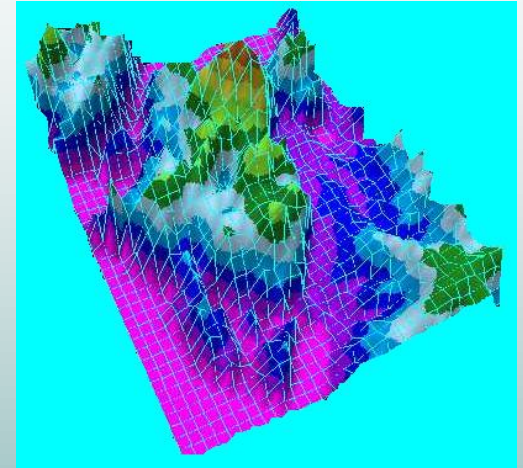




INSTITUTO MEXICANO DEL PETRÓLEO

GEOESTADÍSTICA APLICADA

Tema: **Geoestadística
Multivariada**



Instructores:

Dr. Martín A. Díaz Viera (mdiazv@imp.mx)

Dr. Ricardo Casar González (rcasar@imp.mx)

2004

Contenido

- Introducción
- Momentos cruzados de segundo orden
- Estimación del covariograma
- Análisis estructural multivariado
- Modelo de correogionalización lineal
- Validación del modelo de semivariograma cruzado
- Cokriging Ordinario
- Cokriging en el caso de funciones aleatorias intrínsecas
- Cokriging en el caso no estacionario
- Cokriging Colocado
- Ejemplos de Aplicaciones del Cokriging
- Dificultades y Aspectos Prácticos del Cokriging
- Métodos alternativos al Cokriging

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.

Introducción

- Las aplicaciones multivariadas que han recibido una mayor atención en la geoestadística 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.

Momentos Cruzados de Segundo Orden

- Estacionaridad de segundo orden para las F.A. $Z_i(\underline{x})$ y $Z_j(\underline{x})$
- Covarianza cruzada

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

- Semivariograma cruzado (Covariograma)

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

donde $m_i = E \left[Z_i(\underline{x}) \right]$ y $m_j = E \left[Z_j(\underline{x}) \right]$ valores esperados

- Cuando $i=j$ los momentos cruzados se convierten en la covarianza y en la semivarianza.

