Predicción Geoestadística: Kriging Simple, Ordinario y con Tendencia

Salvador Pintos. Applied Computing Institute, University of Zulia

Noviembre/2012

Resumen

Este reporte contiene una formulación alternativa basada en vectores aleatorios de los resultados fundamentales de predicción geoestadística. Para evitar los complejos sistemas lineales y el exceso de sumatorias que les están asociados -tal como aparecen en los textos habituales- se recurre al poder de síntesis del algebra lineal. Si bien su aplicación directa es en ciencias de la tierra, los resultados de Kriging se presentan para cualquier dimensión finita, permitiendo su uso en problemas de aproximación y optimización en (R^p) . El capítulo **1** es una introducción a procesos estocásticos y a los conceptos de continuidad espacial y estacionaridad. El capítulo **2** contiene los resultados fundamentales de Kriging Simple. El **3,** Kriging Ordinario; el **4,** Kriging con Tendencia. En el **apéndice** se incluye un resumen de algebra lineal y vectores aleatorios con los instrumentos necesarios para derivar los resultados fundamentales de Kriging. Pensando en aquel lector a quien las demostraciones le provocan pánico, éstas aparecen deliberadamente entre los símbolos ▼ y ▲ para que, en una primera lectura, pueda conocer los resultados y sus consecuencias sin necesidad de atascarse en el cómo.

Capítulo 1

Procesos estocásticos en ciencias de la tierra

Una de las características observables de las variables que se consideran en ciencias de la tierra, como la porosidad o la permeabilidad en la caracterización de yacimientos petrolíferos, es la de presentar patrones de continuidad espacial debido al proceso crono-depositacional. Por otra parte sólo se conocen escasas mediciones -directas o indirectas- de estas propiedades, por ejemplo, mediciones en los pozos.

Es así que la predición de estas variables en puntos arbitrarios del campo en estudio deberá tener en cuenta la incertidumbre intrínseca de esas predicciones y establecer una medida de la misma.

Figura 1.1: Imagen geológica

Los métodos de predicción deterministicos como interpolación inversa, el k-vecino próximo o splines hacen énfasis en la predicción; en cambio la propuesta geoestadística, Kriging -que se fundamenta en los patrones de continuidad- privilegia el control de la incertidumbre.

Antes de explicar los fundamentos de Kriging se presentarán algunos conceptos básicos de procesos estocásticos.

1.1. El paradigma estocástico

Sea D un subconjunto del plano, R^2 o del espacio R^3 y $z(x)$ una función (propiedad petrofísica, por ejemplo) objeto de estudio definida en D. El *paradigma estocástico* conciste en considerar a $z(x)$ como una realización de un proceso estocástico $Z(x, w)$. Se entiende que es un proceso estocástico si para cada $x_o \in D$, fijo, $Z(x_o, w)$ es una variable aleatoria, y para cada w_o fijo $Z(x, w_o)$ -*realización*- es una función de x. Más precisamente, existe w_k tal que la función de interés, $z(x)$, es la realización $z(x) = Z(x, w_k)$. En lo que sigue se notará z_k a la variable aleatoria $Z(x_k, w)$ salvo que sea necesario enfatizar la posición x_k y en cuyo caso se indicará el $z(x)$.

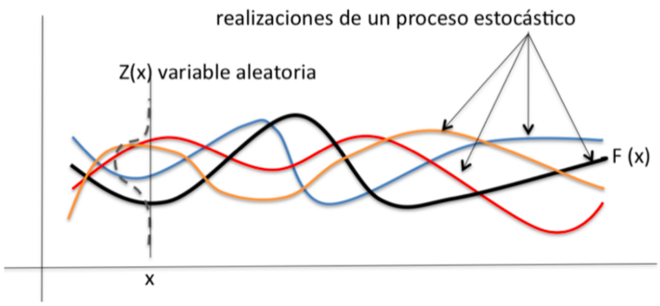

Figura 1.2: Realización de un proceso estocástico

En la Figura 1.2 una variable $F(x)$, por ejemplo saturación de agua, se la identifica con una realización de un proceso estocástico; para un x fijo la incertidumbre acerca del verdadero valor de la saturación queda contemplada en la variable aleatoria $Z(x)$.

