# EPG3343 - Seminario de Estadística III

## Clase 5

Jonathan Acosta

Pontificia Universidad Católica de Chile

Segundo Semestre, 2022

# Esquema

1. Introducción a la Predicción Espacial (Kriging)
   - Nociones de Esperanza Condicional
   - Nociones de la Distribución Normal Multivariada

2. Kriging Lineales
   - Kriging Simple
   - Kriging Ordinario
   - Kriging Universal

# Esquema

1. Introducción a la Predicción Espacial (Kriging)
   - Nociones de Esperanza Condicional
   - Nociones de la Distribución Normal Multivariada
2. Kriging Lineales
   - Kriging Simple
   - Kriging Ordinario
   - Kriging Universal

# Introducción a la Predicción Espacial (Kriging)

**Objetivo General**: Predecir un proceso $Z$ en un punto, denotado por $\boldsymbol{s}_0$ donde el proceso no ha sido observado, a partir de un conjunto de $n$ observaciones del proceso.

# Introducción a la Predicción Espacial (Kriging)

**Objetivo General**: Predecir un proceso $Z$ en un punto, denotado por $\boldsymbol{s}_0$ donde el proceso no ha sido observado, a partir de un conjunto de $n$ observaciones del proceso.

- Esta idea fue planteada originalmente por Krige (1951).

# Introducción a la Predicción Espacial (Kriging)

**Objetivo General**: Predecir un proceso $Z$ en un punto, denotado por $\boldsymbol{s}_0$ donde el proceso no ha sido observado, a partir de un conjunto de $n$ observaciones del proceso.

- Esta idea fue planteada originalmente por Krige (1951).
- Los trabajos empíricos de Krige para evaluar recursos mineros, fueron formalizados en la década del 60 por el ingeniero francés Georges Matheron.

# Introducción a la Predicción Espacial (Kriging)

**Objetivo General**: Predecir un proceso $Z$ en un punto, denotado por $\boldsymbol{s}_0$ donde el proceso no ha sido observado, a partir de un conjunto de $n$ observaciones del proceso.

- Esta idea fue planteada originalmente por Krige (1951).
- Los trabajos empíricos de Krige para evaluar recursos mineros, fueron formalizados en la década del 60 por el ingeniero francés Georges Matheron.
- Georges Matheron desarrolló la técnica denominada **kriging** en honor a la investigación de Krige en el campo de la geoestadística.

# Introducción a la Predicción Espacial (Kriging)

**Notacion**:

- $\boldsymbol{Z} = (Z(\boldsymbol{s}_1), Z(\boldsymbol{s}_2), \ldots, Z(\boldsymbol{s}_n))^\top$ datos observados.
- $Z(\boldsymbol{s}_0)$ desconocido pero fijo.
- $P(\boldsymbol{Z}, g(Z(\boldsymbol{s}_0)))$: Predictor de $g(Z(\boldsymbol{s}_0))$ en el punto $\boldsymbol{s}_0$.
- Caso de interés es $g(Z(\boldsymbol{s}_0)) = Z(\boldsymbol{s}_0)$
- $P(\boldsymbol{Z}, Z(\boldsymbol{s}_0)) = P(Z, \boldsymbol{s}_0)$.
- $L = L(\boldsymbol{s}_0, \boldsymbol{Z}) = \big(Z(\boldsymbol{s}_0) - P(Z, \boldsymbol{s}_0)\big)^2$: Pérdida Cuadrática.
- $E[L] = \mathbb{E}\left[\big(Z(\boldsymbol{s}_0) - P(Z, \boldsymbol{s}_0)\big)^2\right]$: función de Riesgo.

**Objetivo Específico**: Obtener $P^0(Z, \boldsymbol{s}_0)$ como aquel que minimiza la función de riesgo, es decir,

$$P^0(Z, \boldsymbol{s}_0) = \text{argmin} \left\{ \mathbb{E} \left[ \left( Z(\boldsymbol{s}_0) - P(Z, \boldsymbol{s}_0) \right)^2 \right] \right\}$$

# Introducción a la Predicción Espacial

Las herramientas que se utilizan para resolver el problema de minimización anterior son:

# Introducción a la Predicción Espacial

Las herramientas que se utilizan para resolver el problema de minimización anterior son:

1. Esperanzas Condicionales

Las herramientas que se utilizan para resolver el problema de minimización anterior son:

1. Esperanzas Condicionales
2. Distribución Normal Multivariada

# Introducción a la Predicción Espacial

Las herramientas que se utilizan para resolver el problema de minimización anterior son:

1. Esperanzas Condicionales
2. Distribución Normal Multivariada
3. Regresión Lineal Multiple

# Introducción a la Predicción Espacial

Las herramientas que se utilizan para resolver el problema de minimización anterior son:

1. Esperanzas Condicionales
2. Distribución Normal Multivariada
3. Regresión Lineal Multiple
4. Multiplicadores de Lagrange

# Esperanza Condicional

**Definición**: Sean $X$ e $Y$ dos variables aleatorias con función de densidad conjunta $f_{X,Y}(x,y)$ y densidades marginales $f_X(x)$ y $f_Y(y)$. Entonces la densidad condicional de $Y|X=x$ es

$$f_{Y|X=x}(y) = \frac{f_{X,Y}(x,y)}{f_X(x)}$$

Además su valor esperado es

$$\mathbb{E}[Y|X=x] = \int y f_{Y|X=x}(y) dy$$

# Esperanza Condicional

**Propiedades**:

1. $Y|X$ es una variable aleatoria.

# Esperanza Condicional

**Propiedades**:

1. $Y|X$ es una variable aleatoria.
2. $\mathbb{E}[Y|X = x]$ es un número.