Estimación del covariograma

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

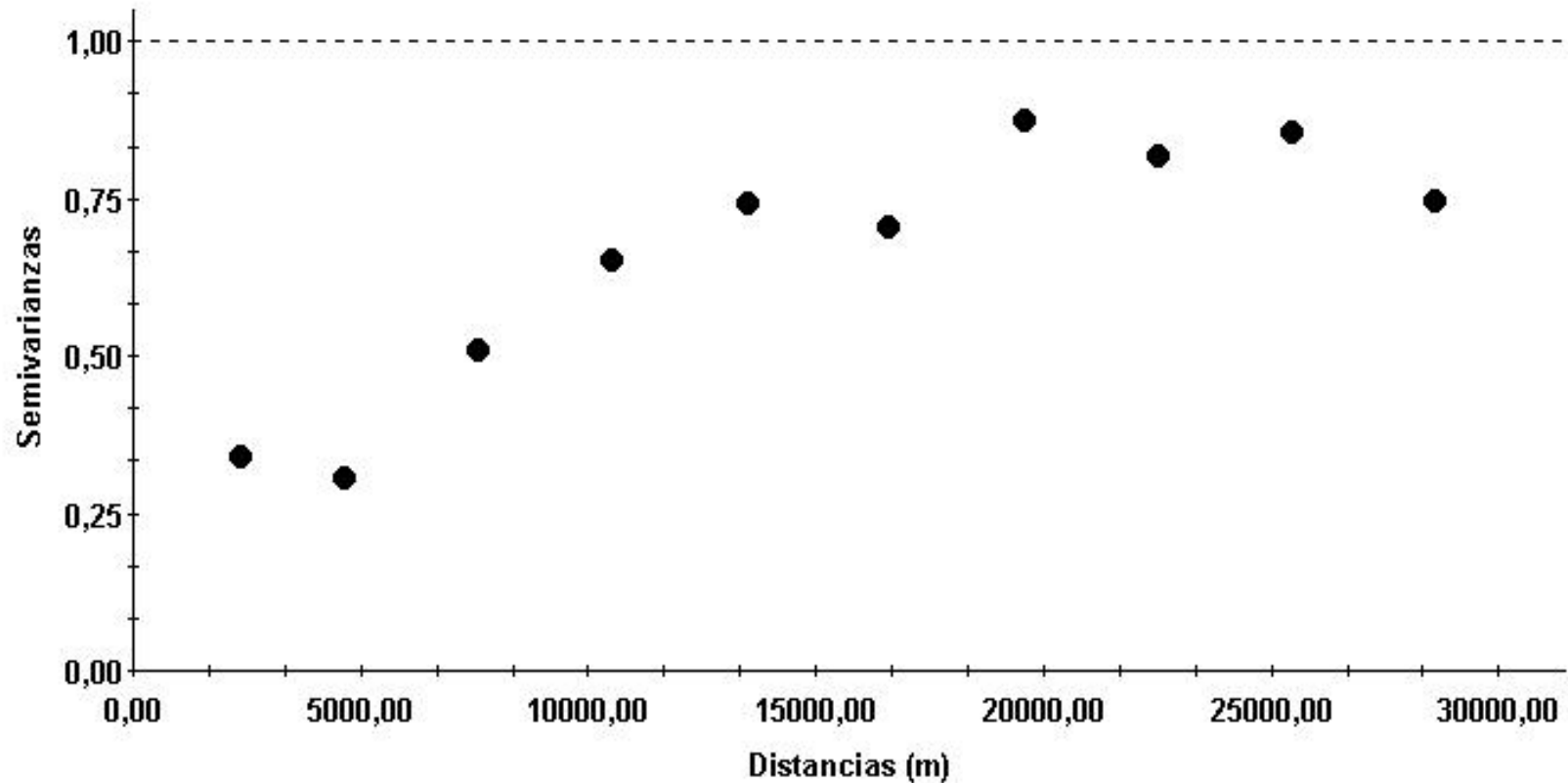
$$\gamma_{ij}^*(\underline{h}) = \frac{1}{2N(\underline{h})} \sum_{k=1}^{N(\underline{h})} [Z_i(\underline{x}_k + \underline{h}) - Z_i(\underline{x}_k)] [Z_j(\underline{x}_k + \underline{h}) - Z_j(\underline{x}_k)]$$

- donde $N(\underline{h})$ es el número de pares y separados a una distancia $h=|\underline{h}|$.
- Es una generalización del estimador del semivariograma simple y por lo tanto adolece de los mismos problemas y limitaciones.

Estimación del covariograma

Ejemplo de covariograma estimado para dos FAs

Variograma Cruzado Estimado de Ln(Pluv.)-Ln(Radar)



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
- Para modelar los variogramas cruzados de n FAs, se deben estimar y modelar (ajustar) un total de $n(n+1)/2$ variogramas simples.
- El uso de modelos de variogramas autorizados o combinaciones de éstos no garantiza que la matriz de covarianzas sea positiva definida.

Análisis estructural multivariado

- La manera más aceptada para realizar un análisis estructural multivariado es mediante un *modelo de correionalización lineal* (Goovaerts, 1997).
- Existen otras metodologías menos difundidas que usan métodos espectrales y están basadas en el teorema de Bochner (Christakos, 1992; Wackernagel, 1995).

Modelo de corregeionalización lineal

- Un *modelo de corregeionalización lineal* está dado por

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

- en términos de las *covarianzas*

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

- en términos de las *semivarianzas*.

Modelo de corregeionalización lineal

- Se interpreta como $S+1$ estructuras anidadas a diferentes escalas.
- Las matrices de corregeionalización \underline{V}_k son las matrices de covarianzas que describen la correlación multivariada a la escala k .
- Note que a cada escala le corresponde una estructura elemental o básica
- Si determinada estructura básica no está presente, se le hace corresponder un coeficiente cero en la matriz

Modelo de corregeionalización lineal

- Para establecer un modelo de corregeionalización lineal se debe probar que las matrices de coeficientes \underline{V}_k son positivas semidefinidas.
- Por definición, una matriz es *positiva semidefinida* (Golub y Van Loan, 1989) si

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

- 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.

