, Z_n}(z_1, z_2, \dots, z_n) = f_{Z_1}(z_1) \cdot f_{Z_2}(z_2) \dots f_{Z_n}(z_n)$, donde $f_{Z_i}(z_i), i = 1, \dots, n$, es la fdp de la distribución marginal correspondiente a la variable aleatoria Z_i .

De manera análoga al caso de una variable aleatoria se pueden definir los momentos de la función de distribución n -variada. En particular, nos interesa estudiar el caso bivariado.

Hasta el momento, sólo hemos considerado en nuestros análisis y discusiones a las variables aleatorias por separado, sin que exista ninguna interrelación entre éstas. En muchos campos de aplicación y en particular, en las Ciencias de la Tierra, es frecuentemente más importante conocer el patrón de dependencia que relaciona a una variable aleatoria X con otra variable aleatoria Y . Esto es imprescindible si se desea estimar el valor de X (porosidad) a partir de los valores conocidos de Y (mediciones de pozo). Por lo que le dedicaremos especial atención al análisis conjunto de dos variables aleatorias, conocido como análisis bivariado.

Función de Distribución de Probabilidad Bivariada

La distribución conjunta de un par de variables aleatorias X y Y se caracteriza por la Función de Distribución de Probabilidad (FDP) bivariada, que se define como:

$$F_{XY}(x, y) = \Pr\{X \leq x, Y \leq y\} \in [0, 1] \quad (\text{A.56})$$

$$F_{XY}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f_{XY}(u, v) du dv \quad (\text{A.57})$$

donde $f_{XY}(x, y) = \frac{d^2 F_{XY}(x, y)}{dx dy}$ es la función de densidad de probabilidad (*fdp*) bivariada.

En la práctica se estima mediante la proporción de pares de valores de X y Y que se encuentran por debajo del umbral x, y respectivamente.

Diagrama de Dispersión (Scattergram)

El equivalente bivariado del histograma es el diagrama de dispersión o scattergram (Fig. 2), donde cada par (x_i, y_i) es un punto.

El grado de dependencia entre las dos variables aleatorias X y Y puede ser caracterizado por el diagrama de dispersión alrededor de cualquier línea de regresión.

Semivariograma

El momento de inercia del diagrama de dispersión con respecto a una línea de 45° de pendiente, permite caracterizar la carencia de dependencia, ver Fig. 2.

$$\gamma_{XY} = \frac{1}{N} \sum_{i=1}^N [d_i]^2 = \frac{1}{2N} \sum_{i=1}^N [x_i - y_i]^2 \quad (\text{A.58})$$

donde N es el número de pares (x_i, y_i) .

Este momento de inercia se conoce como semivariograma. Mientras mayor sea el valor del semivariograma más dispersos estarán los valores en el diagrama de dispersión y menor será la dependencia entre las dos variables aleatorias X y Y .

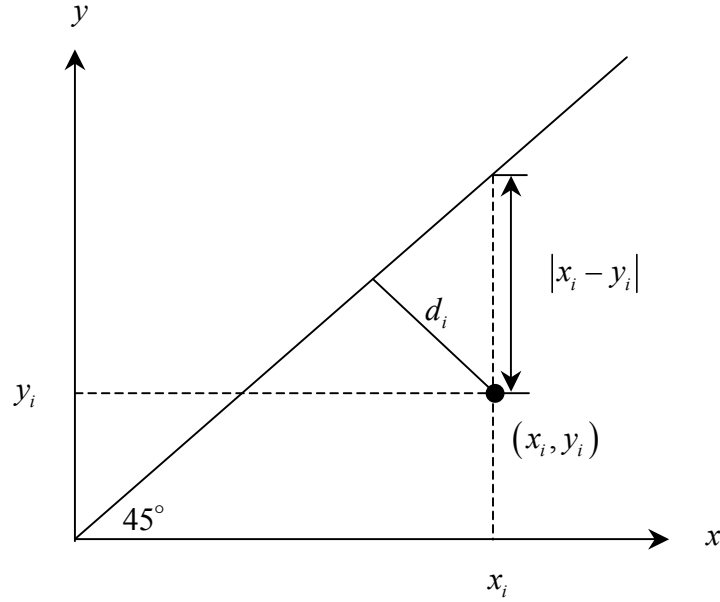


Figura 2: Diagrama de dispersión

Momentos de la Distribución de Probabilidad Bivariada

Valor esperado o valor medio

Se define como:

1.- VA discreta

$$m_{XY} = E\{XY\} = \sum_{i=1}^N \sum_{j=1}^M p_{ij} x_i y_j \quad (\text{A.59})$$

2.- VA continua

$$m_{XY} = E\{XY\} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xy d^2 F_{XY}(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xy f_{XY}(x, y) dx dy \quad (\text{A.60})$$

En la práctica se estima mediante:

$$\hat{m}_{XY} = \sum_{i=1}^N x_i y_i \quad (\text{A.61})$$

El momento bivariado $m_{XY} = E\{XY\}$ es conocido como **covarianza no centrada**.

