
Capítulo Cuatro

Predicción Espacial

4.1. Predicción Espacial Óptima.

De la teoría de la decisión se conoce que si Z_0 es una cantidad aleatoria y Z_0^* es su predictor ², entonces $L(Z_0; Z_0^*)$ representa la pérdida en que se incurre cuando se predice Z_0 con Z_0^* y el mejor predictor será el que minimice $E\{L(Z_0; Z_0^*)/Z\}$ con $Z = \{Z_1, Z_2, \dots, Z_n\}'$, es decir el predictor óptimo es el que minimice la *esperanza condicional* de la función de pérdida. Si $L(Z_0; Z_0^*) = [Z_0 - Z_0^*]^2 \Rightarrow Z_0^* = E(Z_0 / Z)$. La expresión anterior indica que para encontrar el predictor óptimo se requiere conocer la distribución conjunta de la $n+1$ variables aleatorias.

4.2. Definición de Kriging.

La palabra kriging³ (expresión anglosajona) procede del nombre del geólogo sudafricano D. G. Krige, cuyos trabajos en la predicción de reservas de oro, realizados en la década del cincuenta, suelen considerarse como pioneros en los métodos de interpolación espacial. Kriging encierra un conjunto de métodos de predicción espacial que se fundamentan en la minimización del error cuadrático medio de predicción. En la tabla 5 se mencionan los tipos de kriging y algunas de sus propiedades. En la secciones 4.3 y 4.4, se hace una presentación detallada de ellos.

Tabla 5. Tipos de predictores kriging y sus propiedades.

TIPO DE PREDICTOR	NOMBRE	PROPIEDADES
LINEAL	<ul style="list-style-type: none">• Simple• Ordinario• Universal	<ul style="list-style-type: none">• Son óptimos si hay normalidad multivariada.• Independiente de la distribución son los mejores predictores linealmente insesgados.
NO LINEAL	<ul style="list-style-type: none">• Indicador• Probabilístico• Log Normal, Trans-Gaussiano• Disyuntivo	<ul style="list-style-type: none">• Son predictores óptimos.

² La palabra *estimación* es utilizada exclusivamente para inferir sobre parámetros fijos pero desconocidos; *predicción* es reservada para inferencia sobre cantidades aleatorias.

³ Algunos textos indican que en español la palabra adecuada sería krigeado.

Los métodos kriging se aplican con frecuencia con el propósito de predicción, sin embargo estas metodologías tienen diversas aplicaciones, dentro de las cuales se destacan la simulación y el diseño de redes óptimas de muestreo (capítulo 5).

4.3. Kriging Ordinario

Suponga que se hacen mediciones de la variable de interés Z en los puntos x_i , $i= 1, 2, \dots, n$, de la región de estudio, es decir se tienen realizaciones de las variables $Z(x_1), \dots, Z(x_n)$, y se desea predecir $Z(x_0)$, en el punto x_0 donde no hubo medición. En esta circunstancia, el método kriging ordinario propone que el valor de la variable puede predecirse como una combinación lineal de las n variables aleatorias así:

$$\begin{aligned} Z^*(x_0) &= \lambda_1 Z(x_1) + \lambda_2 Z(x_2) + \lambda_3 Z(x_3) + \lambda_4 Z(x_4) + \lambda_5 Z(x_5) + \dots + \lambda_n Z(x_n) \\ &= \sum_{i=1}^n \lambda_i Z(x_i) \end{aligned}$$

en donde los λ_i representan los pesos o ponderaciones de los valores originales. Dichos pesos se calculan en función de la distancia entre los puntos muestreados y el punto donde se va a hacer la correspondiente predicción. La suma de los pesos debe ser igual a uno para que la esperanza del predictor sea igual a la esperanza de la variable. Esto último se conoce como el requisito de insesgamiento.

Estadísticamente la propiedad de insesgamiento se expresa a través de:

$$E(Z^*(x_0)) = E(Z(x_0))$$

Asumiendo que el proceso es estacionario de media m (desconocida) y utilizando las propiedades del valor esperado, se demuestra que la suma de las ponderaciones debe ser igual a uno:

$$E\left(\sum_{i=1}^n \lambda_i Z(x_i)\right) = m$$

$$\sum_{i=1}^n \lambda_i E(Z(x_i)) = m$$

$$\sum_{i=1}^n \lambda_i m = m$$

$$m \sum_{i=1}^n \lambda_i = m \Rightarrow \sum_{i=1}^n \lambda_i = 1$$

Se dice que $Z^*(x_0)$ es el mejor predictor, lineal en este caso, porque los pesos se obtienen de tal manera que minimicen la varianza del error de predicción, es decir que minimicen la expresión:

$$V(Z^*(x_0) - Z(x_0))$$

Esta última es la característica distintiva de los métodos kriging, ya que existen otros métodos de interpolación como el de distancias inversas o el poligonal, que no garantizan varianza mínima de predicción (Samper y Carrera, 1990). La estimación de los pesos se

obtiene minimizando $V[Z^*(x_0) - Z(x_0)]$ sujeto a $\sum_{i=1}^n \lambda_i = 1$.