# Esperanza Condicional

**Propiedades**:

1. $Y|X$ es una variable aleatoria.
2. $\mathbb{E}[Y|X=x]$ es un número.
3. $\mathbb{E}[Y|X]$ es una variable aleatoria, función de $X$

# Esperanza Condicional

**Propiedades**:

1. $Y|X$ es una variable aleatoria.
2. $\mathbb{E}[Y|X = x]$ es un número.
3. $\mathbb{E}[Y|X]$ es una variable aleatoria, función de $X$
4. $\mathbb{E}[\mathbb{E}[Y|X]] = \mathbb{E}[Y]$

# Esperanza Condicional

**Resultado**: Sea $r$ una función tal que $r(X)$ está bien definida. Entonces se tienen los siguientes resultados:

# Esperanza Condicional

**Resultado**: Sea $r$ una función tal que $r(X)$ está bien definida. Entonces se tienen los siguientes resultados:

1. $\mathbb{E}[r(X)Y|X] = r(X)\mathbb{E}[Y|X]$

**Resultado**: Sea $r$ una función tal que $r(X)$ está bien definida. Entonces se tienen los siguientes resultados:

1. $\mathbb{E}[r(X)Y|X] = r(X)\mathbb{E}[Y|X]$
2. $\mathbb{E}[r(X)\mathbb{E}[Y|X]] = \mathbb{E}[r(X)Y]$

# Esperanza Condicional

**Resultado**: Sea $r$ una función tal que $r(X)$ está bien definida. Entonces se tienen los siguientes resultados:

1. $\mathbb{E}[r(X)Y|X] = r(X)\mathbb{E}[Y|X]$
2. $\mathbb{E}[r(X)\mathbb{E}[Y|X]] = \mathbb{E}[r(X)Y]$
3. $Y - \mathbb{E}[Y|X]$ y $r(X)$ son no correlacionados.

# Esperanza Condicional

**Resultado**: Sea $r$ una función tal que $r(X)$ está bien definida. Entonces se tienen los siguientes resultados:

1. $\mathbb{E}[r(X)Y|X] = r(X)\mathbb{E}[Y|X]$
2. $\mathbb{E}[r(X)\mathbb{E}[Y|X]] = \mathbb{E}[r(X)Y]$
3. $Y - \mathbb{E}[Y|X]$ y $r(X)$ son no correlacionados.

**Demostración**:(1) por definición. (2) directo desde la Propiedad (4). Para (3) calcular la covarianza entre $Y - \mathbb{E}[Y|X]$ y $r(X)$, luego concluir.

# Esperanza Condicional

**Teorema**: (Mejor Predictor) Sea $Y$ una variable aleatoria, sea $\boldsymbol{U}$ un vector aleatorio y sea $g$ una función. Entonces una de las siguientes condiciones se cumple:

(a) $\mathbb{E}\left[\left(Y - g(\boldsymbol{U})\right)^2\right] = \infty$

(b) $\mathbb{E}\left[\left(Y - \mathbb{E}[Y|\boldsymbol{U}]\right)^2\right] \leq \mathbb{E}\left[\left(Y - g(\boldsymbol{U})\right)^2\right]$

**Teorema**: (Mejor Predictor) Sea $Y$ una variable aleatoria, sea $\boldsymbol{U}$ un vector aleatorio y sea $g$ una función. Entonces una de las siguientes condiciones se cumple:

(a) $\mathbb{E}\left[\big(Y - g(\boldsymbol{U})\big)^2\right] = \infty$

(b) $\mathbb{E}\left[\big(Y - \mathbb{E}[Y|\boldsymbol{U}]\big)^2\right] \leq \mathbb{E}\left[\big(Y - g(\boldsymbol{U})\big)^2\right]$

**Demostración**: Sumar y restar $\mathbb{E}[Y|\boldsymbol{U}]$ en $\mathbb{E}\left[\big(Y - g(\boldsymbol{U})\big)^2\right]$ y luego concluir usando el resultado anterior.

# Esperanza Condicional

**Consecuencia**: $\mathbb{E}[Y|\boldsymbol{U}]$ minimiza la función de pérdida cuadrática.

**Consecuencia**: $\mathbb{E}[Y|\boldsymbol{U}]$ minimiza la función de pérdida cuadrática.

Por lo tanto, se tiene que

$$P^0(Z, \boldsymbol{s}_0) = \mathbb{E}[Z(\boldsymbol{s}_0)|\boldsymbol{Z}]$$

**Consecuencia**: $\mathbb{E}[Y|\boldsymbol{U}]$ minimiza la función de pérdida cuadrática.

Por lo tanto, se tiene que

$$P^0(Z, \boldsymbol{s}_0) = \mathbb{E}[Z(\boldsymbol{s}_0)|\boldsymbol{Z}]$$

Más aún,

$$\mathbb{E}\left[\left(Z(\boldsymbol{s}_0) - P^0(Z, \boldsymbol{s}_0)\right)^2\right] = \mathbb{V}[Z(\boldsymbol{s}_0)] - \mathbb{V}\left[P^0(Z, \boldsymbol{s}_0)\right]$$

En el caso Gaussiano, la esperanza condicional tiene forma explicita.

# Distribución Normal Multivariada

En el caso Gaussiano, la esperanza condicional tiene forma explicita.

Suponga que $\boldsymbol{W} = (\boldsymbol{V}, \boldsymbol{U})^\top \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ tal que

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_V \\ \boldsymbol{\mu}_U \end{pmatrix}; \qquad \boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_V & \boldsymbol{\Sigma}_{VU} \\ \boldsymbol{\Sigma}_{UV} & \boldsymbol{\Sigma}_U \end{bmatrix}$$

Entonces, $\boldsymbol{V}|\boldsymbol{U} \sim \mathcal{N}\,(\mathbb{E}[\boldsymbol{V}|\boldsymbol{U}], \mathbb{V}[\boldsymbol{V}|\boldsymbol{U}])$, donde

$$\begin{aligned} \mathbb{E}[\boldsymbol{V}|\boldsymbol{U}] &= \boldsymbol{\mu}_V + \boldsymbol{\Sigma}_{VU}\boldsymbol{\Sigma}_U^{-1}(\boldsymbol{U} - \boldsymbol{\mu}_U) \\ \mathbb{V}[\boldsymbol{V}|\boldsymbol{U}] &= \boldsymbol{\Sigma}_V - \boldsymbol{\Sigma}_{VU}\boldsymbol{\Sigma}_U^{-1}\boldsymbol{\Sigma}_{UV} \end{aligned}$$