Modelo de corregeionalización lineal

- Modelo de corregeionalización lineal para dos FAs

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

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

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

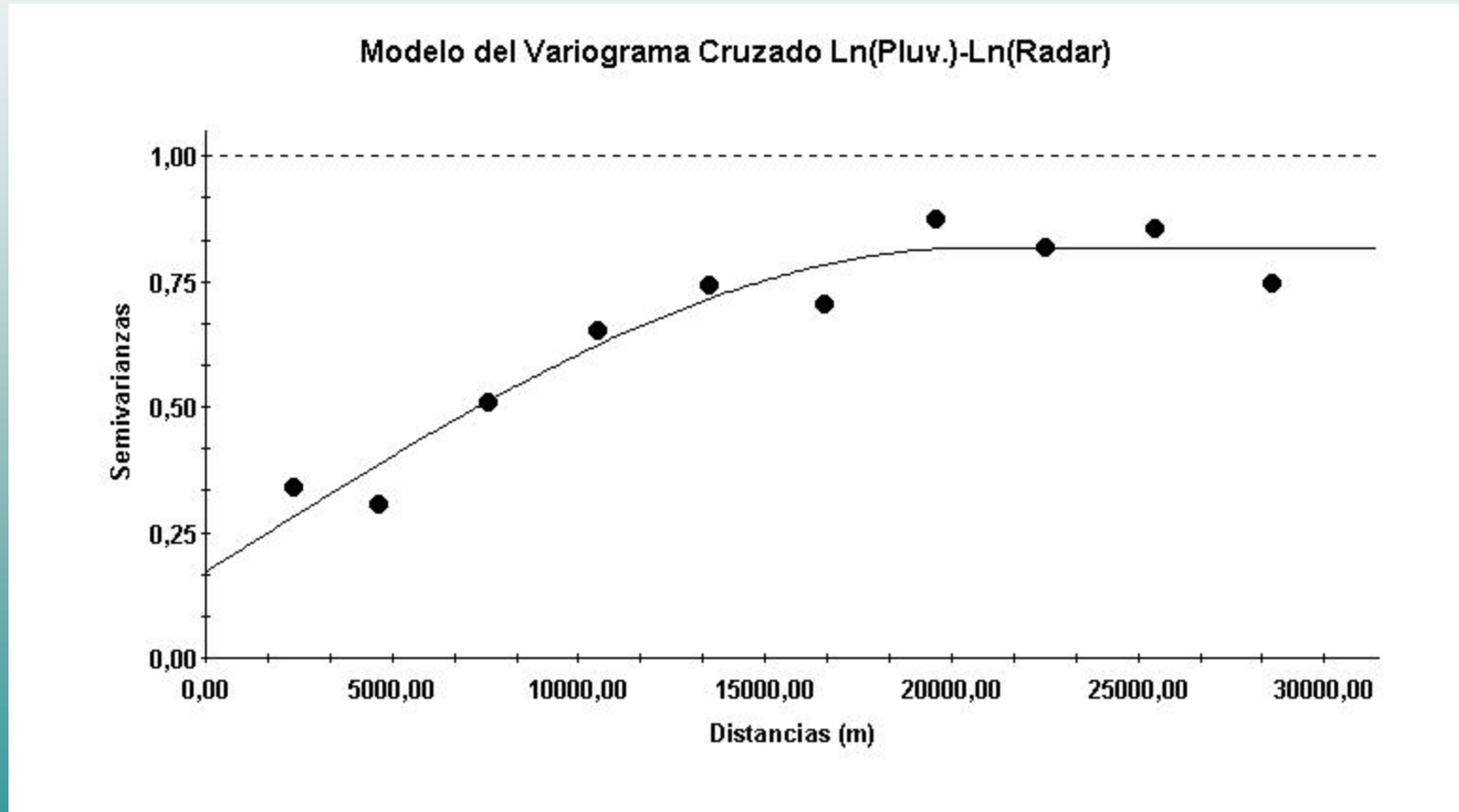
Modelo de correionalización lineal

- El esquema general del Análisis Estructural Multivariado
 1. Modelar cada semivariograma simple y semivariograma cruzado individualmente
 2. Determinar el número de estructuras anidadas de manera que sea mínimo (es deseable que sea cuanto más tres)
 3. Comprobar que todos los determinantes de los menores de orden dos son no negativos.
 4. Verificar que todas las matrices de correionalización sean positivas semidefinidas, en caso contrario hacer los cambios necesarios hasta satisfacer la condición o volver al paso 2.