1.1.1. La continuidad espacial

Lejos de ser un ruido blanco, estos procesos asociados a variables geofísicas presentan una alta correlación en puntos, x_1, x_2, cercanos del dominio, y la covarianza entre las variables aleatorias $Z(x_1, w)$ y $Z(x_2, w)$ asociadas,

$$Cov(Z(x_1, w),\ Z(x_2, w)) \quad para\, x_1,\ x_2\, arbitrarios \tag{1.1}$$

caracteriza la continuidad espacial. La posibilidad de predecir en un punto arbitrario a partir de una muestra escasa dada -tal como es común en ciencias de la tierra- y reducir la incertidumbre inherente al proceso estocástico depende del alcance de la covarianza.

1.1.2. Geoestadística

Si el dominio D es una región compacta y conexa del plano o el espacio, $(E = R^2,\ o\, R^3)$, como es el caso de una región de la tierra, entonces los procesos estocásticos fundamentales se les define como Campos Aleatorios (Random Fields) y son objeto de estudio de la Geoestadística. Otro ejemplo de procesos estocásticos son las Series de Tiempo donde el conjunto D es un conjunto de tiempos igualmente espaciados $D = \{t : \quad t = n \times t_o\}$.

1.2. Procesos estacionarios

Si la función de covarianza, ecu. 1.1, es conocida para cualquier par de puntos del dominio es posible realizar predicciones lineales óptimas tal como se presentan en los capítulos 2 y 3. Sin embargo, puesto que en general sólo se dispone de una muestra escasa de n pares $A = \{(z_k, x_k), \quad 1 \leq k \leq n\}$ de una realización $z(x)$, la inferencia de la estructura de covarianza ecu. 1.1, se fundamenta en una hipótesis adicional, *estacionaridad.*

Se define que el proceso $Z(x, w)$ es *estacionario* si para cualquier conjunto finito $\{x_1, ...x_j, ...x_m\}$ arbitrario la distribución conjunta de las variables aleatorias asociadas $\{z_1, ...z_j, ...z_m\}$ es invariante a traslaciones del conjunto $\{x_1, ...x_j, ...x_m\}$.

Más precisamente para cualquier vector h, los conjuntos $\{x_1, ...x_j, ...x_n\}$ y $\{x_1+h, ...x_j+h, .., x_n+h\}$ tienen sus variables aleatorias asociadas $\{z_1, ...z_j, ...z_n\}$ con idéntica distribución conjunta de probabilidad.

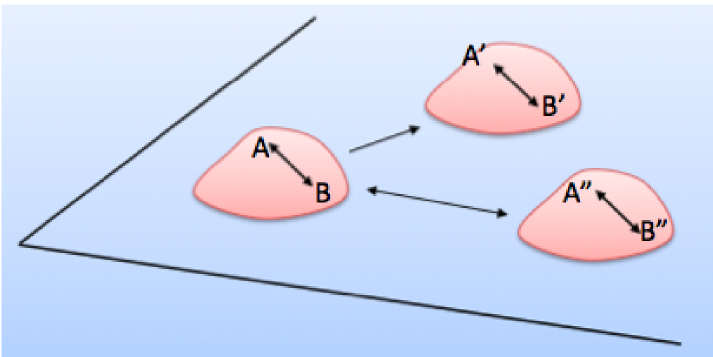

Figura 1.3: Proceso estacionario

La figura 1.3 presenta tres regiones que se corresponden en una traslación. Si el proceso es estacionario se cumple: que las medias y varianzas en puntos homólogos son iguales $E(Z(A)) = E(Z(A')) = E(Z(A"))$; $Var(Z(A)) = Var(Z(A')) = Var(Z(A")$; que las covarianzas entre pares homólogos son iguales: $Cov(\ Z(A),\ Z(B)) = Cov(\ Z(A'),\ Z(B')) = Cov(\ Z(A"),\ Z(B"))$. Estos tres resultados dan origen a una condición más debil que la estacionaridad estricta que se presenta a continuación.