# Distribución Normal Multivariada

En términos de la notación anterior,

$$V = Z(\boldsymbol{s}_0); \qquad \boldsymbol{U} = \boldsymbol{Z}$$

$$\mathbb{E}[Z(\boldsymbol{s}_0)] = \mu(\boldsymbol{s}_0); \qquad \mathbb{E}[\boldsymbol{Z}] = \boldsymbol{\mu}(\boldsymbol{s})$$

$$\mathbb{V}[\boldsymbol{Z}] = \boldsymbol{\Sigma}; \qquad \operatorname{cov}(\boldsymbol{Z}, Z(\boldsymbol{s}_0)) = \boldsymbol{C}$$

# Distribución Normal Multivariada

En términos de la notación anterior,

$$V = Z(\boldsymbol{s}_0); \qquad \boldsymbol{U} = \boldsymbol{Z}$$

$$\mathbb{E}[Z(\boldsymbol{s}_0)] = \mu(\boldsymbol{s}_0); \qquad \mathbb{E}[\boldsymbol{Z}] = \boldsymbol{\mu}(\boldsymbol{s})$$

$$\mathbb{V}[\boldsymbol{Z}] = \boldsymbol{\Sigma}; \qquad \operatorname{cov}(\boldsymbol{Z}, Z(\boldsymbol{s}_0)) = \boldsymbol{C}$$

Entonces

$$P^0(Z, \boldsymbol{s}_0) = \mu(\boldsymbol{s}_0) + \boldsymbol{C}^\top \boldsymbol{\Sigma}^{-1} \left(\boldsymbol{Z} - \boldsymbol{\mu}(\boldsymbol{s})\right)$$

# Distribución Normal Multivariada

En términos de la notación anterior,

$$V = Z(\boldsymbol{s}_0); \qquad \boldsymbol{U} = \boldsymbol{Z}$$

$$\mathbb{E}[Z(\boldsymbol{s}_0)] = \mu(\boldsymbol{s}_0); \qquad \mathbb{E}[\boldsymbol{Z}] = \boldsymbol{\mu}(\boldsymbol{s})$$

$$\mathbb{V}[\boldsymbol{Z}] = \boldsymbol{\Sigma}; \qquad \mathrm{cov}(\boldsymbol{Z}, Z(\boldsymbol{s}_0)) = \boldsymbol{C}$$

Entonces

$$P^0(Z, \boldsymbol{s}_0) = \mu(\boldsymbol{s}_0) + \boldsymbol{C}^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{Z} - \boldsymbol{\mu}(\boldsymbol{s}))$$

$$\Rightarrow \quad P^0(Z, \boldsymbol{s}_0) = \lambda_0 + \sum_{i=1}^{n} \lambda_i Z(\boldsymbol{s}_i)$$

# Distribución Normal Multivariada

En términos de la notación anterior,

$$V = Z(\boldsymbol{s}_0); \qquad \boldsymbol{U} = \boldsymbol{Z}$$

$$\mathbb{E}[Z(\boldsymbol{s}_0)] = \mu(\boldsymbol{s}_0); \qquad \mathbb{E}[\boldsymbol{Z}] = \boldsymbol{\mu}(\boldsymbol{s})$$

$$\mathbb{V}[\boldsymbol{Z}] = \boldsymbol{\Sigma}; \qquad \operatorname{cov}(\boldsymbol{Z}, Z(\boldsymbol{s}_0)) = \boldsymbol{C}$$

Entonces

$$P^0(Z, \boldsymbol{s}_0) = \mu(\boldsymbol{s}_0) + \boldsymbol{C}^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{Z} - \boldsymbol{\mu}(\boldsymbol{s}))$$

$$\Rightarrow \quad P^0(Z, \boldsymbol{s}_0) = \lambda_0 + \sum_{i=1}^{n} \lambda_i Z(\boldsymbol{s}_i)$$

Además,

$$\mathbb{E}\left[\left(Z(\boldsymbol{s}_0) - P^0(Z, \boldsymbol{s}_0)\right)^2\right] = \mathbb{V}[Z(\boldsymbol{s}_0)] - \boldsymbol{C}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{C}$$

# Esquema

# Kriging Simple

Sea $\boldsymbol{Z}(\boldsymbol{s}) = (Z(\boldsymbol{s}_1), \ldots, Z(\boldsymbol{s}_n))^\top$ tal que

$$\boldsymbol{Z}(\boldsymbol{s}) = \boldsymbol{\mu}(\boldsymbol{s}) + \boldsymbol{e}(\boldsymbol{s}), \qquad \text{donde} \quad \boldsymbol{e}(\boldsymbol{s}) \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$$

# Kriging Simple

Sea $\boldsymbol{Z}(\boldsymbol{s}) = (Z(\boldsymbol{s}_1), \ldots, Z(\boldsymbol{s}_n))^\top$ tal que

$$\boldsymbol{Z}(\boldsymbol{s}) = \boldsymbol{\mu}(\boldsymbol{s}) + \boldsymbol{e}(\boldsymbol{s}), \qquad \text{donde} \quad \boldsymbol{e}(\boldsymbol{s}) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$$

Los supuestos sobre el Kriging Simple:

- $\boldsymbol{\mu}(\boldsymbol{s})$ es conocido
- $\boldsymbol{\Sigma}$ es conocido

# Kriging Simple

Sea $\boldsymbol{Z}(\boldsymbol{s}) = (Z(\boldsymbol{s}_1), \ldots, Z(\boldsymbol{s}_n))^\top$ tal que

$$\boldsymbol{Z}(\boldsymbol{s}) = \boldsymbol{\mu}(\boldsymbol{s}) + \boldsymbol{e}(\boldsymbol{s}), \qquad \text{donde} \quad \boldsymbol{e}(\boldsymbol{s}) \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$$