Modelo de correogionalización lineal

Ejemplo de ajuste del variograma cruzado

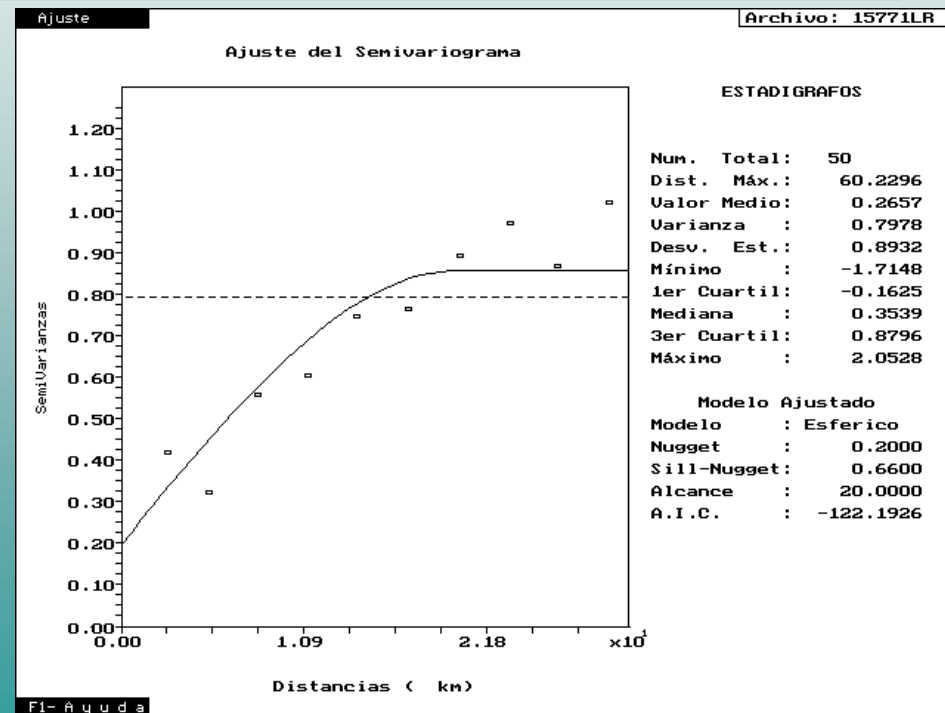
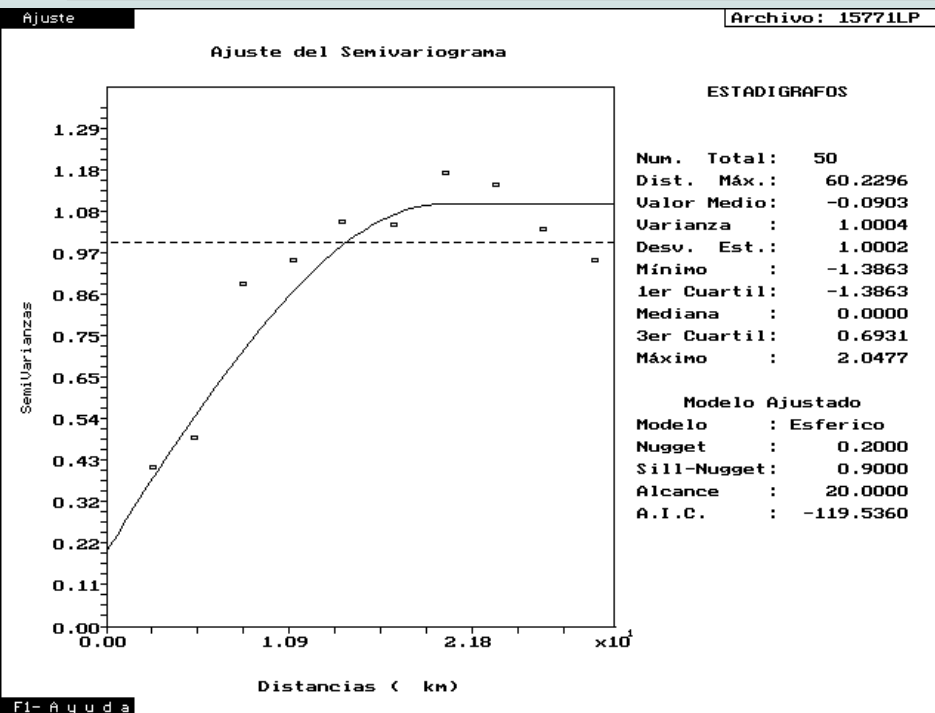
Variables	Modelo	Nugget	Sill-Nugget	Alcance	AIC
Ln Pluv.- Ln Radar	Esférico	0.17	0.647	20.00	-119