Se tiene que $V[Z^*(x_0) - Z(x_0)] = V[Z^*(x_0)] - 2COV[Z^*(x_0), Z(x_0)] + V[Z(x_0)]$

Desagregando las componentes de la ecuación anterior se obtiene lo siguiente:

$$V[Z^*(x_0) - Z(x_0)] = V\left[\sum_{i=1}^n \lambda_i Z(x_i)\right] = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j COV[Z(x_i), Z(x_j)]$$

En adelante se usará la siguiente notación: $COV[Z(x_i), Z(x_j)] = C_{ij}$ y $V[Z(x_0)] = \sigma^2$

$$\begin{aligned} \text{De lo anterior } COV[Z^*(x_0), Z(x_0)] &= COV\left[\sum_{i=1}^n \lambda_i Z(x_i), Z(x_0)\right] \\ &= \sum_{i=1}^n \lambda_i COV[Z(x_i), Z(x_0)] = \sum_{i=1}^n \lambda_i C_{i0} \end{aligned}$$

Entonces reemplazando, se tiene que:

$$V[Z^*(x_0) - Z(x_0)] = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j C_{ij} - 2 \sum_{i=1}^n \lambda_i C_{i0} + \sigma^2 \quad (0)$$

Luego se debe minimizar la función anterior sujeta a la restricción $\sum_{i=1}^n \lambda_i = 1$. Este problema de minimización con restricciones se resuelve mediante el método de multiplicadores de Lagrange.

$$\sigma_k^2 = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j C_{ij} - 2 \sum_{i=1}^n \lambda_i C_{i0} + \sigma^2 + \underbrace{2\mu}_{\text{Multiplicador de Lagrange}} \underbrace{\left(\sum_{i=1}^n \lambda_i - 1\right)}_0$$

Siguiendo el procedimiento acostumbrado para obtener valores extremos de una función, se deriva e iguala a cero, en este caso con respecto a λ_i y μ :

$$\begin{aligned} \frac{\partial(\sigma_k^2)}{\partial \lambda_1} &= \frac{\partial \left[(\lambda_1^2 C_{11} + 2\lambda_1 \sum_{j=2}^n \lambda_j C_{1j} + \sum_{i=2}^n \sum_{j=1}^n \lambda_i \lambda_j C_{ij}) - 2 \sum_{i=1}^n \lambda_i C_{i0} + \sigma^2 + 2\mu \left(\sum_{i=1}^n \lambda_i - 1 \right) \right]}{\partial \lambda_1} \\ &= \underbrace{\left(2\lambda_1 C_{11} + 2 \sum_{j=2}^n \lambda_j C_{1j} \right)}_{\downarrow} - 2C_{10} + 2\mu \\ &= 2 \sum_{j=1}^n \lambda_j C_{1j} - 2C_{10} + 2\mu = 0 \Rightarrow \sum_{j=1}^n \lambda_j C_{1j} + \mu = C_{10} \quad (1) \end{aligned}$$

De manera análoga se determinan las derivadas con respecto a $\lambda_2, \dots, \lambda_n$:

$$\frac{\partial(\sigma_k^2)}{\partial \lambda_2} = 2 \sum_{j=1}^n \lambda_j C_{2j} - 2C_{20} + 2\mu = 0 \Rightarrow \sum_{j=1}^n \lambda_j C_{2j} + \mu = C_{20} \quad (2)$$

\vdots

$$\frac{\partial(\sigma_k^2)}{\partial \lambda_n} = 2 \sum_{j=1}^n \lambda_j C_{nj} - 2C_{n0} + 2\mu = 0 \Rightarrow \sum_{j=1}^n \lambda_j C_{nj} + \mu = C_{n0} \quad (3)$$

por último derivamos con respecto a μ :

$$\frac{\partial(\sigma_k^2)}{\partial \mu} = 2 \sum_{i=1}^n \lambda_i - 2 = 0 \Rightarrow \sum_{i=1}^n \lambda_i = 1 \quad (4)$$

De (1), (2), (3), (4) resulta un sistema de $(n + 1)$ ecuaciones con $(n + 1)$ incógnitas, que matricialmente puede ser escrito como:

$$\begin{pmatrix} C_{11} & \cdot & \cdot & \cdot & C_{1n} & 1 \\ \cdot & \cdot & & & \cdot & \cdot \\ \cdot & & \cdot & & \cdot & \cdot \\ \cdot & & & \cdot & \cdot & \cdot \\ C_{n1} & \cdot & \cdot & \cdot & C_{nn} & 1 \\ 1 & \cdot & \cdot & \cdot & 1 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \cdot \\ \cdot \\ \cdot \\ \lambda_n \\ \mu \end{pmatrix} = \begin{pmatrix} C_{10} \\ \cdot \\ \cdot \\ \cdot \\ C_{n0} \\ 1 \end{pmatrix}$$

$$C_{ij} \bullet \lambda = C_{i0}$$

por lo cual los pesos que minimizan el error de predicción se determinan mediante la función de covariograma a través de:

$$\lambda = C_{ij}^{-1} \bullet C_{i0}.$$

Encontrando los pesos se calcula la predicción en el punto x_o . De forma análoga se procede para cada punto donde se quiera hacer predicción.