Covarianza centrada

Dadas dos variables aleatorias X y Y , se puede definir la covarianza de manera análoga a los momentos centrales univariados, como

$$\begin{aligned} Cov(X, Y) &= E\{(X - m_X)(Y - m_Y)\} \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - m_X)(y - m_Y) f_{XY}(x, y) dx dy \end{aligned} \quad (A.62)$$

Se puede designar como

$$\sigma_{XY} = Cov(X, Y) = E\{(X - m_X)(Y - m_Y)\} = E\{XY\} - m_X m_Y \quad (A.63)$$

Y se estima como:

$$\hat{\sigma}_{XY} = \frac{1}{N} \sum_{i=1}^N (x_i - m_X)(y_i - m_Y) = \frac{1}{N} \sum_{i=1}^N x_i y_i - m_X m_Y \quad (A.64)$$

En el caso particular cuando $X = Y$ la autocovarianza es la varianza de X

$$\sigma_X^2 = Var\{X\} = Cov\{X, X\} = E\{[X - m_X]^2\} \geq 0 \quad (A.65)$$

Si la covarianza es nula, es decir $Cov(X, Y) = 0$, se dice que las variables aleatorias X y Y no están correlacionadas.

Coefficiente de correlación

Se define como:

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{Cov\{X, Y\}}{\sqrt{Var\{X\} Var\{Y\}}} \in [-1, 1] \quad (A.66)$$

Caracteriza el grado de dependencia lineal o correlación entre dos variables aleatorias X y Y . Por ejemplo si $Y = aX + b$, entonces se cumple que:

$$\rho_{XY} = \begin{cases} 1, & \text{para } a > 0 \\ -1, & \text{para } a < 0 \end{cases} \quad (A.67)$$

Nótese que si X y Y son independientes, entonces $\rho_{XY} = 0$, pero lo contrario no siempre es cierto. Es decir, si $\rho_{XY} = 0$ no implica que las VAs X y Y sean independientes, sólo se puede afirmar que no puede ser descrita su dependencia mediante un modelo lineal. Para el caso en que X y Y posean una distribución normal bidimensional si es cierto el recíproco.

Relación entre covarianza y variograma.

El variograma esta dado por

$$2\gamma_{XY} = \frac{1}{N} \sum_{i=1}^N [x_i - y_i]^2 \quad (\text{A.68})$$

entonces

$$2\gamma_{XY} = \frac{1}{N} \sum_{i=1}^N [x_i^2 - m_X^2] + \frac{1}{N} \sum_{i=1}^N [y_i^2 - m_Y^2] - 2 \left(\frac{1}{N} \sum_{i=1}^N x_i y_i - m_X m_Y \right) + (m_X - m_Y)^2 \quad (\text{A.69})$$

y por lo tanto

$$2\gamma_{XY} = \sigma_X^2 + \sigma_Y^2 - 2\sigma_{XY} + (m_X - m_Y)^2 \geq 0 \quad (\text{A.70})$$

Si $\uparrow \gamma_{XY} \Rightarrow \downarrow \sigma_{XY}$ y viceversa.

Relación entre coeficiente de correlación y variograma

Si tomamos las variables X y Y estandarizadas: $X' = \frac{X - m_X}{\sigma_X}$ y $Y' = \frac{Y - m_Y}{\sigma_Y}$, donde

$m_{X'} = m_{Y'} = 0$, $\sigma_{X'} = \sigma_{Y'} = 1$. La covarianza $\sigma_{X'Y'}$ coincide con el coeficiente de correlación ρ_{XY}

$$\sigma_{X'Y'} = E \left\{ \left[\frac{X - m_X}{\sigma_X} \right] \left[\frac{Y - m_Y}{\sigma_Y} \right] \right\} = \rho_{XY} \quad (\text{A.71})$$

Aplicando la Ec. (A.70), se obtiene la relación

$$\gamma_{X'Y'} = 1 - \rho_{XY} \quad (\text{A.72})$$

Tanto la covarianza como el coeficiente de correlación son medidas de la dependencia lineal entre dos variables aleatorias, mientras que el variograma es una medida de la variabilidad (no dependencia) entre ambas.

A.3 REGRESIÓN LINEAL Y MÍNIMOS CUADRADOS

Sea Y una variable aleatoria que puede expresarse como:

$$Y = m + e \quad (\text{A.73})$$

donde m - es el valor esperado de Y , $m = E[Y]$, y e - es el error.