Los supuestos sobre el Kriging Simple:

- $\boldsymbol{\mu}(\boldsymbol{s})$ es conocido
- $\boldsymbol{\Sigma}$ es conocido

Asumir que

$$P(Z, \boldsymbol{s}_0) = \lambda_0 + \boldsymbol{\lambda}^\top \boldsymbol{Z}, \qquad \boldsymbol{\lambda}^\top = (\lambda_1, \ldots, \lambda_n)\,, \quad \lambda_i \in \mathbb{R}.$$

Notar que $\lambda_0$ y $\boldsymbol{\lambda}$ son parámetros desconocidos que se obtienen minimizando

$$\mathbb{E}\left[\big(Z(\boldsymbol{s}_0) - P(Z, \boldsymbol{s}_0)\big)^2\right]$$

# Kriging Simple

La condición de insesgamiento:

$$\mathbb{E}[P(Z, \boldsymbol{s}_0)] = \mathbb{E}[Z(\boldsymbol{s}_0)] \quad \Longleftrightarrow \quad \lambda_0 + \boldsymbol{\lambda}^\top \boldsymbol{\mu}(\boldsymbol{s}) = \mu(\boldsymbol{s}_0)$$

Se satisface si $\lambda_0 = \mu(\boldsymbol{s}_0) - \boldsymbol{\lambda}^\top \boldsymbol{\mu}(\boldsymbol{s})$.

# Kriging Simple

La condición de insesgamiento:

$$\mathbb{E}[P(Z, \boldsymbol{s}_0)] = \mathbb{E}[Z(\boldsymbol{s}_0)] \quad \iff \quad \lambda_0 + \boldsymbol{\lambda}^\top \boldsymbol{\mu}(\boldsymbol{s}) = \mu(\boldsymbol{s}_0)$$

Se satisface si $\lambda_0 = \mu(\boldsymbol{s}_0) - \boldsymbol{\lambda}^\top \boldsymbol{\mu}(\boldsymbol{s})$.

Luego,

$$\mathbb{E}\left[\left(P(Z, \boldsymbol{s}_0) - Z(\boldsymbol{s}_0)\right)^2\right] = \mathbb{V}\left[\boldsymbol{\lambda}^\top \boldsymbol{Z} - Z(\boldsymbol{s}_0)\right] = \underbrace{\boldsymbol{\lambda}^\top \boldsymbol{\Sigma} \boldsymbol{\lambda} + \sigma^2 - 2\boldsymbol{\lambda}^\top \boldsymbol{C}}_{g(\boldsymbol{\lambda})}$$

donde $\sigma^2 = \mathbb{V}[Z(\boldsymbol{s}_0)]$, $\boldsymbol{\Sigma} = \mathbb{V}[\boldsymbol{Z}]$ y $\boldsymbol{C} = \operatorname{cov}(Z(\boldsymbol{s}_0), \boldsymbol{Z})$.

# Kriging Simple

$$\nabla g(\boldsymbol{\lambda}) = \mathbf{0} \quad \Leftrightarrow \quad 2\boldsymbol{\Sigma}\boldsymbol{\lambda} - 2\boldsymbol{C} = 0 \quad \Rightarrow \quad \boldsymbol{\lambda} = \boldsymbol{\Sigma}^{-1}\boldsymbol{C}$$

# Kriging Simple

$$\nabla g(\boldsymbol{\lambda}) = \mathbf{0} \quad \Leftrightarrow \quad 2\boldsymbol{\Sigma}\boldsymbol{\lambda} - 2\boldsymbol{C} = 0 \quad \Rightarrow \quad \boldsymbol{\lambda} = \boldsymbol{\Sigma}^{-1}\boldsymbol{C}$$

Por lo tanto, el predictor por Kriging Simple de $Z(\boldsymbol{s}_0)$ es

$$P_{SK}(Z, \boldsymbol{s}_0) = \mu(\boldsymbol{s}_0) + \boldsymbol{C}^\top \boldsymbol{\Sigma}^{-1} \left(\boldsymbol{Z} - \boldsymbol{\mu}(\boldsymbol{s})\right)$$

$$\nabla g(\boldsymbol{\lambda}) = \mathbf{0} \quad \Leftrightarrow \quad 2\boldsymbol{\Sigma}\boldsymbol{\lambda} - 2\boldsymbol{C} = 0 \quad \Rightarrow \quad \boldsymbol{\lambda} = \boldsymbol{\Sigma}^{-1}\boldsymbol{C}$$

Por lo tanto, el predictor por Kriging Simple de $Z(\boldsymbol{s}_0)$ es

$$P_{SK}(Z, \boldsymbol{s}_0) = \mu(\boldsymbol{s}_0) + \boldsymbol{C}^\top \boldsymbol{\Sigma}^{-1} \left(\boldsymbol{Z} - \boldsymbol{\mu}(\boldsymbol{s})\right)$$

Además, la varianza de Predicción esta dada por

$$\sigma^2_{SK}(\boldsymbol{s}_0) = \sigma^2 - \boldsymbol{C}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{C}$$

# Kriging Simple

**Comentario**: En la práctica asumir que la media del proceso es conocida es un supuesto demasiado grande. Pero asumiendo que la media del proceso es $\boldsymbol{\mu}(\boldsymbol{s}) = \boldsymbol{X}(\boldsymbol{s})\boldsymbol{\beta}$ ha sido bien especificada, se puede plantear el siguiente método:

# Kriging Simple

**Comentario**: En la práctica asumir que la media del proceso es conocida es un supuesto demasiado grande. Pero asumiendo que la media del proceso es $\boldsymbol{\mu}(\boldsymbol{s}) = \boldsymbol{X}(\boldsymbol{s})\boldsymbol{\beta}$ ha sido bien especificada, se puede plantear el siguiente método:

1. Ajustar el modelo: $\widehat{\boldsymbol{\beta}}_{LS} = \left(\boldsymbol{X}(\boldsymbol{s})^\top \boldsymbol{X}(\boldsymbol{s})\right)^{-1} \boldsymbol{X}(\boldsymbol{s})^\top \boldsymbol{Z}$

