INTRODUCCIÓN AL ANÁLISIS GEOESTADÍSTICO DE DATOS EN GEOCIENCIAS: TEORÍA Y APLICACIÓN

Maximiliano Garnier

Escuela Centroamericana de Geología, Universidad de Costa Rica.

maximiliano.garniervillarreal@ucr.ac.cr

**RESUMEN:** Kriging se ha usado y ha sido propuesto como el mejor método de interpolación, muchas veces sin realmente entender cómo es que se usa adecuadamente y dejando que el software que lo brinda decida cómo implementarlo. Esta aseveración tiene fundamento cuando se procede de la manera correcta, realizando los pasos necesarios durante el análisis y modelado geoestadístico, por lo que es necesario entender cómo aplicar Kriging correctamente para que los resultados obtenidos sean relevantes y confiables. Estos pasos se detallan en este trabajo, abordando la teoría y mediante un ejemplo se pone en práctica el método usando el software estadístico libre **R**. Adicionalmente, se presenta una aplicación web de libre acceso para quienes no se sientan cómodos usando lenguajes de programación.

Palabras clave: Geoestadística, Kriging , R, Variograma, Interpolación

# INTRODUCCIÓN

En las ciencias que tienen una fuerte componente espacial (dentro de ellas geología) es común recolectar muestras, describirlas y tener la ubicación de dónde se recolectaron. En muchos casos el muestreo se hace con el fin de caracterizar una variable o proceso en el espacio, con lo que se tiene en mente pasar de puntos a una superficie (mapa).

Para poder generar estas superficies se pueden emplear diferentes [Métodos de interpolación](#geostats-met-interp), donde comúnmente se ha dado a entender que Kriging es el método por excelencia a usar (casi que indiscriminadamente), pero no se ha profundizado en cómo usar el método apropiadamente y cuándo es adecuado o no utilizarlo.

# MÉTODOS DE INTERPOLACIÓN

De manera resumida y sin entrar en mucho detalle se mencionan diferentes métodos de interpolación comúnmente usados, para ellos se puede consultar Webster & Oliver ([2007](#ref-webster2007)). De manera general se tienen:

# GEOESTADÍSTICA

## Conceptos básicos

### Correlación espacial

### Semivarianza

Es la medida que se usa para determinar la disimilitud entre observaciones que varían con la distancia, y se representa mediante la Ecuación (1), donde es el valor de la variable en la posición , es el valor de la variable a una distancia , es el número total de puntos (observaciones), y es el número de pares de puntos que se encuentran a una distancia específica. **Se recomienda tener más de 30 pares de puntos por cada distancia , y no calcular la semivarianza más allá de la mitad de la máxima distancia entre observaciones** (Chilès & Delfiner, [1999](#ref-chiles1999); Goovaerts, [1997](#ref-goovaerts1997); Isaaks & Srivastava, [1989](#ref-isaaks1989); Webster & Oliver, [2007](#ref-webster2007)).