En el caso de regresión lineal m puede expresarse como una combinación lineal de p variables aleatorias:

$$m = b_0 + b_1X_1 + b_2X_2 + \dots + b_{p-1}X_{p-1} = \underline{\mathbf{b}} \cdot \underline{\mathbf{X}} \quad (\text{A.74})$$

donde

$(1, X_1, X_2, \dots, X_{p-1})$ - es el vector de las p variables aleatorias,

$(b_0, b_1, b_2, \dots, b_{p-1})$ - es el vector de los p parámetros.

Dadas n observaciones (Y_1, Y_2, \dots, Y_n) de la variable dependiente Y , correspondientes a n experimentos, el problema que se plantea es la estimación de los parámetros $(b_0, b_1, b_2, \dots, b_{p-1})$ de manera que sea mínima la suma de los cuadrados de los errores:

$$S_e = \sum_{i=1}^n e_i^2 \quad (\text{A.75})$$

Usando notación matricial:

$$\underline{\mathbf{Y}} = (Y_1, Y_2, \dots, Y_n) \quad (\text{A.76})$$

$$\underline{\mathbf{b}} = (b_0, b_1, b_2, \dots, b_{p-1}) \quad (\text{A.77})$$

$$\underline{\mathbf{X}} = \begin{pmatrix} 1 & X_{11} & \dots & X_{1p} \\ 1 & X_{21} & \dots & X_{2p} \\ \vdots & \vdots & \dots & \vdots \\ 1 & X_{n1} & \dots & X_{np} \end{pmatrix} \quad (\text{A.78})$$

El vector de errores posee esperanza matemática nula y una matriz de covarianzas:

$$E[\underline{\mathbf{e}}] = 0, \quad (\text{A.79})$$

$$\sigma^2 \underline{\underline{\mathbf{C}}}_{\mathbf{e}}, \quad (\text{A.80})$$

Mínimos cuadrados ordinarios

Si los errores e_i son independientes, su matriz de covarianzas puede expresarse como $\sigma^2 \underline{\underline{\mathbf{I}}}$, entonces el estimador de $\underline{\mathbf{b}}$ por mínimos cuadrados ordinarios está dado por:

$$\underline{\mathbf{b}}^* = (\underline{\underline{\mathbf{X}}}^T \underline{\underline{\mathbf{X}}})^{-1} \underline{\underline{\mathbf{X}}}^T \underline{\mathbf{Y}} \quad (\text{A.81})$$

donde la suma de los cuadrados de los errores S_e es:

$$S_e = \underline{\mathbf{Y}}^T \underline{\mathbf{Y}} - \underline{\mathbf{b}}^{*T} \underline{\underline{\mathbf{X}}}^T \underline{\underline{\mathbf{X}}} \underline{\mathbf{b}}^* \quad (\text{A.82})$$

Un estimador insesgado de σ^2 es:

$$S^2 = \frac{1}{n-p} S_e \quad (\text{A.83})$$

Y la matriz de covarianzas de $\underline{\mathbf{b}}$ está dada por:

$$\underline{\underline{\mathbf{C}}}_{\mathbf{b}} = \sigma^2 (\underline{\underline{\mathbf{X}}}^T \underline{\underline{\mathbf{X}}})^{-1} \quad (\text{A.84})$$

Mínimos cuadrados generalizados

Si los errores e_i no son independientes, entonces el estimador de $\underline{\mathbf{b}}$ está dado por:

$$\underline{\mathbf{b}}^* = (\underline{\underline{\mathbf{X}}}^T \underline{\underline{\mathbf{C}}}_{\mathbf{e}}^{-1} \underline{\underline{\mathbf{X}}})^{-1} \underline{\underline{\mathbf{X}}}^T \underline{\underline{\mathbf{C}}}_{\mathbf{e}}^{-1} \underline{\mathbf{Y}} \quad (\text{A.85})$$

donde la suma de los cuadrados de los errores S_e es:

$$S_e = (\underline{\mathbf{Y}} - \underline{\underline{\mathbf{X}}} \underline{\mathbf{b}}^*)^T \underline{\underline{\mathbf{C}}}_{\mathbf{e}}^{-1} (\underline{\mathbf{Y}} - \underline{\underline{\mathbf{X}}} \underline{\mathbf{b}}^*) \quad (\text{A.86})$$

Y la matriz de covarianzas de $\underline{\mathbf{b}}$ está dada por:

$$\underline{\underline{\mathbf{C}}}_{\mathbf{b}} = \sigma^2 (\underline{\underline{\mathbf{X}}}^T \underline{\underline{\mathbf{C}}}_{\mathbf{e}}^{-1} \underline{\underline{\mathbf{X}}})^{-1} \quad (\text{A.87})$$

Nota: Los mínimos cuadrados ponderados es un caso particular de mínimos cuadrados generalizados cuando la matriz $\underline{\underline{\mathbf{C}}}_{\mathbf{e}}$ es diagonal.