# Kriging Simple

**Comentario**: En la práctica asumir que la media del proceso es conocida es un supuesto demasiado grande. Pero asumiendo que la media del proceso es $\boldsymbol{\mu}(\boldsymbol{s}) = \boldsymbol{X}(\boldsymbol{s})\boldsymbol{\beta}$ ha sido bien especificada, se puede plantear el siguiente método:

1. Ajustar el modelo: $\widehat{\boldsymbol{\beta}}_{LS} = \left(\boldsymbol{X}(\boldsymbol{s})^\top \boldsymbol{X}(\boldsymbol{s})\right)^{-1} \boldsymbol{X}(\boldsymbol{s})^\top \boldsymbol{Z}$
2. Determinar los Residuos: $\widehat{\boldsymbol{e}} = \boldsymbol{Z} - \boldsymbol{X}(\boldsymbol{s})\widehat{\boldsymbol{\beta}}$

# Kriging Simple

**Comentario**: En la práctica asumir que la media del proceso es conocida es un supuesto demasiado grande. Pero asumiendo que la media del proceso es $\boldsymbol{\mu}(\boldsymbol{s}) = \boldsymbol{X}(\boldsymbol{s})\boldsymbol{\beta}$ ha sido bien especificada, se puede plantear el siguiente método:

1. Ajustar el modelo: $\widehat{\boldsymbol{\beta}}_{LS} = \left(\boldsymbol{X}(\boldsymbol{s})^\top \boldsymbol{X}(\boldsymbol{s})\right)^{-1} \boldsymbol{X}(\boldsymbol{s})^\top \boldsymbol{Z}$
2. Determinar los Residuos: $\widehat{\boldsymbol{e}} = \boldsymbol{Z} - \boldsymbol{X}(\boldsymbol{s})\widehat{\boldsymbol{\beta}}$
3. Realizar kriging simple sobre los residuos, obtener $P_{SK}(\widehat{e}, \boldsymbol{s}_0)$

# Kriging Simple

**Comentario**: En la práctica asumir que la media del proceso es conocida es un supuesto demasiado grande. Pero asumiendo que la media del proceso es $\boldsymbol{\mu}(\boldsymbol{s}) = \boldsymbol{X}(\boldsymbol{s})\boldsymbol{\beta}$ ha sido bien especificada, se puede plantear el siguiente método:

1. Ajustar el modelo: $\widehat{\boldsymbol{\beta}}_{LS} = \left(\boldsymbol{X}(\boldsymbol{s})^\top \boldsymbol{X}(\boldsymbol{s})\right)^{-1} \boldsymbol{X}(\boldsymbol{s})^\top \boldsymbol{Z}$
2. Determinar los Residuos: $\widehat{\boldsymbol{e}} = \boldsymbol{Z} - \boldsymbol{X}(\boldsymbol{s})\widehat{\boldsymbol{\beta}}$
3. Realizar kriging simple sobre los residuos, obtener $P_{SK}(\widehat{e}, \boldsymbol{s}_0)$
4. Obtener el predictor de $Z(\boldsymbol{s}_0)$, $P(Z, \boldsymbol{s}_0) = \boldsymbol{X}(\boldsymbol{s})\widehat{\boldsymbol{\beta}} + P_{SK}(\widehat{e}, \boldsymbol{s}_0)$

# Kriging Simple

**Criticas**:

- $\widehat{\boldsymbol{\beta}}_{LS}$ no considera $\boldsymbol{\Sigma}$

# Kriging Simple

**Criticas**:

- $\widehat{\boldsymbol{\beta}}_{LS}$ no considera $\boldsymbol{\Sigma}$
- Aún cuando, se utilice

$$\widehat{\boldsymbol{\beta}}_{GLS} = \left(\boldsymbol{X}(\boldsymbol{s})^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{X}(\boldsymbol{s})\right)^{-1} \boldsymbol{X}(\boldsymbol{s})^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{Z}$$

$\boldsymbol{\Sigma}$ en general es desconocido.

# Kriging Simple

**Criticas**:

- $\widehat{\boldsymbol{\beta}}_{LS}$ no considera $\boldsymbol{\Sigma}$
- Aún cuando, se utilice

$$\widehat{\boldsymbol{\beta}}_{GLS} = \left(\boldsymbol{X}(\boldsymbol{s})^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{X}(\boldsymbol{s})\right)^{-1} \boldsymbol{X}(\boldsymbol{s})^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{Z}$$

  $\boldsymbol{\Sigma}$ en general es desconocido.

- Obtener $\widehat{\boldsymbol{\Sigma}}_{\text{variograma}} \longrightarrow \widehat{\boldsymbol{\beta}}$ sesgado.

# Kriging Ordinario

Considere el modelo

$$\boldsymbol{Z}(\boldsymbol{s}) = \boldsymbol{\mu}(\boldsymbol{s}) + \boldsymbol{e}(\boldsymbol{s}), \qquad \text{donde} \quad \boldsymbol{e}(\boldsymbol{s}) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$$

# Kriging Ordinario

Considere el modelo

$$\boldsymbol{Z}(\boldsymbol{s}) = \boldsymbol{\mu}(\boldsymbol{s}) + \boldsymbol{e}(\boldsymbol{s}), \qquad \text{donde} \quad \boldsymbol{e}(\boldsymbol{s}) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$$

Los supuestos sobre el Kriging Ordinario:

- $\mu(\boldsymbol{s}) = \mu, \;\; \forall \boldsymbol{s}$, con $\mu \in \mathbb{R}$ desconocido
- $\boldsymbol{\Sigma}$ es conocido

# Kriging Ordinario

Considere el modelo

$$\boldsymbol{Z}(\boldsymbol{s}) = \boldsymbol{\mu}(\boldsymbol{s}) + \boldsymbol{e}(\boldsymbol{s}), \qquad \text{donde} \quad \boldsymbol{e}(\boldsymbol{s}) \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$$

