Institución Universitaria
Acreditada en Alta Calidad

Análisis de datos espacio-temporales

Docente: Cristian Guarnizo-Lemus
Correo: cristianguarnizo@itm.edu.co

Somos Innovación Tecnológica con Sentido Humano

Alcaldía de Medellín

Contenido

1. Definiciones
2. Krigging
3. Interpoladores
4. Análisis Espacio-Temporal

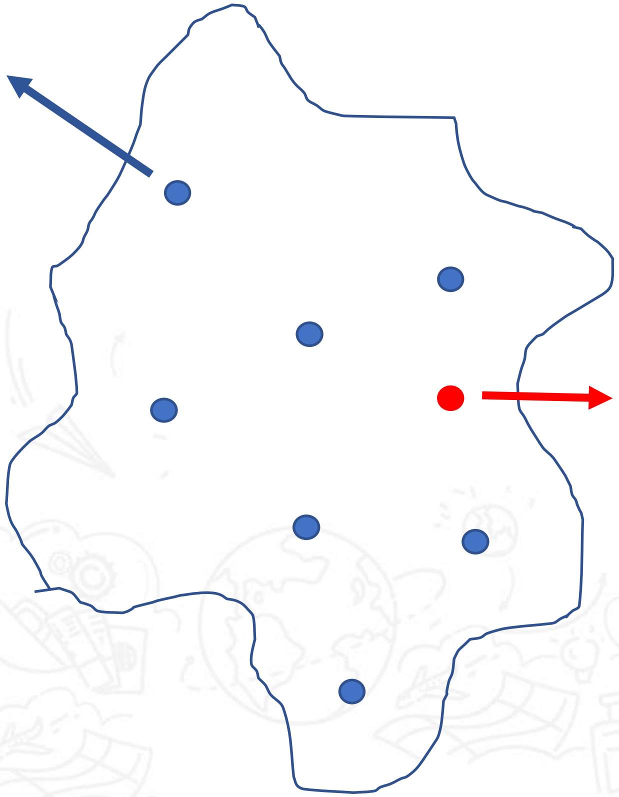
1. Campos Aleatorios

Un proceso estocástico es una familia o colección de variables aleatorias, los miembros pueden ser localizados (indexados) a partir de una métrica. Por ejemplo una serie de tiempo $Y(t)$, $t = t_1, \ldots, t_n$, es indexada por instantes de tiempo en los cuales la serie es observada. Similarmente, un proceso espacial es una colección de variables aleatorias indexadas por coordenadas espaciales. Para un plano en 2D, $s = [x, y]^T$.

1. Modelo espacial

z es un proceso aleatorio que depende de $\boldsymbol{s}$.

$$z = f(\boldsymbol{s}_1)$$
$$\boldsymbol{s}_1 = [x_1, y_1]$$



$$z_* = f(\boldsymbol{s}_*)$$
$$\boldsymbol{s}_* = [x_*, y_*]$$

$$\boldsymbol{z} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_N \end{bmatrix}$$

$$\boldsymbol{S} = \begin{bmatrix} \boldsymbol{s}_1 \\ \boldsymbol{s}_2 \\ \vdots \\ \boldsymbol{s}_N \end{bmatrix} = \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ \vdots & \vdots \\ x_N & y_N \end{bmatrix}$$

$$z = f(\boldsymbol{s}) + \epsilon$$

$$f(\boldsymbol{s}) \sim GP(0, k(\boldsymbol{s}, \boldsymbol{s}'))$$

$$\epsilon \sim N(0, \sigma^2)$$

$$\boldsymbol{z} = N(0, \boldsymbol{K} + \sigma^2 \boldsymbol{I})$$

1. Modelo espacial

z es un proceso aleatorio que depende de $\boldsymbol{s}$.