ANEXO B: GLOSARIO DE TÉRMINOS GEOESTADÍSTICOS

Alcance (range): Para el modelo esférico, la distancia en la que el modelo alcanza el valor máximo, o sill (meseta). Para los modelos gaussianos y exponencial, que se aproximan a la meseta asintóticamente. Se usa el término alcance para designar el alcance ‘práctico’ o ‘efectivo’ donde la porción alcanza aproximadamente el 95 % del máximo. El modelo Nugget tiene una meseta con un rango de cero; el modelo lineal usa ‘meseta/alcance’ simplemente para definir la pendiente.

Anisotropía: En Geoestadística, es la situación cuando el variograma exhibe un comportamiento que varía con la dirección.

Anisotropía Geométrica: Es el caso de anisotropía cuando los alcances del variograma forman una elipse, donde el mayor y el menor alcance son perpendiculares y corresponden al mayor y al menor semieje de la elipse.

Covarianza: Es una medida estadística de la correlación entre dos variables. En Geoestadística, la covarianza es usualmente tratada como la simple inversión del variograma, calculado como la varianza total de la muestra menos el valor del variograma. Estos valores de covarianza, así como los valores del variograma, se utilizan en las ecuaciones de la matriz de Kriging para una mayor eficiencia de cálculo.

Cuadrante de Búsqueda: La elipse de la vecindad de búsqueda puede ser dividida en cuatro sectores de ángulos iguales, lo que pueden tener especificada un número mínimo y máximo de muestras. También puede especificarse un número limitado de sectores consecutivos vacíos. Cuando el criterio especificado no se satisface para un punto o bloque particular, entonces no se produce el estimado por Kriging.

Deriva (drift): El valor esperado de una función aleatoria puede ser constante o depender de las coordenadas de la posición. La deriva es una característica de la función aleatoria y no de los datos.

Desviación Estándar del Kriging: Error estándar de la estimación calculada para el estimado del Kriging. Por definición, Kriging es el estimador lineal ponderado con una serie particular de pesos los que minimizan el valor de la varianza de la estimación.

Discretización : En Kriging, el proceso de aproximación al área de un bloque por un arreglo finito de puntos.

Estacionaridad: Es una propiedad de la función aleatoria. Se dice que una función aleatoria es *estrictamente estacionaria* si su función de distribución de probabilidad es invariante a cualquier traslación respecto a un vector \underline{h} .

Función Aleatoria: Puede ser vista como una colección de variables aleatorias que dependen de la posición.

Geoestadística: Metodología para el análisis de datos espacialmente correlacionados. El rasgo característico es el uso de variogramas de técnicas relacionadas para cuantificar y modelar la correlación espacial de la estructura. También incluye diferentes técnicas como Kriging, la cual utiliza modelos de correlación espacial. Los métodos geoestadísticos son aplicables en todas las ciencias de la Tierra. Pueden aplicarse para explorar los procesos responsables de la variación espacial. También pueden aplicarse donde existe una información completa obtenida por percepción remota u otra fuente, para determinar un muestreo eficiente, así como también para estimar el valor de propiedades en localidades no muestreadas.

Hipótesis Intrínseca: Suposición de que la media y la varianza de las diferencias son estacionarias, es decir no dependen de una traslación \underline{h} de la función aleatoria.

Intervalo (lag): Intervalos de distancia de la clase usada para calcular el variograma.

Kriging de bloques: Estimación del valor de un bloque a partir de los valores de una muestra continua usando Kriging. El área de un bloque es un arreglo de aproximadamente 2x2, 3x3, ó 4x4 puntos con centro en cada nodo de la malla especificada. Se dice que se obtiene un valor suavizado de la estimación.

Kriging Ordinario: Es un tipo de Kriging que asume que la media local no está necesariamente cercana a la media de la población, y que usa solamente para el estimado la muestra para la vecindad local. Es el método usado mas comúnmente por su robustez.

Kriging Puntual: Estimación del valor de un punto de los valores de la muestra cercana usando kriging. El estimado para un punto será casi similar al estimado por un bloque relativamente cercano centrado en el punto, pero la desviación estándar del kriging calculada será alta. Cuando el punto del Kriging coincide con el lugar de la muestra, el estimado tendrá un valor igual al de la muestra.

Kriging Simple: Variedad de kriging que asume que las medias locales son relativamente constantes e iguales a la media de la población, la que es bien conocida. La media de la población es usada como un factor en cada estimado local, con todas las muestras en la misma vecindad. No es un método muy usado.

Kriging: Método de interpolación del valor medio ponderado donde los pesos asignados a las muestras minimizan la varianza del error, la que se calcula como una función del modelo de variograma y localizaciones de las muestras relacionadas unas con las otras, y del punto o bloque que está siendo estimado.