Modelo de correogionalización lineal

Ejemplo de ajuste del modelo de correogionalización lineal

Variables	Modelo	Nugget	Sill-Nugget	Alcance	AIC
Ln Pluv.	Esférico	0.20	0.90	20.00	-119
Ln Radar	Esférico	0.20	0.66	20.00	-122
Ln Pluv. - Ln Radar	Esférico	0.17	0.65	20.00	-119



Modelo de correogionalización lineal

Ejemplo de ajuste del modelo de correogionalización lineal

El modelo de correogionalización lineal resultante de Ln Pluv. y Ln Radar es:

$$\begin{pmatrix} \gamma_{PP}(\underline{h}) & \gamma_{PR}(\underline{h}) \\ \gamma_{RP}(\underline{h}) & \gamma_{RR}(\underline{h}) \end{pmatrix} = \begin{pmatrix} 0.20 & 0.17 \\ 0.17 & 0.20 \end{pmatrix} \gamma_0(\underline{h}) + \begin{pmatrix} 0.90 & 0.65 \\ 0.65 & 0.66 \end{pmatrix} \gamma_1(\underline{h})$$

donde $\gamma_0(\underline{h})$ es el modelo nugget, y $\gamma_1(\underline{h})$ es el modelo esférico con alcance 20 Km.

Se puede observar que el modelo es válido, ya que los determinantes son positivos:

$$\det \begin{pmatrix} 0.20 & 0.17 \\ 0.17 & 0.20 \end{pmatrix} = 0.011 > 0, \quad \det \begin{pmatrix} 0.90 & 0.65 \\ 0.65 & 0.66 \end{pmatrix} = 0.172 > 0$$

Validación del modelo de semivariograma cruzado

- Consiste en estimar por Cokriging los valores en los puntos muestrales usando el procedimiento de *leave one out*.
- Con los valores estimados y sus correspondientes varianzas de la estimación se calculan los criterios convencionales de la validación cruzada para una variable (error medio, error cuadrático medio, etc,...)
- Primero validar los semivariogramas simples por separado y luego los cruzados de manera conjunta.

Cokriging Ordinario

- Sistema de ecuaciones:

$$\left\{ \begin{array}{l} \sum_{j=1}^n \underline{\underline{\Lambda}}^j \underline{\underline{C}}(\underline{x}_i - \underline{x}_j) + \underline{\underline{M}} = \underline{\underline{C}}(\underline{x}_i - \underline{x}) \\ \sum_{j=1}^n \underline{\underline{\Lambda}}^j = \underline{\underline{I}}, \quad i = 1, \dots, n \end{array} \right.$$

- Estimador:

$$Z_1^*(\underline{x}) = \sum_{j=1}^n \left(\lambda_{11}^j Z_1(\underline{x}_j) + \lambda_{12}^j Z_2(\underline{x}_j) \right)$$

$$Z_2^*(\underline{x}) = \sum_{j=1}^n \left(\lambda_{21}^j Z_1(\underline{x}_j) + \lambda_{22}^j Z_2(\underline{x}_j) \right)$$

Cokriging Ordinario

- Varianza total de la estimación

$$\sigma_{CK}^2 = \text{Tr } \underline{\underline{C}}(\underline{\underline{0}}) - \text{Tr } \sum_{j=1}^n \underline{\underline{\Lambda}}^j \underline{\underline{C}}(\underline{\underline{x}} - \underline{\underline{x}}_j) - \text{Tr } \underline{\underline{M}}$$

la cual representa una varianza acumulada

- Varianza de la estimación de cada variable

$$\sigma_{CK1}^2 = C_{11}(\underline{\underline{0}}) - \sum_{j=1}^n \lambda_{11}^j C_{11}(\underline{\underline{x}} - \underline{\underline{x}}_j) + \sum_{j=1}^n \lambda_{21}^j C_{21}(\underline{\underline{x}} - \underline{\underline{x}}_j) - \mu_{11}$$

$$\sigma_{CK2}^2 = C_{22}(\underline{\underline{0}}) - \sum_{j=1}^n \lambda_{12}^j C_{12}(\underline{\underline{x}} - \underline{\underline{x}}_j) + \sum_{j=1}^n \lambda_{22}^j C_{22}(\underline{\underline{x}} - \underline{\underline{x}}_j) - \mu_{22}$$