• Varianza de Predicción del Kriging Ordinario

Multiplicando (1), (2) y (3) por λ_i se obtiene:

$$\lambda_i \left(\sum_{j=1}^n \lambda_j C_{ij} + \mu \right) = \lambda_i C_{i0} \quad \forall i, i = 1, 2, \dots, n.$$

Sumando las n ecuaciones

$$\sum_{i=1}^n \lambda_i \sum_{j=1}^n \lambda_j C_{ij} + \sum_{i=1}^n \lambda_i \mu = \sum_{i=1}^n \lambda_i C_{i0}$$

$$\sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j C_{ij} = \sum_{i=1}^n \lambda_i C_{i0} - \sum_{i=1}^n \lambda_i \mu$$

Sustituyendo la expresión anterior en (0)

$$\sigma_k^2 = \sigma^2 + \sum_{i=1}^n \lambda_i C_{i0} - \sum_{i=1}^n \lambda_i \mu - 2 \sum_{i=1}^n \lambda_i C_{i0}$$

$$\sigma_k^2 = \sigma^2 - \sum_{i=1}^n \lambda_i C_{i0} - \mu \quad (5)$$

• **Estimación de Ponderaciones por medio de la Función de Semivarianza**

Los pesos λ pueden ser estimados a través de la función de semivarianza, para lo cual se requiere conocer la relación entre las funciones de covariograma y de semivarianza. Antes de esto conveniente tener en cuenta la siguiente notación:

$\sigma^2 = V(Z(x))$, $\gamma_{ij} = \gamma(h)$, donde h es la distancia entre los puntos i y j y análogamente $C_{ij} = C(h)$.

La relación entre las dos funciones en cuestión es la siguiente:

$$\begin{aligned} \gamma_{ij} &= \frac{1}{2} E[(Z(x_j) - Z(x_i))^2] \\ &= \frac{1}{2} E[(Z(x_j))^2 - 2Z(x_j)Z(x_i) + (Z(x_i))^2] \\ &= \frac{1}{2} E[(Z(x_j))^2] - E[Z(x_j)Z(x_i)] + \frac{1}{2} E[(Z(x_i))^2] \\ &= \frac{1}{2} [E(Z(x_j))^2 - k^2] + \frac{1}{2} [E(Z(x_i))^2 - k^2] - [E(Z(x_j)Z(x_i)) - k^2] \\ &= \frac{1}{2} [V(Z(x)) + \mu^2] + \frac{1}{2} [V(Z(x)) + \mu^2] - [COV[Z(x_j)Z(x_i)] + \mu^2] \\ &= V[Z(x)] - COV[Z(x_j)Z(x_i)] \\ &= \sigma^2 - C_{ij} \Rightarrow C_{ij} = \sigma^2 - \gamma_{ij} \quad (6) \end{aligned}$$

Reemplazando (6) en (1), (2) y (3) se determinan los pesos óptimos λ en términos de la función de semivarianza:

$$\begin{aligned} \frac{\partial(\sigma_k^2)}{\partial \lambda_1} &= \sum_{j=1}^n \lambda_j C_{1j} + \mu - C_{10} = \sum_{j=1}^n \lambda_j (\sigma^2 - \gamma_{1j}) + \mu - (\sigma^2 - \gamma_{10}) \\ &= \sigma^2 \sum_{j=1}^n \lambda_j - \sum_{j=1}^n \lambda_j \gamma_{1j} + \mu - \sigma^2 + \gamma_{10} \\ &= \sigma^2 - \sum_{j=1}^n \lambda_j \gamma_{1j} + \mu - \sigma^2 + \gamma_{10} \Rightarrow \sum_{j=1}^n \lambda_j \gamma_{1j} - \mu = \gamma_{10} \end{aligned}$$

Similarmente,

$$\frac{\partial(\sigma_k^2)}{\partial\lambda_2} = \sum_{j=1}^n \lambda_j \gamma_{2j} - \mu = \gamma_{20}$$

\vdots

$$\frac{\partial(\sigma_k^2)}{\partial\lambda_n} = \sum_{j=1}^n \lambda_j \gamma_{nj} - \mu = \gamma_{n0}$$

El sistema de ecuaciones se completa con (4). De acuerdo con lo anterior los pesos se obtienen en términos del semivariograma a través del sistema de ecuaciones:

$$\begin{pmatrix} \gamma_{11} & . & . & . & \gamma_{1n} & 1 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ \gamma_{n1} & . & . & . & \gamma_{nn} & 1 \\ 1 & . & . & . & 1 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ . \\ . \\ . \\ \lambda_n \\ -\mu \end{pmatrix} = \begin{pmatrix} \gamma_{10} \\ . \\ . \\ . \\ \gamma_{n0} \\ 1 \end{pmatrix}$$