$$\boldsymbol{K} = \begin{bmatrix} k(\boldsymbol{s}_1, \boldsymbol{s}_1) & \dots & k(\boldsymbol{s}_1, \boldsymbol{s}_N) \\ \vdots & \ddots & \vdots \\ k(\boldsymbol{s}_N, \boldsymbol{s}_1) & \dots & k(\boldsymbol{s}_N, \boldsymbol{s}_N) \end{bmatrix}$$

$$\begin{bmatrix} \boldsymbol{z} \\ \boldsymbol{z}_* \end{bmatrix} = N\left(\begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{0} \end{bmatrix} \middle| \begin{bmatrix} \boldsymbol{K} & \boldsymbol{K}_{\boldsymbol{z}\boldsymbol{z}_*} \\ \boldsymbol{K}_{\boldsymbol{z}_*\boldsymbol{z}} & \boldsymbol{K}_{\boldsymbol{z}_*\boldsymbol{z}_*} \end{bmatrix}\right)$$

$$\boldsymbol{K}_{\boldsymbol{z}_*\boldsymbol{z}} = \begin{bmatrix} k(\boldsymbol{s}_1^*, \boldsymbol{s}_1) & \dots & k(\boldsymbol{s}_1^*, \boldsymbol{s}_N) \\ \vdots & \ddots & \vdots \\ k(\boldsymbol{s}_T^*, \boldsymbol{s}_1) & \dots & k(\boldsymbol{s}_T^*, \boldsymbol{s}_N) \end{bmatrix}$$

$$\boldsymbol{z}_* | \boldsymbol{z} \sim N(\boldsymbol{\mu}_*, \boldsymbol{\Sigma}_*)$$

$$\boldsymbol{\mu} = \boldsymbol{K}_{\boldsymbol{z}_*\boldsymbol{z}} \boldsymbol{K}^{-1} \boldsymbol{z}$$

$$\boldsymbol{\Sigma} = \boldsymbol{K}_{\boldsymbol{z}_*\boldsymbol{z}_*} - \boldsymbol{K}_{\boldsymbol{z}_*\boldsymbol{z}} \boldsymbol{K}^{-1} \boldsymbol{K}_{\boldsymbol{z}\boldsymbol{z}_*}$$

Institución Universitaria
Acreditada en Alta Calidad

1. Krigging

Se asume un proceso estacionario, se asume conocida la media y se basa en un función de covarianza.

$$z_* = f(\boldsymbol{s}_*) = \sum_{n=1}^{N} \omega_n f_n(\boldsymbol{s}_n) + e(\boldsymbol{s}_*)$$

$$E[z_*] = \sum_{n=1}^{N} \omega_n f_n(\boldsymbol{s}_n) = \boldsymbol{\omega}^T \boldsymbol{z} = \boldsymbol{z}^T \boldsymbol{\omega}$$

$$\begin{bmatrix} \omega_1 \\ \omega_2 \\ \vdots \\ \omega_N \end{bmatrix} = \begin{bmatrix} c(\boldsymbol{s}_1, \boldsymbol{s}_1) & \dots & c(\boldsymbol{s}_1, \boldsymbol{s}_N) \\ \vdots & \ddots & \vdots \\ c(\boldsymbol{s}_N, \boldsymbol{s}_1) & \dots & c(\boldsymbol{s}_N, \boldsymbol{s}_N) \end{bmatrix}^{-1} \begin{bmatrix} c(\boldsymbol{s}_1, \boldsymbol{s}_*) \\ c(\boldsymbol{s}_2, \boldsymbol{s}_*) \\ \vdots \\ c(\boldsymbol{s}_N, \boldsymbol{s}_*) \end{bmatrix}$$

$$\boldsymbol{\omega}^T = \boldsymbol{K}_{\boldsymbol{z}\boldsymbol{z}_*} \boldsymbol{K}^{-1}$$

$$E[z_*] = \boldsymbol{\omega}^T \boldsymbol{z} = \boldsymbol{K}_{\boldsymbol{z}\boldsymbol{z}_*} \boldsymbol{K}^{-1} \boldsymbol{z}$$