Madograma: Gráfico de la diferencia media absoluta de pares de muestras medidas como una función de la distancia y la dirección. Los madogramas no son verdaderos variogramas, y generalmente no deben ser usados en Kriging. En caso de usarse los estimados del Kriging tienen que ser ‘razonables’, pero las desviaciones estándar del Kriging no tendrán significado.

Meseta (sill): Límite superior de cualquier modelo de variograma acotado, al que tiende asintóticamente para grandes distancias. Para el modelo lineal, la relación ‘sill/rango’ se usa para definir la pendiente.

Modelo Esférico: Es una de las funciones autorizadas que es usada frecuentemente como modelo de variograma, en combinación con el modelo Nugget.

Modelo Exponencial: Es una de las funciones autorizadas que es usada frecuentemente como modelo de variograma, en combinación con el modelo Nugget.

Modelo Lineal: Es una de las funciones autorizadas que es usada frecuentemente como modelo de variograma, en combinación con el modelo Nugget.

Modelo Nugget: Modelo de varianza constante comúnmente usado en combinación con uno o mas funciones cuando se ajustan modelos matemáticos a variogramas experimentales.

Modelos de variograma anidado: Aquel que es la suma de dos o mas modelos tales como Nugget, esférico, etc. Agregando el componente Nugget a uno de los otros modelos se obtiene el modelo anidado mas común, aunque muy pocas veces se usan combinaciones muy complejas.

Octante de búsqueda: La elipse de búsqueda de vecindad en Kriging puede ser dividida en 8 sectores de ángulos iguales, los que pueden tener especificadas número de muestras mínimas y máximas. También puede especificarse un número consecutivo de sectores vacíos. Cuando no se satisface el criterio especificado, para un punto o un bloque, entonces no se produce el estimado por Kriging.

Semi-Variograma: Es sinónimo de ‘variograma’. No hay acuerdo en la literatura geoestadística de cual término debe usarse y se usan ambos indistintamente.

Soporte: El término es usado en ambos sentidos: matemático y físico. Con frecuencia los valores experimentales están definidos en soportes puntuales o casi puntuales, en otros casos son los valores medios $Z_{V_i}(\underline{x}_i)$ definidos en los soportes V_i centrados en los puntos \underline{x}_i , donde los soportes pueden ser todos diferentes.

Tendencia (trend): En muchas ocasiones se usa el término intercambiable con el de deriva (drift). Aunque también se asocia con la representación determinística del valor medio obtenida mediante el procedimiento de Análisis de Superficie de Tendencia (Trend Surface Analysis) usualmente como polinomios de las coordenadas mediante ajuste con mínimos cuadrados.

Validación Cruzada: Es una técnica para probar la validez de un modelo de variograma obtenido por Kriging de cada localización de la muestra con todas las otras muestras en la vecindad de búsqueda, y comparando los estimados con los valores reales de la muestra. La interpretación de los resultados, sin embargo, es en ocasiones muy difícil. La existencia de grandes diferencias entre los valores reales y estimados, pueden indicar la presencia de ‘outliers espaciales’, o puntos que aparentemente no pertenecen a esos contornos.

Variograma No ergódico: Variograma calculado restando a la función de covarianzas la varianza de la muestra. Esta aproximación compensa los casos donde existen variogramas direccionales, digamos, en aquellos donde los pares de datos del oeste tienen medias diferentes de los pares del este. ‘No ergódico’ es en ocasiones un término oscuro que se refiere al echo de algunas suposiciones probabilísticas relacionadas con el cálculo del variograma no son tan necesarias. Los variogramas no ergódicos pueden ser modelados y usados en Kriging de la misma forma que los variogramas ordinarios.

Variograma relativo: El variograma en el que el valor del variograma ordinario de cada intervalo ha sido dividido por el cuadrado de la media de los valores de la muestra usada en el cálculo del intervalo. Es en ocasiones útil cuando se presenta un ‘efecto proporcional’, cuando hay áreas de varianza mayor que la varianza promedio. Cuando los modelos de

variogramas relativos son usados en Kriging, la desviación estándar del Kriging resultante representa fracciones decimales de los valores estimados.

Variograma: Gráfico de la semivarianza (la mitad de la diferencia media cuadrada) como una función de la distancia (y de la dirección) entre pares de valores muestrales. Típicamente, son examinadas todos los pares de muestras y agrupadas en clases (intervalos) de distancia y dirección aproximadamente iguales. Los variogramas aportan un significado de cuantificación de la relación observada comúnmente de que las muestras cercanas entre ellas tienden a tener valores mas similares que las muestras mas alejadas. Es la herramienta central de la Geoestadística, ya que da una descripción adecuada de la escala y del patrón de la variación espacial dentro de una región dada. Los procesos de kriging, optimización del muestreo y los cálculos de probabilidades condicionadas utilizan la información de la variación espacial obtenida de los variogramas.