Cokriging Ordinario

- Ecuaciones del Cokriging en forma matricial

$$\begin{bmatrix} \underline{\underline{C}}(\underline{x}_1 - \underline{x}_1) & \dots & \underline{\underline{C}}(\underline{x}_1 - \underline{x}_n) & \underline{\underline{I}} \\ \dots & \dots & \dots & \dots \\ \underline{\underline{C}}(\underline{x}_n - \underline{x}_1) & \dots & \underline{\underline{C}}(\underline{x}_n - \underline{x}_n) & \underline{\underline{I}} \\ \underline{\underline{I}} & \dots & \underline{\underline{I}} & \underline{\underline{0}} \end{bmatrix} \begin{bmatrix} \underline{\underline{\Lambda}}^1 \\ \dots \\ \underline{\underline{\Lambda}}^n \\ \underline{\underline{M}} \end{bmatrix} = \begin{bmatrix} \underline{\underline{C}}(\underline{x}_1 - \underline{x}) \\ \dots \\ \underline{\underline{C}}(\underline{x}_n - \underline{x}) \\ \underline{\underline{I}} \end{bmatrix}$$

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

$$\underline{\underline{\Lambda}}^k = \begin{bmatrix} \lambda_{11}^k & \lambda_{12}^k \\ \lambda_{21}^k & \lambda_{22}^k \end{bmatrix}, \quad \underline{\underline{I}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \underline{\underline{M}} = \begin{bmatrix} \mu_{11} & \mu_{12} \\ \mu_{21} & \mu_{22} \end{bmatrix}$$

Cokriging Ordinario

- Observaciones prácticas:
- La matriz de coeficientes es simétrica a pesar de que la asimetría de $\underline{C}(\underline{x} - \underline{y})$
- Todas las entradas son invertibles (excepto $\underline{0}$) de forma tal que puede ser reducida a una matriz triangular mediante la operación con matrices de menor dimensión y así simplificar el cómputo.
- Cuando las FAs no están correlacionadas entonces el sistema se convierte en n sistemas de ecuaciones de Kriging separados.

Cokriging en el caso de funciones aleatorias intrínsecas

- La hipótesis intrínseca es:

$$E \left[Z_i(\underline{x} + \underline{h}) - Z_i(\underline{x}) \right] = 0, \quad i = 1, 2$$

$$\text{Cov} \left[Z_i(\underline{x} + \underline{h}) - Z_i(\underline{x}), Z_j(\underline{x} + \underline{h}) - Z_j(\underline{x}) \right] = 2\gamma_{ij}(\underline{h}), \quad i, j = 1, 2$$

- Entonces la matriz de semivarianzas cruzadas sustituye en el sistema de ecuaciones del cokriging a la matriz de covarianzas.

$$\underline{\underline{\gamma}}(\underline{x} - \underline{y}) \leftrightarrow \underline{\underline{C}}(\underline{x} - \underline{y})$$

- El sistema de ecuaciones resultante es completamente análogo al caso con estacionaridad de segundo orden

Cokriging en el caso no estacionario

- Cuando las FAs son no estacionarias, entonces las ecuaciones del Cokriging de la sección anterior pueden ser extendidas de forma análoga al Kriging Universal para una variable.
- Pero como en el caso de una variable resulta poco práctico su aplicación debido al no conocimiento *a priori* de los órdenes de las tendencias y los modelos de los covariogramas.
- Es preferible aplicar un enfoque del tipo Kriging Residual a cada FA por separado y luego aplicar el Cokriging Ordinario a los residuos en conjunto.

Ejemplos de Aplicaciones del Cokriging

➤ Estimación de una combinación lineal de FAs

Existen dos enfoques posibles:

- **Estimación directa:** Se toma el 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.
 - **Estimación conjunta:** 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 de manera conjunta con Cokriging.
- ## ➤ 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.

Dificultades y Aspectos Prácticos del Cokriging

- Estimar varias variables correogionalizadas simultáneamente usando el Cokriging es el enfoque más riguroso y el que se basa en un menor número de hipótesis.
- 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.