Del GP se tenia:

$$\boldsymbol{\mu} = \boldsymbol{K}_{\boldsymbol{z}_*\boldsymbol{z}} \boldsymbol{K}^{-1} \boldsymbol{z}$$

1. Función de covarianza espacial

Para determinar las correlaciones entre los puntos del espacio, usamos RBF (*radial basis function*), también conocida como Exponencial cuadrática.

$$k(\boldsymbol{s_i}, \boldsymbol{s}_j') = \sigma_s^2 \exp\left(-\sum_{m=1}^{d} \frac{\left(s_{im} - s_{jm}'\right)^2}{l_m^2}\right)$$

Donde σ_s^2 es la varianza, d es el numero de dimensiones y l_m es la longitud de escala espacial (*lenghtscale*) para cada dimensión.

1. Función de covarianza espacial

También se puede emplear la función kernel matern.

$$k(\boldsymbol{s_i}, \boldsymbol{s_j'}) = \sigma_s^2 (1 + \sqrt{3}\, r) \exp(-\sqrt{3}\, r)\, ,\ r = \sqrt{\sum_{m=1}^{d} \frac{\left(\boldsymbol{s}_{im} - \boldsymbol{s}'_{jm}\right)^2}{l_m^2}}$$

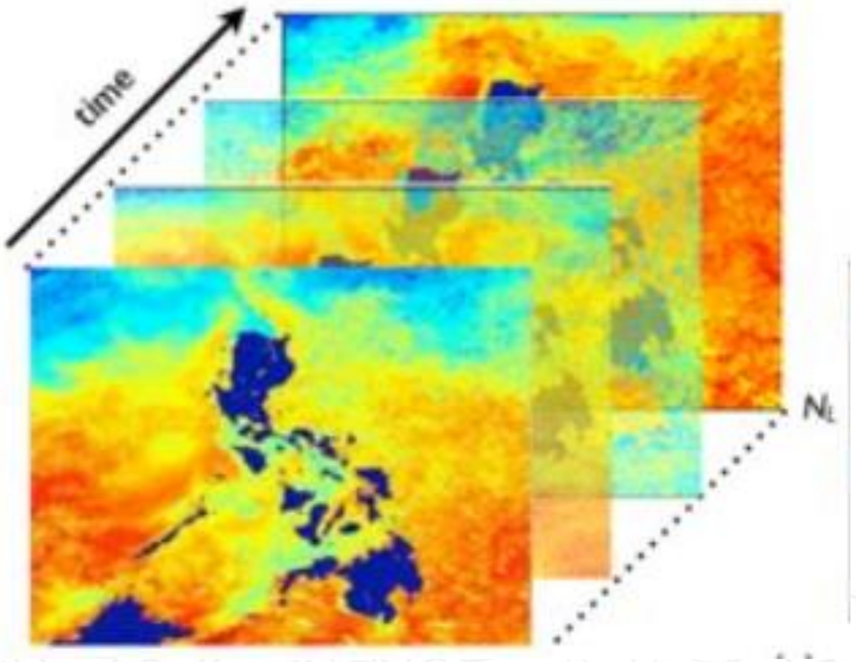
Donde σ_s^2 es la varianza, d es el numero de dimensiones y l_m es la longitud de escala espacial (*lenghtscale*) de la dimensión m-esima.