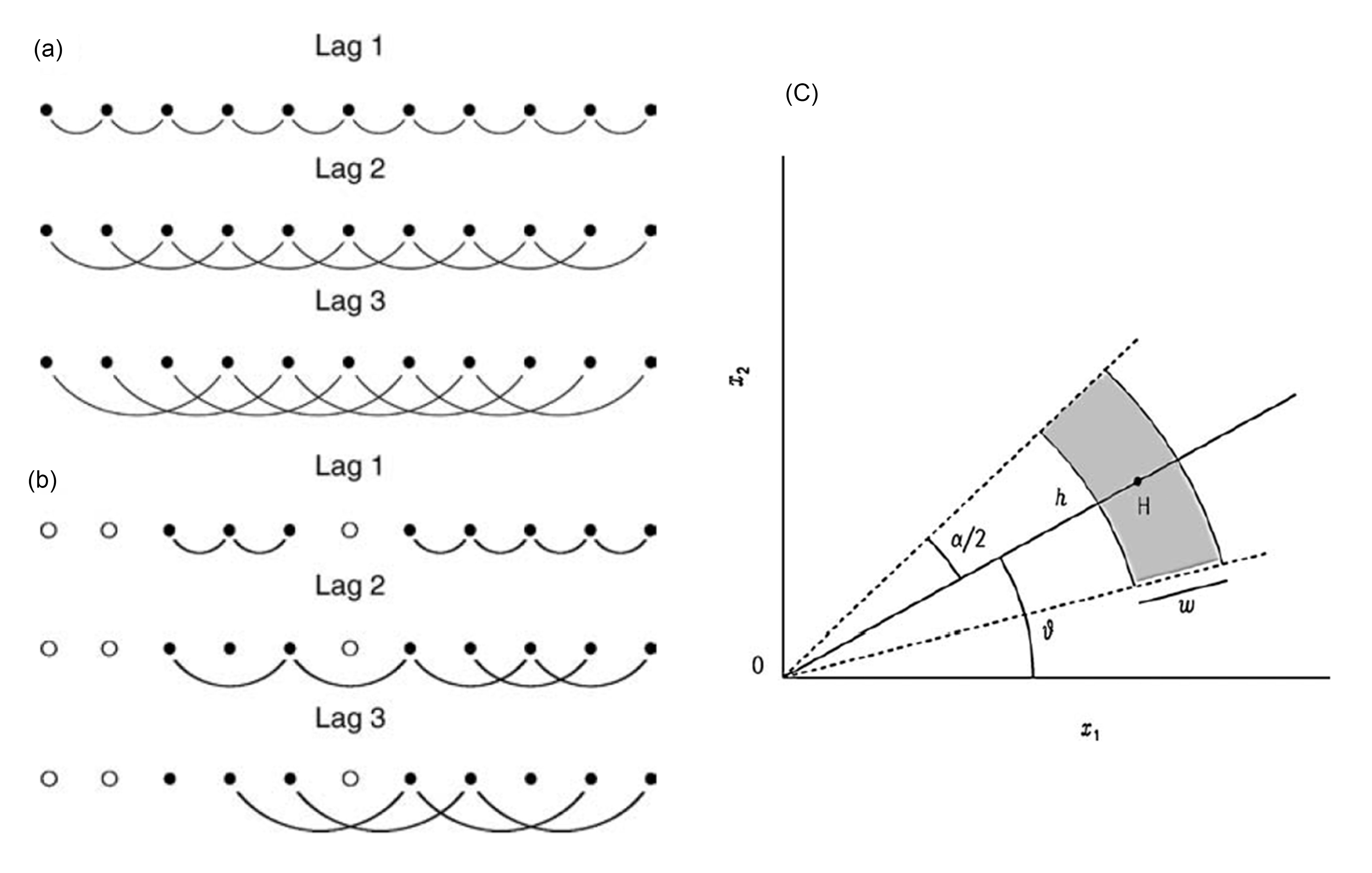


Figura 1: Esquema del calculo de la semivarianza para datos regularmente espaciados, donde los datos están completos (a) y donde hay datos faltantes (b); para datos irregularmente espaciados (c). Modificado de Webster & Oliver ([*2007*](#ref-webster2007)).

### Variograma experimental

### Modelo de variograma

#### Tipos

### Anisotropía

### Validación cruzada

### Kriging

# Análisis geoestadístico

## Análisis Exploratorio de Datos

Tabla 1: Resumen estadístico de la variable.

|  |  |
| --- | --- |
| Estadistica | z |
| Mean | 29.90 |
| Std.Dev | 0.86 |
| Min | 28.19 |
| Q1 | 29.20 |
| Median | 29.86 |
| Q3 | 30.38 |
| Max | 31.95 |
| MAD | 0.94 |
| IQR | 1.15 |
| CV | 0.03 |
| Skewness | 0.47 |
| N.Valid | 60.00 |

## Modelado geoestadístico

### Variograma experimental

myformula = as.formula(paste(myvar,'~1'))  
g = gstat(formula = myformula,   
 data = datos\_sf) # objeto gstat para hacer geoestadistica  
  
# variograma experimental cada cierta distancia (width), y hasta cierta distancia (cutoff)  
dat.vgm = variogram(g,   
 width = 5,  
 cutoff = 60)

### Ajuste de modelo de variograma

### Validación cruzada

## Interpolación (Kriging)

# Aplicación web

# Conclusiones

# Referencias

Borradaile, G. J. (2003). *Statistics of Earth Science Data: Their Distribution in Time, Space and Orientation*. Springer-Verlag Berlin Heidelberg.