Para establecer la expresión de la correspondiente varianza del error de predicción en términos de la función de semivarianza se reemplaza (6) en (5), de donde:

$$\sigma_k^2 = \sigma^2 - \left[\sum_{i=1}^n \lambda_i (\sigma^2 - \gamma_{ij}) \right] + \mu$$

$$\sigma_k^2 = \sigma^2 - \sigma^2 \sum_{i=1}^n \lambda_i + \sum_{i=1}^n \lambda_i \gamma_{ij} + \mu$$

$$\sigma_k^2 = \sum_{i=1}^n \lambda_i \gamma_{io} + \mu$$

Los pesos de kriging ordinario también pueden ser estimados mediante el uso del correlograma aplicando la siguiente relación: $\rho_{ij} = C_{ij} / \sigma^2$, caso en el que la correspondiente varianza de predicción estaría dada por (Isaaks y Srivastava, 1989):

$$\sigma_k^2 = \sigma^2 \left(1 - \sum \lambda_i \gamma_{io} + \mu \right)$$

- **Validación del kriging.**

Existen diferentes métodos para evaluar la bondad de ajuste del modelo de semivariograma elegido con respecto a los datos muestrales y por ende de las predicciones hechas con kriging. El más empleado es el de validación cruzada, que consiste en excluir la observación de uno de los n puntos muestrales y con los $n-1$ valores restantes y el modelo

de semivariograma escogido, predecir vía kriging el valor de la variable en estudio en la ubicación del punto que se excluyó. Se piensa que si el modelo de semivarianza elegido describe bien la estructura de autocorrelación espacial, entonces la diferencia entre el valor observado y el valor predicho debe ser pequeña. Este procedimiento se realiza en forma secuencial con cada uno de los puntos muestrales y así se obtiene un conjunto de n “errores de predicción”. Lo usual es calcular medidas que involucren a estos errores de predicción para diferentes modelos de semivarianza y seleccionar aquel que optimice algún criterio como por ejemplo el del mínimo error cuadrático medio (MECM). Este procedimiento es similar a la conocida técnica de remuestreo Jackknife (Efron, 1982) empleada en diversos contextos estadísticos para calcular varianzas de estimación, entre otros aspectos. Una forma descriptiva de hacer la validación cruzada es mediante un gráfico de dispersión de los valores observados contra los valores predichos. En la medida en que la nube de puntos se ajuste más a una línea recta que pase por el origen, mejor será el modelo de semivariograma utilizado para realizar el kriging.

- **Representación de las predicciones**

Una vez se ha hecho la predicción en un conjunto de puntos diferentes de los muestrales vía kriging, se debe elaborar un mapa que dé una representación global del comportamiento de la variable de interés en la zona estudiada. Los más empleados son los mapas de contornos, los mapas de residuos y los gráficos tridimensionales. En el caso de los mapas de contornos, en primer lugar se divide el área de estudio en un enmallado y se hace la predicción en cada uno de los nodos de éste mismo. Posteriormente se unen los valores predichos con igual valor, generando así las líneas de contorno (isolíneas de distribución). Este gráfico permite identificar la magnitud de la variable en toda el área de estudio. Es conveniente acompañar el mapa de interpolaciones de la variable con los correspondientes mapas de isolíneas de los errores y de las varianzas de predicción (posiblemente estimados a través de métodos matemáticos), con el propósito de identificar zonas de mayor incertidumbre respecto a las predicciones.

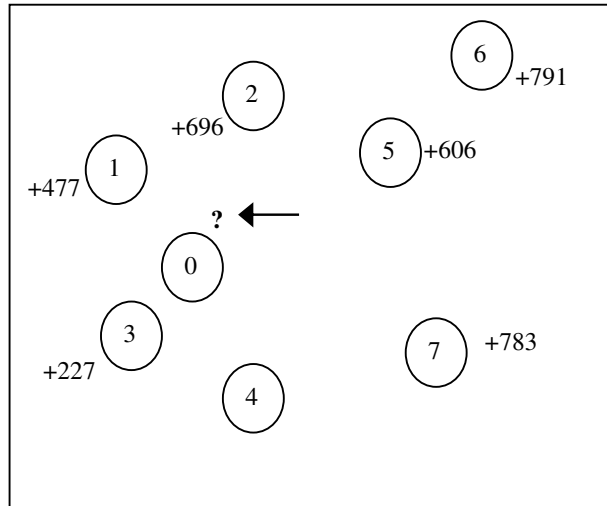
- **Intervalos de Confianza.**

Asumiendo que los errores de predicción siguen una distribución normal estándar y que son independientes, un intervalo de confianza del $100(1-\alpha)\%$, $0 < \alpha < 1$, para $Z(x)$ es:

$\left[z^*(x) - z_{1-\alpha/2} \sigma_k, z^*(x) + z_{1-\alpha/2} \sigma_k \right]$ con $z^*(x)$ el valor calculado de la predicción y $z_{1-\alpha/2}$ el percentil de una normal estándar.

- **Ilustración**

Suponga que se tiene una configuración de datos como la que se presenta en el esquema de abajo. Con base en siete datos observados (valores al lado del signo + por fuera de los círculos numerados de 1 a 7) se quiere predecir un valor de la variable en el punto donde se encuentra el signo de interrogación, por fuera del círculo con el número cero.