Los supuestos sobre el Kriging Ordinario:

- $\mu(\boldsymbol{s}) = \mu, \;\; \forall \boldsymbol{s}$, con $\mu \in \mathbb{R}$ desconocido
- $\boldsymbol{\Sigma}$ es conocido

Es decir,

$$\boldsymbol{Z}(\boldsymbol{s}) = \mu\boldsymbol{1} + \boldsymbol{e}(\boldsymbol{s}), \qquad \text{donde} \quad \boldsymbol{e}(\boldsymbol{s}) \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$$

# Kriging Ordinario

Considere el modelo

$$\boldsymbol{Z}(\boldsymbol{s}) = \boldsymbol{\mu}(\boldsymbol{s}) + \boldsymbol{e}(\boldsymbol{s}), \qquad \text{donde} \quad \boldsymbol{e}(\boldsymbol{s}) \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$$

Los supuestos sobre el Kriging Ordinario:

- $\mu(\boldsymbol{s}) = \mu, \ \ \forall \boldsymbol{s}$, con $\mu \in \mathbb{R}$ desconocido
- $\boldsymbol{\Sigma}$ es conocido

Es decir,

$$\boldsymbol{Z}(\boldsymbol{s}) = \mu\mathbf{1} + \boldsymbol{e}(\boldsymbol{s}), \qquad \text{donde} \quad \boldsymbol{e}(\boldsymbol{s}) \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$$

Nuevamente asumir que

$$P(Z, \boldsymbol{s}_0) = \lambda_0 + \boldsymbol{\lambda}^\top \boldsymbol{Z}, \qquad \boldsymbol{\lambda}^\top = (\lambda_1, \ldots, \lambda_n)\,, \quad \lambda_i \in \mathbb{R}.$$

La condición de insesgamiento:

$$\mathbb{E}[P(Z, \boldsymbol{s}_0)] = \mathbb{E}[Z(\boldsymbol{s}_0)] \quad \Longleftrightarrow \quad \lambda_0 + \boldsymbol{\lambda}^\top \mu \mathbf{1} = \mu$$

pero ahora $\lambda_0 = \mu(\boldsymbol{s}_0) - \boldsymbol{\lambda}^\top \mu \mathbf{1}$ no es posible ya que $\mu$ es desconocido .

# Kriging Ordinario

La condición de insesgamiento:

$$\mathbb{E}[P(Z, \boldsymbol{s}_0)] = \mathbb{E}[Z(\boldsymbol{s}_0)] \quad \Longleftrightarrow \quad \lambda_0 + \boldsymbol{\lambda}^\top \mu \mathbf{1} = \mu$$

pero ahora $\lambda_0 = \mu(\boldsymbol{s}_0) - \boldsymbol{\lambda}^\top \mu \mathbf{1}$ no es posible ya que $\mu$ es desconocido .

Al observar la condición de insesgamiento como un polinomio en $\mu$, obtenemos

$$\lambda_0 = 0, \quad \wedge \quad \boldsymbol{\lambda}^\top \mathbf{1} - 1 = 0 \implies \sum_{i=1}^{n} \lambda_i = 1$$

# Kriging Ordinario

El problema es

$$\min_{\boldsymbol{\lambda}} \mathbb{E}\left[\left(\boldsymbol{\lambda}^\top \boldsymbol{Z} - Z(\boldsymbol{s}_0)\right)^2\right] \quad \text{s.a.} \quad \boldsymbol{\lambda}^\top \mathbf{1} - 1 = 0$$

Utilizando multiplicadores de Lagrange, la función objetivo es

$$Q(\boldsymbol{\lambda}) = \boldsymbol{\lambda}^\top \boldsymbol{\Sigma} \boldsymbol{\lambda} + \sigma^2 - 2\boldsymbol{\lambda}^\top \boldsymbol{C} - 2m\left(\boldsymbol{\lambda}^\top \mathbf{1} - 1\right)$$

donde $m$ es el multiplicador de Lagrange.

# Kriging Ordinario

Al resolver $\nabla Q = \mathbf{0}$ se obtiene que

$$m = \frac{1 - \mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{C}}{\mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\mathbf{1}} \qquad \wedge \qquad \boldsymbol{\lambda}^\top = \left(\boldsymbol{C} + \mathbf{1}\frac{1 - \mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{C}}{\mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\mathbf{1}}\right)\boldsymbol{\Sigma}^{-1}$$

Obteniendo,

$$P_{OK}(Z, \boldsymbol{s}_0) \;\; = \;\; \widehat{\mu} + \boldsymbol{C}^\top \boldsymbol{\Sigma}^{-1}\left(\boldsymbol{Z} - \widehat{\mu}\mathbf{1}\right)$$

donde $\widehat{\mu} = \left(\mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\mathbf{1}\right)^{-1} \mathbf{1}\boldsymbol{\Sigma}^{-1}\boldsymbol{Z}$.

# Kriging Ordinario

Al resolver $\nabla Q = \mathbf{0}$ se obtiene que

$$m = \frac{1 - \mathbf{1}^\top \mathbf{\Sigma}^{-1} \boldsymbol{C}}{\mathbf{1}^\top \mathbf{\Sigma}^{-1} \mathbf{1}} \qquad \wedge \qquad \boldsymbol{\lambda}^\top = \left( \boldsymbol{C} + \mathbf{1} \frac{1 - \mathbf{1}^\top \mathbf{\Sigma}^{-1} \boldsymbol{C}}{\mathbf{1}^\top \mathbf{\Sigma}^{-1} \mathbf{1}} \right) \mathbf{\Sigma}^{-1}$$

Obteniendo,

$$P_{OK}(Z, \boldsymbol{s}_0) \;\; = \;\; \widehat{\mu} + \boldsymbol{C}^\top \mathbf{\Sigma}^{-1} \left( \boldsymbol{Z} - \widehat{\mu} \mathbf{1} \right)$$

donde $\widehat{\mu} = \left( \mathbf{1}^\top \mathbf{\Sigma}^{-1} \mathbf{1} \right)^{-1} \mathbf{1} \mathbf{\Sigma}^{-1} \boldsymbol{Z}$.