Dificultades y Aspectos Prácticos del Cokriging

- 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
- Exige un tiempo de cálculo considerable en la modelación de los semivariogramas cruzados mediante la modelación y validación de múltiples semivariogramas
- Aumenta la complejidad y el tamaño de los sistemas de ecuaciones a resolver

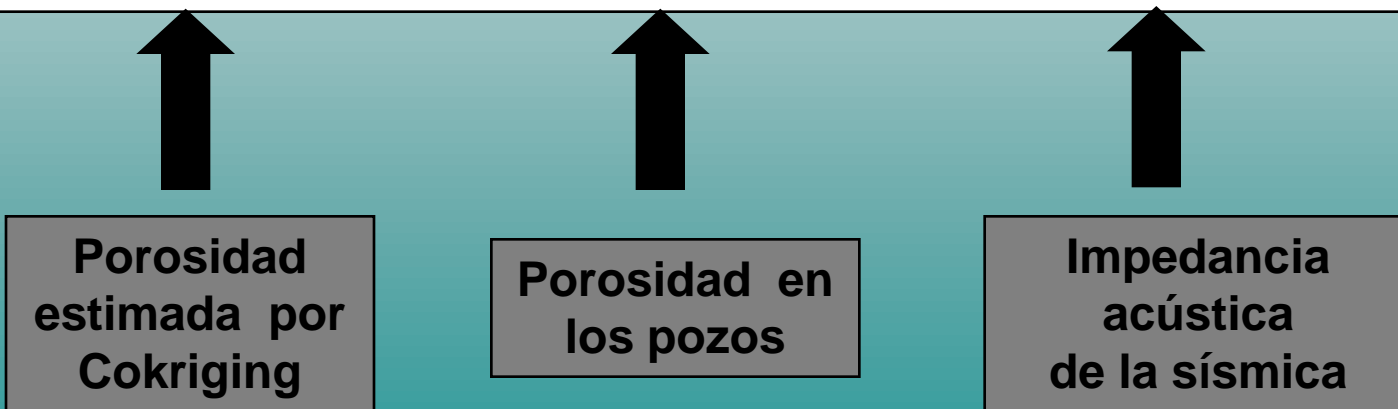
Cokriging Colocado

- Es un caso particular del Cokriging
- La variable de interés es conocida en unos pocos puntos y la variable auxiliar es conocida en todos los puntos de la malla de estimación
- La vecindad de la variable auxiliar es reducida a un solo punto: el punto de estimación
- No se requiere del conocimiento del modelo de correogionalización lineal sino del coeficiente de correlación entre las variables.
- Es computacionalmente más simple y eficiente comparado con el Cokriging

Cokriging Colocado

- Es un caso particular de Cokriging
- Ejemplo: Dos variables $Z_1(\underline{x})$ y $Z_2(\underline{x})$ (porosidad & impedancia acústica)

$$Z_1^* (\underline{x}_0) = \sum_{j=1}^n \left(\lambda_{11}^j Z_1 (\underline{x}_j) + \lambda_{12}^j Z_2 (\underline{x}_j) \right)$$



Cokriging Colocado

- **COKRIGING**
 - Sistema de ecuaciones grande
 - Requiere variogramas de Z_1 , Z_2 , variograma cruzado de Z_1 y Z_2 (Modelo de Correogionalización Lineal)
- **COKRIGING COLOCADO**
 - Sistema de ecuaciones mas simple
 - No requiere Modelo de Correogionalización Lineal
 - Sólo variogramas de Z_1 , Z_2 y coeficiente de correlación

$$Z_1^* (\underline{x}_0) = \sum_{j=1}^n \lambda_{11}^j Z_1 (\underline{x}_j) + \lambda_{12}^0 Z_2 (\underline{x}_0)$$

Métodos alternativos al Cokriging

- **Kriging combinado con Regresión Lineal:** (Delhomme,1976) Consiste en establecer un modelo de regresión lineal entre dos FAs Z y Y

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

- Aplicando la regresión se puede estimar la variable Z en los puntos donde hay valores muestrales para la otra variable Y

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

- Luego se aplica el procedimiento del Kriging a cada una por separado, el cual es computacionalmente mucho más sencillo que el Cokriging.

Métodos alternativos al Cokriging

- **Kriging con una Deriva Externa:** Se requiere que el valor esperado de una FA Z sea una función lineal conocida dependiente de otra FA Y , como sigue:

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

- Se necesita que la segunda variable Y haya sido muestreada en un gran número de puntos.
- Entonces se puede aplicar un Kriging con deriva

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