La matriz de distancia euclidianas entre los sitios es la siguiente:

Sitio	0	1	2	3	4	5	6	7
0	0.00	4.47	3.61	8.06	4.49	6.71	8.94	13.45
1		0.00	2.24	10.44	13.04	10.05	12.17	17.80
2			0.00	11.05	13.00	8.00	10.05	16.97
3				0.00	4.12	13.04	15.00	11.05
4					0.00	12.37	13.93	7.00
5						0.00	2.24	12.65
6							0.00	13.15
7								0.00

Asumiendo que la estructura de correlación espacial entre los datos es estimada por un modelo exponencial $\gamma(h) = 10(1 - \exp(-3h/10))$ (pepita cero, meseta 10 y rango 10) o en términos de la función de autocovarianza por $C(h) = 10(\exp(-3h/10))$, se encuentran las siguientes matrices que permiten encontrar los pesos para la predicción:

$$C_{ij} = \begin{pmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} & C_{17} & 1 \\ C_{21} & C_{22} & C_{23} & C_{24} & C_{25} & C_{26} & C_{27} & 1 \\ C_{31} & C_{32} & C_{33} & C_{34} & C_{35} & C_{36} & C_{37} & 1 \\ C_{41} & C_{42} & C_{43} & C_{44} & C_{45} & C_{46} & C_{47} & 1 \\ C_{51} & C_{52} & C_{53} & C_{54} & C_{55} & C_{56} & C_{57} & 1 \\ C_{61} & C_{62} & C_{63} & C_{64} & C_{65} & C_{66} & C_{67} & 1 \\ C_{71} & C_{72} & C_{73} & C_{74} & C_{75} & C_{76} & C_{77} & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 10 & 5.11 & 0.44 & 0.20 & 0.49 & 0.26 & 0.05 & 1 \\ 5.11 & 10 & 0.36 & 0.20 & 0.91 & 0.49 & 0.06 & 1 \\ 0.44 & 0.36 & 10 & 2.90 & 0.20 & 0.11 & 0.36 & 1 \\ 0.20 & 0.20 & 2.90 & 10 & 0.24 & 0.15 & 1.22 & 1 \\ 0.49 & 0.91 & 0.20 & 0.24 & 10 & 5.11 & 0.22 & 1 \\ 0.26 & 0.49 & 0.11 & 0.15 & 5.11 & 10 & 0.19 & 1 \\ 0.05 & 0.06 & 0.36 & 1.22 & 0.22 & 0.19 & 10 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \end{pmatrix}$$

$$C_{ij}^{-1} = \begin{pmatrix} 0.127 & -0.077 & -0.013 & -0.009 & -0.008 & -0.009 & -0.012 & 0.136 \\ -0.077 & 0.129 & -0.010 & -0.008 & -0.015 & -0.008 & -0.011 & 0.121 \\ -0.013 & -0.010 & 0.098 & -0.042 & -0.010 & -0.010 & -0.014 & 0.156 \\ 0.009 & -0.008 & -0.042 & 0.102 & -0.009 & -0.009 & -0.024 & 0.139 \\ -0.008 & -0.015 & -0.010 & -0.009 & 0.130 & -0.077 & -0.012 & 0.118 \\ -0.009 & -0.008 & -0.010 & -0.009 & -0.077 & 0.126 & -0.013 & 0.141 \\ -0.012 & -0.011 & -0.014 & -0.024 & -0.012 & -0.013 & 0.085 & 0.188 \\ 0.136 & 0.121 & 0.156 & 0.139 & 0.118 & 0.141 & 0.188 & -2.180 \end{pmatrix}$$

$$C_{io} = \begin{pmatrix} C_{10} \\ C_{10} \\ C_{10} \\ C_{10} \\ C_{10} \\ C_{10} \\ C_{10} \\ 1 \end{pmatrix} = \begin{pmatrix} 2.61 \\ 3.39 \\ 0.89 \\ 0.58 \\ 1.34 \\ 0.68 \\ 0.18 \\ 1 \end{pmatrix}, \text{ de donde } \lambda = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_{41} \\ \lambda_5 \\ \lambda_6 \\ \lambda_7 \\ \mu \end{pmatrix} = C_{ij}^{-1} C_{io} = \begin{pmatrix} 0.173 \\ 0.318 \\ 0.129 \\ 0.086 \\ 0.151 \\ 0.057 \\ 0.086 \\ 0.907 \end{pmatrix}.$$

con base en el vector estimado de parámetros se encuentra que

$$Z_0^* = \sum_{i=1}^7 \lambda_i Z_i = (0.173)(477) + (0.318)(696) + \dots + (0.086)(0.18) = 592.$$

$$\text{con } \sigma_k^2 = \sigma^2 - \sum_{i=1}^7 \lambda_i C_{io} - \mu = 10 - [(0.173)(2.61) + \dots + (0.086)(0.18)] - 0.907$$

4.4. Otros Métodos Kriging

A continuación se mencionan algunos aspectos generales de otros métodos de predicción espacial. Un estudio riguroso de ellos puede hacerse en Cressie (1993), Deutsch y Journel (1998) y Samper y Carrera (1990).

4.4.1. Kriging Simple

Suponga que hay una variable regionalizada estacionaria con media (m) y covarianza conocidas. De manera análoga a como se define en modelos lineales (por ejemplo en diseño de experimentos) el modelo establecido en este caso es igual a la media más un error aleatorio con media cero. La diferencia es que en este caso los errores no son independientes.

Sea $Z(x)$ la variable de interés medida en el sitio x .

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

$$Z(x) = m + \varepsilon(x), \text{ con } E[\varepsilon(x)] = 0.$$

El predictor de la variable de interés en un sitio x_0 donde no se tiene información se define como:

$$Z^*(x_0) = m + \varepsilon^*(x_0),$$

con $\varepsilon^*(x_0)$ que corresponde a la predicción del error aleatorio en el sitio x_0 . Despejando de la ecuación anterior $\varepsilon^*(x_0) = Z^*(x_0) - m$.