Contenido

1. Campos Aleatorios
2. Análisis Espacio-Temporal

2. Modelo espacio-temporal

z es un proceso aleatorio que depende de $\boldsymbol{s}$.



$$z = f(\boldsymbol{s}, t) + \epsilon(\boldsymbol{s}) + n(t)$$

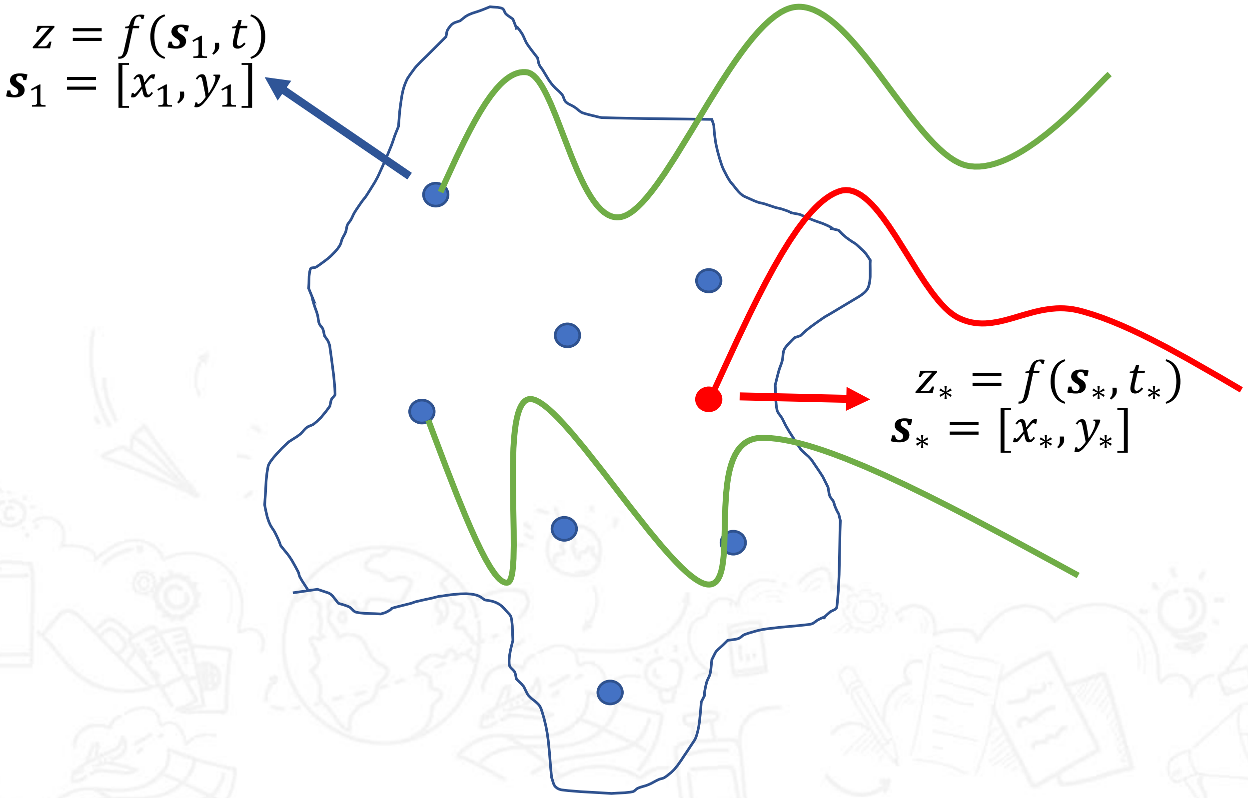
$$f(\boldsymbol{s}) \sim GP(0, k(\boldsymbol{s}, \boldsymbol{s}'))$$

$$\epsilon \sim N(0, \sigma^2) \quad n \sim N(0, \sigma_t^2)$$

$$\boldsymbol{z} = N(0, \boldsymbol{K} + \sigma^2 \boldsymbol{I} + \sigma_t^2 \boldsymbol{I})$$

2. Modelo espacio-temporal

z es un proceso aleatorio que depende de $\boldsymbol{s}$.