Además,

$$\sigma^2_{OK}(\boldsymbol{s}_0) = \sigma^2 - \boldsymbol{C}^\top \mathbf{\Sigma}^{-1} \boldsymbol{C} + \frac{\left( 1 - \mathbf{1}^\top \mathbf{\Sigma}^{-1} \boldsymbol{C} \right)^2}{\mathbf{1}^\top \mathbf{\Sigma}^{-1} \mathbf{1}}$$

**Observaciones**: Sea $C(\cdot)$ una función de covarianza isotrópica asociada el proceso $Z$.

**Observaciones**: Sea $C(\cdot)$ una función de covarianza isotrópica asociada el proceso $Z$.

1. $\mathbb{V}[Z(\boldsymbol{s}_0)] = \sigma^2 = C(0)$.
2. $\operatorname{cov}\left(Z(\boldsymbol{s}_i), Z(\boldsymbol{s}_0)\right) = C(\|\boldsymbol{s}_i - \boldsymbol{s}_0\|)$.
3. $\sigma^2_{SK}(\boldsymbol{s}_0) < \sigma^2_{OK}(\boldsymbol{s}_0)$.
4. Estimar $\boldsymbol{\Sigma}$ a partir del variograma.
5. Es posible escribir $P_{OK}$ y $\sigma^2_{OK}$ en términos del variograma.

**Modelamiento**:

1. Ajustar un variograma a la muestra.
2. Obtener $\widehat{\boldsymbol{\Sigma}}$, $\widehat{\mu}$, $\widehat{\boldsymbol{C}}$ y $\widehat{\sigma}^2$.
3. Calcular $P_{OK}(Z, \boldsymbol{s}_0)$ y $\sigma^2_{OK}(\boldsymbol{s}_0)$ para varios valores de $\boldsymbol{s}_0$

# Kriging Ordinario

**Modelamiento**:

1. Ajustar un variograma a la muestra.
2. Obtener $\widehat{\boldsymbol{\Sigma}}$, $\widehat{\mu}$, $\widehat{\boldsymbol{C}}$ y $\widehat{\sigma}^2$.
3. Calcular $P_{OK}(Z, \boldsymbol{s}_0)$ y $\sigma^2_{OK}(\boldsymbol{s}_0)$ para varios valores de $\boldsymbol{s}_0$

**Critica**: Asumir media constante es aún muy restrictivo.

# Kriging Universal

Considere el modelo

$$\boldsymbol{Z}(\boldsymbol{s}) = \boldsymbol{\mu}(\boldsymbol{s}) + \boldsymbol{e}(\boldsymbol{s}), \qquad \text{donde} \quad \boldsymbol{e}(\boldsymbol{s}) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$$

# Kriging Universal

Considere el modelo

$$\boldsymbol{Z}(\boldsymbol{s}) = \boldsymbol{\mu}(\boldsymbol{s}) + \boldsymbol{e}(\boldsymbol{s}), \qquad \text{donde} \quad \boldsymbol{e}(\boldsymbol{s}) \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$$

Los supuestos sobre el Kriging Universal:

- $\boldsymbol{\mu}(\boldsymbol{s}) = \boldsymbol{X}(\boldsymbol{s})\boldsymbol{\beta}, \ \ \forall \boldsymbol{s}$, con $\boldsymbol{\beta} \in \mathbb{R}^p$ desconocido.
- $\boldsymbol{\Sigma}$ es conocido

# Kriging Universal

Considere el modelo

$$\boldsymbol{Z}(\boldsymbol{s}) = \boldsymbol{\mu}(\boldsymbol{s}) + \boldsymbol{e}(\boldsymbol{s}), \qquad \text{donde} \quad \boldsymbol{e}(\boldsymbol{s}) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$$

Los supuestos sobre el Kriging Universal:

- $\boldsymbol{\mu}(\boldsymbol{s}) = \boldsymbol{X}(\boldsymbol{s})\boldsymbol{\beta}, \ \ \forall \boldsymbol{s}$, con $\boldsymbol{\beta} \in \mathbb{R}^p$ desconocido.
- $\boldsymbol{\Sigma}$ es conocido

Es decir,

$$\boldsymbol{Z}(\boldsymbol{s}) = \boldsymbol{X}(\boldsymbol{s})\boldsymbol{\beta} + \boldsymbol{e}(\boldsymbol{s}) \quad \text{y} \quad Z(\boldsymbol{s}_0) = \boldsymbol{X}(\boldsymbol{s}_0)\boldsymbol{\beta} + e(\boldsymbol{s}_0)$$

# Kriging Universal

Considere el modelo

$$\boldsymbol{Z}(\boldsymbol{s}) = \boldsymbol{\mu}(\boldsymbol{s}) + \boldsymbol{e}(\boldsymbol{s}), \qquad \text{donde} \quad \boldsymbol{e}(\boldsymbol{s}) \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$$

Los supuestos sobre el Kriging Universal:

- $\boldsymbol{\mu}(\boldsymbol{s}) = \boldsymbol{X}(\boldsymbol{s})\boldsymbol{\beta}, \;\; \forall \boldsymbol{s}$, con $\boldsymbol{\beta} \in \mathbb{R}^p$ desconocido.
- $\boldsymbol{\Sigma}$ es conocido

Es decir,

$$\boldsymbol{Z}(\boldsymbol{s}) = \boldsymbol{X}(\boldsymbol{s})\boldsymbol{\beta} + \boldsymbol{e}(\boldsymbol{s}) \quad \text{y} \quad Z(\boldsymbol{s}_0) = \boldsymbol{X}(\boldsymbol{s}_0)\boldsymbol{\beta} + e(\boldsymbol{s}_0)$$

**Obs**: $\boldsymbol{X}(\boldsymbol{s})$ tiene dimensión $n \times p$. $\boldsymbol{X}(\boldsymbol{s}_0)$ tiene dimensión $1 \times p$