El predictor del error aleatorio se define por:

$$\varepsilon^*(x_0) = \sum_{i=1}^n \lambda_i \varepsilon(x_i) = \sum_{i=1}^n \lambda_i (Z(x_i) - m).$$

de donde el predictor de la variable de estudio es:

$$Z^*(x_0) = m + \left[\sum_{i=1}^n \lambda_i (Z(x_i) - m) \right] = m + \sum_{i=1}^n \lambda_i \varepsilon(x_i)$$

El predictor es insesgado si:

$$E(Z^*(x_0)) = E(Z(x_0)) = m. \text{ Luego el predictor será insesgado cuando } E(\varepsilon^*(x_0)) = 0.$$

$$E(\varepsilon^*(x_0)) = \sum_{i=1}^n \lambda_i \varepsilon(x_i) = \sum_{i=1}^n \lambda_i (0) = 0. \text{ Por consiguiente, a diferencia del kriging ordinario,}$$

en este caso no existen restricciones para las ponderaciones tendientes al cumplimiento de la condición de insesgamiento.

La estimación de los pesos del método kriging ordinario se obtiene de tal forma que se minimice $V(\varepsilon^*(x_0) - \varepsilon(x_0))$.

$$\begin{aligned}
V(\varepsilon^*(x_0) - \varepsilon(x_0)) &= E(\varepsilon^*(x_0) - \varepsilon(x_0))^2 \\
&= E\left(\left(\sum_{i=1}^n \lambda_i \varepsilon(x_i)\right) - \varepsilon(x_0)\right)^2 \\
&= E\left[\left(\sum_{i=1}^n \lambda_i \varepsilon(x_i)\right)\left(\sum_{j=1}^n \lambda_j \varepsilon(x_j)\right) - 2E\left(\left(\sum_{i=1}^n \lambda_i \varepsilon(x_i)\right)(\varepsilon(x_0))\right) + E(\varepsilon(x_0))^2\right] \\
&= \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j E(\varepsilon(x_i) \varepsilon(x_j)) - 2 \sum_{i=1}^n \lambda_i E(\varepsilon(x_i) \varepsilon(x_0)) + E(\varepsilon(x_0))^2
\end{aligned}$$

usando:

- i. $E[\varepsilon(x_0)] = 0$
- ii. $E(\varepsilon(x_i) \varepsilon(x_j)) = COV(\varepsilon(x_i), \varepsilon(x_j)) = C_{ij}$
- iii. $E(\varepsilon(x_0))^2 = \sigma^2$

$$V(\varepsilon^*(x_0) - \varepsilon(x_0)) = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j C_{ij} - 2 \sum_{i=1}^n \lambda_i C_{i0} + \sigma^2 \quad (7)$$

derivando respecto a λ_1 se tiene:

$$\begin{aligned}
\frac{\partial V(\varepsilon^*(x_0) - \varepsilon(x_0))}{\partial \lambda_1} &= \frac{\partial}{\partial \lambda_1} \left(\lambda_1^2 C_{11} + 2\lambda_1 \sum_{j=2}^n \lambda_j C_{1j} + \sum_{i=2}^n \sum_{j=2}^n \lambda_i \lambda_j C_{ij} - 2\lambda_1 C_{10} - 2 \sum_{i=2}^n \lambda_i C_{i0} + \sigma^2 \right) \\
&= 2\lambda_1 C_{11} + 2 \sum_{j=2}^n \lambda_j C_{1j} - 2C_{10} \\
&= 2 \sum_{i=1}^n \lambda_i C_{1i} - 2C_{10}
\end{aligned}$$

igualando a cero

$$\sum_{i=1}^n \lambda_i C_{1i} = C_{10}.$$

En general para cualquier i , $i = 1, 2, \dots, n$, se obtiene:

$$\frac{\partial}{\partial \lambda_i} = \sum_{j=1}^n \lambda_j C_{ij} = C_{i0} \quad (8)$$

Con las n ecuaciones resultantes se construye el siguiente sistema de ecuaciones:

$$\begin{pmatrix} C_{11} & C_{12} & \cdots & C_{1n} \\ C_{21} & C_{22} & \cdots & C_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ C_{n1} & C_{n2} & \cdots & C_{nn} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} C_{10} \\ C_{20} \\ \vdots \\ C_{n0} \end{pmatrix}$$

- **Varianza de Predicción Kriging Simple.**

Se tiene de (7) que:

$$V(\varepsilon^*(x_0) - \varepsilon(x_0)) = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j C_{ij} - 2 \sum_{i=1}^n \lambda_i C_{i0} + \sigma^2$$

$$\sigma_k^2 = \sum_{i=1}^n \lambda_i \sum_{j=1}^n \lambda_j C_{ij} - 2 \sum_{i=1}^n \lambda_i C_{i0} + \sigma^2$$

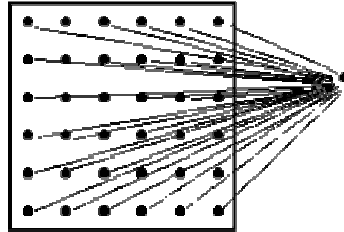
reemplazando (8) en (7)

$$\sigma_k^2 = \sum_{i=1}^n \lambda_i C_{i0} - 2 \sum_{i=1}^n \lambda_i C_{i0} + \sigma^2$$

$$\sigma_k^2 = \sigma^2 - \sum_{i=1}^n \lambda_i C_{i0}$$

4.4.2. Kriging en Bloques.

En los dos métodos kriging hasta ahora descritos el objetivo ha estado centrado en la predicción puntual. A menudo, sin embargo, se requiere estimar un bloque, o más precisamente, estimar el valor promedio de la variable dentro de un área local.