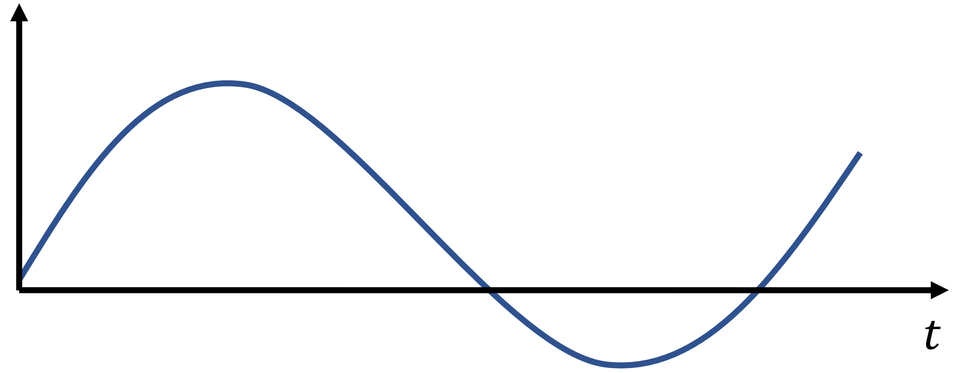
$$z = f(\boldsymbol{s}, t) + \epsilon + n(t)$$

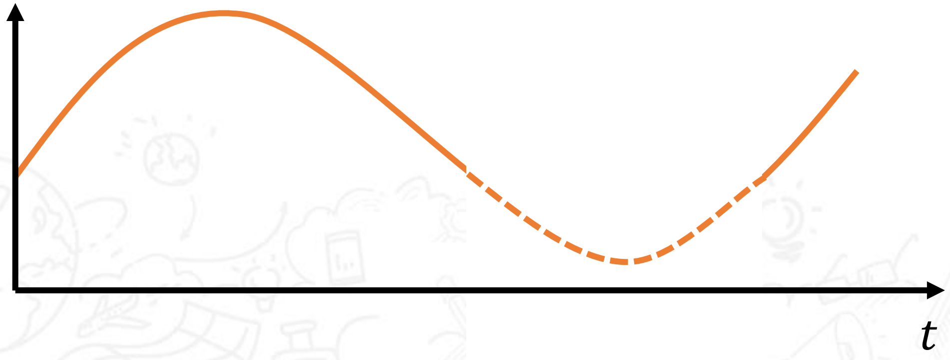
$$f(\boldsymbol{s}) \sim N(0, k(\boldsymbol{s}, \boldsymbol{s}'))$$

$$\epsilon \sim N(0, \sigma^2) \quad n \sim N(0, \sigma_t^2)$$

$$\boldsymbol{z} = N(0, \boldsymbol{K} + \sigma^2 \boldsymbol{I} + \sigma_t^2 \boldsymbol{I})$$

2. Correlación temporal





2. Función de covarianza temporal

Si se requiere modelar un comportamiento cíclico, se puede emplear un kernel periódico, así

$$k(t,t') = \theta_t \exp\left(-\frac{1}{2l_t^2}\sin\left(\frac{\pi(t-t')}{T}\right)\right)$$

donde l_t es la longitud de escala del proceso; θ_t refleja su variaza y T es el periodo de oscilación.

2. Función de covarianza temporal

Dependiendo del comportamiento de los datos, se pueden emplar:

$$k(t,t') = \sigma_t^2 \exp\left(-\frac{(t-t')^2}{l_t^2}\right)$$

$$k(t,t') = \sigma_t^2\left(1+\sqrt{3}\,r\right)\exp\left(-\sqrt{3}\,r\right), \quad r = \sqrt{\frac{(t-t')^2}{l_t^2}}$$

2. Función de covarianza compuesta

Dependiendo del comportamiento de los datos, se pueden emplar:

$$k(t,t') = \sigma_t^2 \exp\left(-\frac{(t-t')^2}{l_t^2}\right)$$

$$k(t,t') = \sigma_t^2\left(1+\sqrt{3}\,r\right)\exp\left(-\sqrt{3}\,r\right), \quad r = \sqrt{\frac{(t-t')^2}{l_t^2}}$$