Vecindad de búsqueda: Es el área elíptica centrada en un punto o bloque que está siendo analizado por el Kriging. Solamente las muestras dentro de la elipse serán usadas para el cálculo de la estimación por Kriging.

GLOSSARY OF GEOSTATISTICAL TERMS

- ANISOTROPIC:** This term is applied both to a random function and to its variogram (or covariance) when the values of the variogram depend on both the distance and the direction. Also see Isotropic.
- ANISOTROPY:** In Geostatistics, the situation where a variogram exhibits a longer range (i.e., better correlation) in one direction than another.
- BLOCK KRIGING:** Estimating the value of a block from a set of nearby sample values using kriging. In Geo-EAS, the block area is approximated by a 2x2, 3x3, or 4x4 array of points centered on each specified grid node.
- COVARIANCE:** A statistical measure of the correlation between two variables. In Geostatistics, covariance is usually treated as the simple inverse of the variogram, computed as the overall sample variance minus the variogram value. These covariance values, rather than variogram values, are actually used in the Geo-EAS kriging matrix equations for greater computational efficiency.
- CROSS VALIDATION:** A technique for testing the validity of a variogram model by kriging each sampled location with all of the other samples in the search neighborhood, and comparing the estimates with the true sample values. Interpretation of results, however, can often be difficult. Unusually large differences between estimated and true values may indicate the presence of "spatial outliers", or points which do not seem to belong with their surroundings. A method for comparing two or more conjectured variogram (or covariance) models. The technique depends on Jackknifing the data and on the exactness of the kriging estimator.
- DISCRETIZATION:** In kriging, the process of approximating the area of a block by a finite array of points.
- DRIFT:** The expected value of a random function, it may be constant or it may depend on the coordinates of the location. In order for a random function to be stationary, second-order stationary or to satisfy the Intrinsic Hypothesis; the drift must be a constant. The drift is a characteristic of a random function and not of data.
- EXPONENT MODEL:** A function frequently used when fitting mathematical models to experimental variograms, often in combination with a nugget model.
- GAUSSIAN MODEL:** A function frequently used when fitting mathematical models to experimental variograms, often in combination with a nugget model.
- GEOSTATISTICS:** A methodology for the analysis of spatially correlated data. The characteristic feature is the use of variograms of related techniques to quantify and model the spatial correlation

structure. Also includes the various techniques such as kriging, which utilize spatial correlation model.

INTRINSIC HYPOTHESIS: A weak form of stationarity for a random function sufficient for deriving the kriging equations corresponding to the (Ordinary) kriging estimator. See (i) and (ii).

INVERTED COVARIANCE VARIOGRAM: A variogram computed by subtracting lag covariances from the sample variance. This approach compensates for cases where in directional variograms, the mean of the, say, west members of the sample pairs is not the same as the mean of the east members. "InvCov" is a rather obscure term referring to the fact that some probabilistic assumptions underlying the variogram computation are no longer necessary. Inverted Covariance variograms may be modeled and used in kriging in the same way as ordinary variograms.

ISOTROPIC: A term applied both to a random function and to its variogram. See anisotropic which is the complementary property.

KRIGING: A weighted-moving-average interpolation method where the set of weights assigned to samples minimizes the estimation variance, which is computed as a function of the variogram model and locations of the samples relative to each other, and to the point or block being estimated.

KRIGING EQUATIONS: A set of linear equations whose solution includes the values of the weights in the kriging estimator.

KRIGING STANDARD DEVIATION: The standard error of estimation computed for a kriged estimate. By definition, kriging is the weighted linear estimate with a particular set of weights which minimizes the computed estimation variance (standard error squared). The relationship of the kriging standard deviation to the actual error of estimation is very dependent on the variogram model used and the validity of the underlying assumptions (kriging standard deviations should be interpreted with caution).

KRIGING VARIANCE: The minimized value of the estimation variance, i.e., the variance of the error of estimation. This variance is not data dependent but rather is determined by the variogram and the sample location pattern as well as the location of the point to be estimated relative to the sample locations.

LAG: A distance class interval used for variogram computation.

LINEAR MODEL: A function frequently used when fitting mathematical models to experimental variograms, often in combination with a nugget model.

MADOGRAM: A plot of mean absolute difference of paired sample measurements as a function of distance and direction. Madograms are not true variograms, and generally should not be used in kriging. If used, the kriged estimates might be

"reasonable", but the kriging standard deviations will be meaningless.