1.2.1. Procesos estacionarios de segundo orden

Si el conjunto finito se reduce a un punto, $m = 1$ todas las variables z_j están igualmente distribuidas entonces:

- $E(z_j) = \mu \quad constante \ \vee z_j$
- $Var(z_j) \quad constante \ \vee z_j$

Si el conjunto finito se reduce a dos puntos $\{x_1,\ x_2\}$, $m = 2$, las covarianzas son iguales ante traslaciones:

$$cov(z(x_1),\ z(x_2)) = cov(z(x_1 + h),\ z(x_2 + h)) \quad \vee h \tag{1.2}$$

El proceso es *estacionario de segundo orden* si se cumplen las 3 propiedades arriba indicadas.

Si bien es una hipótesis más debil que la estacionaridad es suficiente para múltiples propósitos en geoestadística, por ejemplo predicción por Kriging. Además, si el proceso es gaussiano (las distribuciones conjuntas Normalmente distribuidas), estacionario de segundo orden y estacionario son equivalentes ya que la Normal multivariada queda caracterizada por los momentos de primer y segundo orden.

1.2.2. Existencia de una función de covarianza $Cov(h)$

La estacionaridad de segundo orden implica la existencia de una función de covarianza $Cov(h)$ tal que:

$$cov(z(x_1),\ z(x_2)) = Cov(h) \quad h = x_2 - x_1 \tag{1.3}$$

siendo h el vector del espacio diferencia entre los puntos dados.

Demostración

▼ Fijado un punto arbitrario del dominio, x_{fijo}, entonces, para $k = x_{fijo} - x_1$ sustituyendo en ecu. 1.2$cov(z(x_1),\ z(x_2)) = cov(z(x_1 + k),\ z(x_2 + k))$

luego, $cov(z(x_{fijo}),\ z(x_2 - x_1 + x_{fijo})) = cov(z(x_{fijo}),\ z(x_{fiijo} + h))$

Esta expresión sólo depende de h; luego, la covarianza entre dos variables $z(x_1)$, $z(x_2)$ depende del vector $h = x_2 - x_1$ y no de las posiciones en sí de x_1, x_2 ▲

1.2.3. Positividad

Si el espacio es el habitual R^3, entonces la función Cov es de $R^3 \Rightarrow R$. $Cov(h)$ no es cualquier función ya que debe tener la característica de generar matrices definidas positivas. Cuando Cov es de $R \Rightarrow R$ (caso isotrópico, ver punto siguiente) son modelos válidos: exponencial, esférico, gaussiano, Matern, etc.

1.2.4. Isotropía

El proceso es *isotrópico* si la dependencia es más simple aún y sólo depende de la distancia $\|h\|$ y no de su dirección. Es decir $cov(z(x_1), z(x_2)) = Cov(\|h\|)$ y es una gran simplificación ya que la función Cov en lugar de ser de $R^3 \Rightarrow R$, es de $R \Rightarrow R$, y es una función más simple y por ende más facil de modelar. La isotropía implica que las iso-superficies son esferas en R^3 o circunferencias en R^2. Si Cov es dependiente de la dirección del vector además de su módulo se afirma que el campo aleatorio es *anisotrópico* y existen distintos tipos de anisotropía.

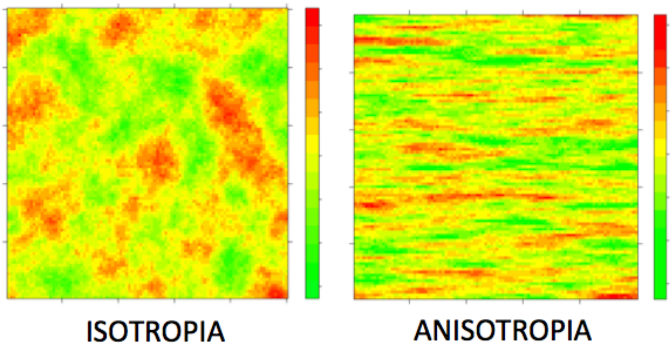

Figura 1.4: Realizaciones de un proceso isotrópico (izq) y anisotrópico (der)

La figura 1.4 muestra a la izquierda una realización de un proceso isotrópico, y a la derecha un anisotrópico donde se observan extensas zonas horizontales casi constantes en contraste con los rápidos cambios en vertical.

1.3. Análisis y modelado de la continuidad espacial

Este tema, fundamental en geoestádistica, que permite inferir un modelo de la estructura de covarianza a partir de una muestra, no se incluye en estas notas orientadas a predicción por Kriging.