2. Modelo Espacio-Temporal

$$\boldsymbol{z} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_N \end{bmatrix} \qquad \boldsymbol{U} = [\boldsymbol{S} \ \ \boldsymbol{t}] = \begin{bmatrix} \boldsymbol{u}_1 \\ \boldsymbol{u}_2 \\ \vdots \\ \boldsymbol{u}_N \end{bmatrix} = \begin{bmatrix} x_1 & y_1 & t_1 \\ x_2 & y_2 & t_2 \\ \vdots & \vdots & \vdots \\ x_N & y_N & t_N \end{bmatrix} \qquad \boldsymbol{\mu} = \boldsymbol{K}_{\boldsymbol{z}_*\boldsymbol{z}}\boldsymbol{K}^{-1}\boldsymbol{z}$$

$$\boldsymbol{K} = \begin{bmatrix} k(\boldsymbol{u}_1, \boldsymbol{u}_1) & \dots & k(\boldsymbol{u}_1, \boldsymbol{u}_N) \\ \vdots & \ddots & \vdots \\ k(\boldsymbol{u}_N, \boldsymbol{u}_1) & \dots & k(\boldsymbol{u}_N, \boldsymbol{u}_N) \end{bmatrix} \qquad \boldsymbol{K}_{\boldsymbol{z}_*\boldsymbol{z}} = \begin{bmatrix} k(\boldsymbol{u}_1^*, \boldsymbol{u}_1) & \dots & k(\boldsymbol{u}_1^*, \boldsymbol{u}_N) \\ \vdots & \ddots & \vdots \\ k(\boldsymbol{u}_T^*, \boldsymbol{u}_1) & \dots & k(\boldsymbol{u}_T^*, \boldsymbol{u}_N) \end{bmatrix}$$

Institución Universitaria
Acreditada en Alta Calidad

2. Modelo Espacio-Temporal

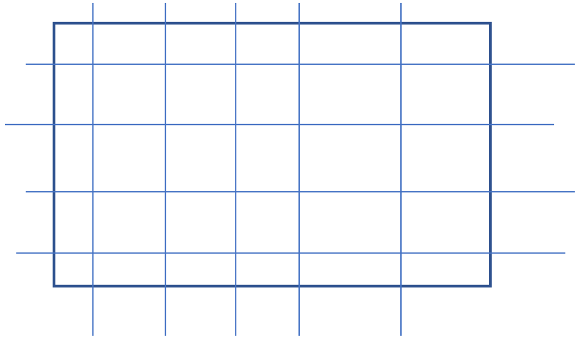
$$k((\boldsymbol{s},t),(\boldsymbol{s}',t')) = k_s(\boldsymbol{s},\boldsymbol{s}';\ \theta_{\boldsymbol{s}})k_t(t,t';\theta_t)$$

http://proceedings.mlr.press/v22/luttinen12.html

$$k((\boldsymbol{s},t),(\boldsymbol{s}',t')) = k_{t,m}(t,t';\theta_t) + k_{s,m}(\boldsymbol{s},\boldsymbol{s}';\ \theta_{\boldsymbol{s}}) + k_{t,i}(t,t';\theta_t) \times k_{s,i}(\boldsymbol{s},\boldsymbol{s}';\ \theta_{\boldsymbol{s}})$$

https://arxiv.org/pdf/1806.10873.pdf

3. Interpoladores



Herramientas:

Análisis Espacial

Python

Análisis: SciPy, PySal, PyInterpolate, PySpatialML

Datos: GeoPandas

https://pythonnumericalmethods.berkeley.edu/notebooks/chapter17.05-Newtons-Polynomial-Interpolation.html

Institución Universitaria
Acreditada en Alta Calidad

¡Gracias!

Somos Innovación Tecnológica con Sentido Humano

Alcaldía de Medellín