NESTED VARIOGRAM MODEL: A model which is the sum of two or more component models such as nugget, spherical, etc. Adding a nugget component to one of the other models is the most common nested model, but not more complex combinations are occasionally used.

NON-ERGODIC VARIOGRAM: A variogram computed by subtracting lag covariance from the sample variance. This approach compensates for cases where in directional variograms, the mean of the, say, west members of the sample pairs is not the same as the mean of the east members. "Non-ergodic" is a rather obscure term referring to the fact that some probabilistic assumptions underlying the variogram computation are no longer necessary. Non-ergodic variograms may be modeled and used in kriging in the same way as ordinary variograms.

NUGGET: The variogram may exhibit an apparent discontinuity at the origin. The magnitude of the discontinuity is called the nugget.

NUGGET MODEL: A constant variance model most often used in combination with one or more other functions when fitting mathematical models to experimental variograms.

OCTANT SEARCH: In Geo-EAS, the kriging search neighborhood ellipse may be divided into eight equal-angle sectors, which may have minimum and maximum numbers of samples specified. A limit on the number of consecutive empty sectors may also be specified. When the specified criteria are not satisfied for a particular point or block then the kriged estimate is not produced.

ORDINARY KRIGING: A variety of kriging which assumes that local means are not necessarily closely related to the population mean, and which therefore uses only the samples in the local neighborhood for the estimate. Ordinary kriging is the most commonly used method for environmental situations.

POINT KRIGING: Estimating the value of a point from a set of nearby sample values using kriging. The kriged estimate for a point will usually be quite similar to the kriged estimate for a relatively small block centered on the point, but the computed kriging standard deviation will be higher. When a kriged point happens to coincide with a sample location, the kriged estimate will equal the sample value.

POSITIVE DEFINITE: A term applied both to matrices and to functions, (Auto) covariance functions must be positive definite whereas the negative of variograms must be conditionally positive definite. Conditional positive definiteness is a weaker condition.

QUADRANT SEARCH: In Geo-EAS, the kriging search neighborhood ellipse may be divided into four equal-angle sectors, which may have minimum and maximum numbers of samples specified. A limit on the number of consecutive empty sectors may also be specified.

When the specified criteria are not satisfied for a particular point or block then the kriged estimate is not produced.

RANDOM FUNCTION: A random function may be seen in two different forms; it may be thought of as a collection of dependent random variables with one for each possible sample location. Alternatively it may be thought of as a "random variable" whose values are functions rather than numbers.

RANGE(of a variogram): The distance at which the variogram becomes a constant. The Power model does not have a (finite) range. The Exponential and Gaussian models have only an apparent range. For a spherical model, the distance at which the model reaches its maximum value, or sill. For the exponential and gaussian models, which approach the sill asymptotically, Geo-EAS uses range to mean the "practical", or "effective" range, where the function reaches approximately 95% of the maximum. The nugget model effectively has a sill with a range of zero; the linear model uses "sill/range" merely to define the slope.

RELATIVE VARIOGRAM: A variogram in which the ordinary variogram value of each lag has been divided by the square of the mean of the sample values used in computing the lag. This is sometimes useful when a "proportional effect" is present; i.e., when areas of higher than average variance. When relative variogram models are used in kriging, the resulting kriging standard deviations represent decimal fractions of the estimated values.

SEARCH NEIGHBORHOOD: In Geo-EAS, an elliptical area centered on a point or block being kriged. Only samples within the ellipse will be used for kriging. When the next point is kriged, the ellipse will be re-centered, and a different(perhaps) set of samples will be used.

SEMI-VARIOGRAM: Identical to the term "variogram" as defined in Geo-EAS. There is disagreement in the geostatistical literature as to which term should be used. Geo-EAS uses "variogram" for simplicity.

SILL(of a variogram): The value of the variogram for distances beyond the range of the variogram. The Power model does not have a sill. The upper limit of any variogram model which has such a limit, i.e., which tends to "level off" at large distances. In Geo-EAS, the spherical, gaussian, exponential, and nugget models have sills. For the linear model, "sill/range" is used merely to define the slope.

SIMPLE KRIGING: A variety of kriging which assumes that local means are relatively constant and equal to the population mean, which is well-known. The population mean is used as a factor in each local estimate, along with the samples in the local neighborhood. This is not usually the most appropriate method for environmental situations.

SPATIAL CORRELATION: Used both as a generic term to denote that data at two locations is correlated in some sense as a function of their locations and also to denote the value of a spatial structure

function such as a variogram or (auto)covariance for a pair of locations.

SPHERICAL MODEL: A function frequently used when fitting mathematical models to experimental variograms, often in combination with a nugget model.

STATIONARITY: Several different forms of stationarity are utilized in geostatistics. Stationarity, in one of its forms, is a property of a random function rather than of a data set. It expresses the property that certain joint distributions are translation invariant or that certain moments of the random function are translation invariant. See second order stationarity and the Intrinsic hypothesis.