# Kriging Universal

Nuevamente asumir que

$$P(Z, \boldsymbol{s}_0) = \lambda_0 + \boldsymbol{\lambda}^\top \boldsymbol{Z}, \qquad \boldsymbol{\lambda}^\top = (\lambda_1, \ldots, \lambda_n)\,, \quad \lambda_i \in \mathbb{R}.$$

# Kriging Universal

Nuevamente asumir que

$$P(Z, \boldsymbol{s}_0) = \lambda_0 + \boldsymbol{\lambda}^\top \boldsymbol{Z}, \qquad \boldsymbol{\lambda}^\top = (\lambda_1, \ldots, \lambda_n)\,, \quad \lambda_i \in \mathbb{R}.$$

La condición de insesgamiento:

$$\mathbb{E}[P(Z, \boldsymbol{s}_0)] = \mathbb{E}[Z(\boldsymbol{s}_0)] \quad \Longleftrightarrow \quad \lambda_0 + \boldsymbol{\lambda}^\top \boldsymbol{X}(\boldsymbol{s})\boldsymbol{\beta} = \boldsymbol{X}(\boldsymbol{s}_0)\boldsymbol{\beta}$$

Al observar la condición de insesgamiento como un polinomio en $\boldsymbol{\beta}$, obtenemos

$$\lambda_0 = 0, \quad \wedge \quad \boldsymbol{\lambda}^\top \boldsymbol{X}(\boldsymbol{s}) - \boldsymbol{X}(\boldsymbol{s}_0) = \boldsymbol{0}$$

Es decir, el problema de minimización tiene $p$ restricciones.

Considere la función objetivo

$$Q(\boldsymbol{\lambda}) = \boldsymbol{\lambda}^\top \boldsymbol{\Sigma} \boldsymbol{\lambda} + \sigma^2 - 2\boldsymbol{\lambda}^\top \boldsymbol{C} - 2\boldsymbol{m}^\top \left( \boldsymbol{\lambda}^\top \boldsymbol{X}(\boldsymbol{s}) - \boldsymbol{X}(\boldsymbol{s}_0) \right)$$

Resolviendo la ecuación $\nabla Q = \mathbf{0}$, se obtiene que

$$P_{UK}(Z, \boldsymbol{s}_0) = \boldsymbol{X}(\boldsymbol{s}_0)\widehat{\boldsymbol{\beta}} + \boldsymbol{C}^\top \boldsymbol{\Sigma}^{-1} \left( \boldsymbol{Z} - \boldsymbol{X}(\boldsymbol{s})\widehat{\boldsymbol{\beta}} \right)$$

donde $\widehat{\boldsymbol{\beta}} = \left( \boldsymbol{X}(\boldsymbol{s})^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{X}(\boldsymbol{s}) \right)^{-1} \boldsymbol{X}(\boldsymbol{s})^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{Z}$.

# Kriging Universal

Considere la función objetivo

$$Q(\boldsymbol{\lambda}) = \boldsymbol{\lambda}^\top \boldsymbol{\Sigma} \boldsymbol{\lambda} + \sigma^2 - 2\boldsymbol{\lambda}^\top \boldsymbol{C} - 2\boldsymbol{m}^\top \left( \boldsymbol{\lambda}^\top \boldsymbol{X}(\boldsymbol{s}) - \boldsymbol{X}(\boldsymbol{s}_0) \right)$$

Resolviendo la ecuación $\nabla Q = \mathbf{0}$, se obtiene que

$$P_{UK}(Z, \boldsymbol{s}_0) = \boldsymbol{X}(\boldsymbol{s}_0)\widehat{\boldsymbol{\beta}} + \boldsymbol{C}^\top \boldsymbol{\Sigma}^{-1} \left( \boldsymbol{Z} - \boldsymbol{X}(\boldsymbol{s})\widehat{\boldsymbol{\beta}} \right)$$

donde $\widehat{\boldsymbol{\beta}} = \left( \boldsymbol{X}(\boldsymbol{s})^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{X}(\boldsymbol{s}) \right)^{-1} \boldsymbol{X}(\boldsymbol{s})^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{Z}$.

Además, si $\boldsymbol{A} = \left( \boldsymbol{X}(\boldsymbol{s}_0)^\top - \boldsymbol{X}(\boldsymbol{s})^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{C} \right)^\top$, entonces

$$\sigma^2_{UK}(\boldsymbol{s}_0) = \sigma^2 - \boldsymbol{C}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{C} + \boldsymbol{A}^\top \left( \boldsymbol{X}(\boldsymbol{s})^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{X}(\boldsymbol{s}) \right)^{-1} \boldsymbol{A}$$

# Kriging Universal

**Observaciones**:

- Si $\boldsymbol{X}(\boldsymbol{s}) = \mathbf{1}$ se recuperan las ecuaciones del Kriging Ordinario.
- $\boldsymbol{X}(\boldsymbol{s})$ es la matriz de diseño y puede contener las direcciones o bien variables exógenas.
- Para estimar $\boldsymbol{\beta}$ y $\boldsymbol{\Sigma}$ al mismo tiempo usar el método de estimación REML.
- En la libreria **geoR** existen las funciones krige.conv y ksline para realizar las predicciones mediante kriging lineales.

# Bibliografia


Banerjee S., Carlin B., and Gelfand A. (2015) *Hierarchical Modeling and Analysis for Spatial Data*, Boca Raton: Chapman Hall/CRC.

Cressie, N. (1993) *Statistics for Spatial Data*. New York: Wiley.

Matheron, G., (1965) Les variables régionalisés et leur estimation. Masson, Paris.

Shabenberger, O., Gotway, C. A. (2005) *Statistical Methods for Spatial Data Analysis*. Boca Raton, FL: Chapman & Hall/CRC.

Rousseeuw, P. J., Croux, C., (1993). Alternatives of the median absolute deviation. *Journal of American Statistics Association*, **88**, 1273–1283.

## ¿Alguna Consulta?