El valor promedio dentro del bloque es estimado por :

$$\bar{Z}(A) = \sum_{i=1}^n \lambda_i Z(x_i)$$

Del sistema de ecuaciones para el kriging ordinario se tiene:

$$\begin{pmatrix} C_{11} & . & . & . & C_{1n} & 1 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ C_{n1} & . & . & . & C_{nn} & 1 \\ 1 & . & . & . & 1 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ . \\ . \\ . \\ \lambda_n \\ \mu \end{pmatrix} = \begin{pmatrix} C_{10} \\ . \\ . \\ . \\ C_{n0} \\ 1 \end{pmatrix}$$

Consecuentemente el vector del lado derecho de la igualdad en el sistema de arriba debe modificarse para incluir las covarianzas respecto al bloque. La covarianza de un punto al bloque corresponde a la covarianza promedio entre el punto muestreado i y todos los puntos dentro del bloque (en la práctica un enmallado regular de puntos dentro del bloque es usado como se muestra en la figura de la página anterior). El sistema de ecuaciones del kriging en bloques está dado por:

$$\begin{pmatrix} C_{11} & . & . & . & C_{1n} & 1 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ C_{n1} & . & . & . & C_{nn} & 1 \\ 1 & . & . & . & 1 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ . \\ . \\ . \\ \lambda_n \\ \mu \end{pmatrix} = \begin{pmatrix} \bar{C}_{1A} \\ . \\ . \\ . \\ \bar{C}_{nA} \\ 1 \end{pmatrix}$$

donde el vector de covarianzas al lado derecho de la igualdad en el sistema anterior es contiene las covarianzas entre las variables $Z(x_1), Z(x_2), \dots, Z(x_n)$ y el bloque A donde se quiere hacer la estimación.

$$\bar{C}_{iA} = \frac{1}{|A|} \sum_{j \in A} C_{iA}.$$

La varianza del error de predicción del kriging en bloques está dada por:

$$\sigma_{kB}^2 = \bar{C}_{AA} - \left(\sum_{i=1}^n \lambda_i \bar{C}_{iA} + \mu \right), \text{ con } \bar{C}_{AA} = \frac{1}{|A|^2} \sum_{i \in A} \sum_{j \in A} C_{ij} \text{ igual a la covarianza entre}$$

pares de puntos dentro del bloque.

Isaaks y Srivastava (1989) muestran a través de ejemplos que el kriging en bloques coincide con el promedio de predicciones hechas por kriging ordinario sobre cada uno de los puntos del enmallado dentro del bloque. Así mismo indican que en la práctica es suficiente con un enmallado cuadrado (6x6) para obtener estimaciones estables en los bloques.

4.4.3. Kriging Universal.

En los supuestos hechos hasta ahora respecto a los métodos kriging se ha asumido que la variable regionalizada es estacionaria (al menos se cumple con la hipótesis intrínseca). En muchos casos, la variable no satisface estas condiciones y se caracteriza por exhibir una tendencia. Por ejemplo en hidrología los niveles piezométricos⁴ de una acuifero pueden mostrar una pendiente global en la dirección del flujo (Samper y Carrera, 1990). Para tratar este tipo de variables es frecuente descomponer la variable $Z(x)$ como la suma de la tendencia, tratada como una función determinística, más una componente estocástica estacionaria de media cero. Asuma que:

$$Z(x) = m(x) + \varepsilon(x)$$

$$\text{con } E(\varepsilon(x)) = 0, \quad V(\varepsilon(x)) = \sigma^2 \text{ y por consiguiente } E(Z(x)) = m(x).$$

La tendencia puede expresarse mediante:

$$m(x) = \sum_{l=1}^p a_l f_l(x)$$

donde las funciones $f_l(x)$ son conocidas y p es el número de términos empleados para ajustar $m(x)$.

El predictor kriging universal se define como:

$$Z^*(x_0) = \sum_{i=1}^n \lambda_i Z(x_i)$$

este será insesgado si:

$$E(Z^*(x_0)) = m(x_0)$$

$$E\left(\sum_{i=1}^n \lambda_i Z(x_i)\right) = m(x_0)$$

$$\left(\sum_{i=1}^n \lambda_i m(x_i)\right) = m(x_0)$$

$$\sum_{i=1}^n \lambda_i \left(\sum_{l=1}^p a_l f_l(x_i)\right) = \sum_{l=1}^p a_l f_l(x_0)$$

$$\sum_{l=1}^p a_l \left(\sum_{i=1}^n \lambda_i f_l(x_i)\right) = \sum_{l=1}^p a_l f_l(x_0) \quad \Rightarrow \quad \sum_{i=1}^n \lambda_i f_l(x_i) = \sum_{l=1}^p f_l(x_0)$$