SUPPORT: The term is used in both a mathematical and in a physical sense. Many, if not most variables of interest in geostatistics, such as the concentrations of chemical elements or compounds only have values at "points" in an idealized sense although the random function treats them in this manner. The data values are usually associated with a physical sample having a length, area or volume; the concentration then represents an average concentration over this length, area or volume. This length, area or volume is called the support. Although it is common to report laboratory analyses in such a way as to not reflect the original support, non-point support has a significant effect on the variogram modeling process and there is a significant difference in estimating the average value over a large volume and in estimating the average value over a small volume. The kriging estimator and equations allow this to be incorporated.

TREND: While sometimes used interchangeably with the term "drift", in geostatistics the two are considered separate. The term is usually reserved to denote the deterministic representation obtained by the use of Trend Surface Analysis, i.e., a functional representation for spatially located data (usually taken as a polynomial in the position coordinates). The "trend" is obtained by a least squares fit to the data. As an estimator to the mean of a random function it is sub-optimal. If the residuals from trend surface analysis are used to model the variogram, a biased variogram estimator results.

VARIOGRAM: A plot of the variance (one-half the mean squared difference) of paired sample measurements as a function of the distance (and optionally of the direction) between samples. Typically, all possible sample pairs are examined, and grouped into classes (lags) of approximately equal distance and direction. Variograms provide a means of quantifying the commonly observed relationship that samples close together will tend to have more similar values than samples far apart.

BIBLIOGRAFÍA

1. Akaike, H., 1974, *A new look at statistical model identification*, IEEE Transactions on Automatic Control, AC-19, 716-722.
2. Akaike, H., 1977, *An entropy maximization principle*, Applications of Statistics, P.R. Krishnaiah, ed., North Holland, Amsterdam.
3. Armstrong, M. y P. Delfiner, 1980, *Towards a more robust variogram: case study on coal*, Note n-671, CGMM, Fontainebleau, Francia.
4. Armstrong, M., 1984, *Problems with Universal Kriging*, Math. Geol., 16(1), 101-108.
5. Clark I. and W. V. Harper , 2000, "Practical Geostatistics 2000".
6. Cressie, N. y J. Hawkins, 1980, *Robust estimation of the variogram*, J. Math. Geol., 12(2), 115-126.
7. Cressie, N., 1985, *Fitting variogram models using weighted least squares*, J. Math. Geol., 17, 693-702.
8. David M., 1977, *Geostatistical Ore Estimation*, Elsevier, 264 pp.
9. David M., 1988, "Handbook Of Applied Advanced Geostatistical Ore Reserve Estimation".
10. David, M., 1976, *The Practice of Kriging*, Advanced Geostatistics in the Mining Industry, M. Guarascio et al., eds. Dordrecht, Holland, 31-48.
11. Díaz-Viera M. y R. Barandela, 1994, *GEOESTAD: Un sistema de computación para el desarrollo de aplicaciones geoestadísticas*, Resúmenes del II Taller Iberoamericano "Informática y Geociencias", GEOINFO'94, C. Habana.
12. "Geostatistics for the Next Century", Proceedings edited by Roussos Dimitrakopoulos December 1993.
13. Hoeksema, R. J. y P. K. Kitanidis, 1985, *Analysis of the spatial structure of Properties of selected aquifers*, Water Resour. Res., 563-572.
14. Hohn M. E., 1998, "Geostatistics and Petroleum Geology", Second Edition, Kluwer Academic Publishers, 235 pp.

15. Isaaks, E. H. and R. M. Srivastava, 1989, "An Introduction to Applied Geostatistics", Oxford Univ. Press, pp. 561.
16. Journel, A. G. y Ch. J. Huijbregts, 1978, "Mining Geostatistics", Academic Press, New York, 590 pp.
17. Journel, A. G., 1986, *Geostatistics, Models and Tools for the Earth Sciences*, Short Course in Geology: Vol. 8, American Geophysical Union, 40 pp.
18. Journel, A. G., 1989, "Fundamentals of Geostatistics in Five Lessons", Math. Geol. 18(1), 119-140.
19. Kitanidis, P. K., 1997, "Introduction to Geostatistics: Applications in Hydrogeology", Cambridge University Press, 249 pp.
20. Omre, H., 1984, *The variogram and its estimation*, Geostatistics for Natural Resources Characterization, Verly et al. (Eds.), NATO ASI Series, Reidel, Hingham, Mass., Vol. 1, 107-125.
21. Samper, F.J. y J. Carrera, 1990, "Geoestadística, aplicaciones a la hidrogeología subterránea", Centro Internacional de Métodos Numéricos en Ingeniería, Barcelona, 484 pp.