⁴ Piezómetro: Instrumento utilizado para medir coeficientes de compresibilidad de sólidos, líquidos y gases

La obtención de los pesos en el kriging universal, análogo a los otros métodos kriging, se hace de tal forma que la varianza del error de predicción sea mínima.

$$\begin{aligned}
V(Z^*(x_0) - Z(x_0)) &= E(Z^*(x_0) - Z(x_0))^2 \\
&= E\left(\left(\sum_{i=1}^n \lambda_i (m(x_i) - \varepsilon(x_i))\right) - (m(x_0) - \varepsilon(x_0))\right)^2 \\
&= E\left[\left(\sum_{i=1}^n \lambda_i m(x_i) - m(x_0)\right) + \left(\sum_{i=1}^n \lambda_i \varepsilon(x_i) - \varepsilon(x_0)\right)\right]^2 \\
&= E\left[\left(\sum_{i=1}^n \lambda_i \varepsilon(x_i) - \varepsilon(x_0)\right)^2\right] \\
&= \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j E(\varepsilon(x_i) \varepsilon(x_j)) - 2 \sum_{i=1}^n \lambda_i E(\varepsilon(x_i) \varepsilon(x_0)) + E(\varepsilon(x_0))^2
\end{aligned}$$

Usando

$$C_{ij} = COV(\varepsilon(x_i), \varepsilon(x_j))$$

$$\sigma^2 = E(\varepsilon(x_0))^2$$

se tiene

$$V(Z^*(x_0) - Z(x_0)) = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j C_{ij} - 2 \sum_{i=1}^n \lambda_i C_{i0} + \sigma^2.$$

Luego incluyendo la restricción dada por la condición de insesgamiento, se debe minimizar:

$$\sigma_{ku}^2 = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j C_{ij} - 2 \sum_{i=1}^n \lambda_i C_{i0} + \sigma^2 + \sum_{l=1}^p \mu_l \left[\sum_{i=1}^n \lambda_i f_l(x_i) - f_l(x_0) \right]$$

o en términos de la función de semivarianza

$$\sigma_{ku}^2 = - \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \gamma_{ij} + 2 \sum_{i=1}^n \lambda_i \gamma_{i0} + \sum_{l=1}^p \mu_l \left[\sum_{i=1}^n \lambda_i f_l(x_i) - f_l(x_0) \right]$$

derivando la expresión anterior respecto a $\lambda_1, \lambda_2, \dots, \lambda_n, \mu_1, \mu_2, \dots, \mu_p$ e igualando a cero las correspondientes derivadas se obtienen las siguientes ecuaciones:

$$\begin{aligned}
\sum_{j=1}^n \lambda_j \gamma_{ij} + \sum_{l=1}^p \mu_l f_l(x_i) &= \gamma_{i0} \quad i = 1, 2, \dots, n \\
\sum_{j=1}^n \lambda_j f_l(x_j) &= f_l(x_0) \quad j = 1, 2, \dots, p
\end{aligned}$$

en términos matriciales

$$\begin{pmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1n} & f_{11} & \cdots & f_{p1} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2n} & f_{12} & \cdots & f_{p2} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \gamma_{n1} & \gamma_{n2} & \cdots & \gamma_{nn} & f_{1n} & \cdots & f_{pn} \\ f_{11} & f_{12} & \cdots & f_{1n} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ f_{p1} & f_{p2} & \cdots & f_{pn} & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} = \begin{pmatrix} \gamma_{10} \\ \gamma_{20} \\ \vdots \\ \gamma_{n0} \\ f_{10} \\ \vdots \\ f_{p0} \end{pmatrix}$$

donde $f_{ij} = f_i(x_j)$ es la i -ésima función en el punto j -ésimo.

La varianza de predicción del kriging universal está dada por (Samper y Carrera, 1990):

$$\sigma_{ku}^2 = \sum_{i=1}^n \lambda_i \gamma_{i0} + \sum_{l=1}^p \mu_l f_l(x_0).$$

Nótese que si $p = 1$ y $f_l(x) = 1$, el sistema de ecuaciones del kriging universal y la varianza de predicción coinciden con las del kriging ordinario. En este orden de ideas puede decirse que el kriging ordinario es un caso particular del kriging universal.

4.4.4. Kriging Residual.

La técnica kriging residuales empleada bajo las mismas circunstancias del kriging universal, es decir en aquellos casos en que la variable regionalizada no es estacionaria debido a la presencia de tendencia espacial en el valor promedio de la variable. La hipótesis central del kriging residual consiste en suponer conocida la tendencia $m(x)$. A partir de ella se calculan los residuos con base en los cuales se aplica kriging ordinario. La estimación de la tendencia es generalmente llevada a cabo por medio de mínimos cuadrados. La predicción en un sitio no muestreado es igual a la tendencia estimada más la predicción del error, es decir:

$$Z^*(x_0) = \hat{m}(x_0) + e^*(x_0)$$

$$e^*(x_0) = \sum_{i=1}^n \lambda_i e(x_i)$$

los pesos o ponderaciones son estimados por kriging ordinario como se muestra en la sección 4.2. La varianza de predicción de la variable de interés coincide con la varianza de predicción de los errores. En la figura 15 se muestra un esquema con el procedimiento kriging residual en el caso de una tendencia lineal.