BST 6200 Spatial Statistics and Disease Mapping

Section 6.8 Kriging

Steven E. Rigdon

Spring 2020

Section 6.8: Kriging

The model:

Let $z(\boldsymbol{x})$ denote the outcome at location $\boldsymbol{x}$, where $\boldsymbol{x}$ is a vector (i.e., a point in the plane).

Let $f(\boldsymbol{x}_i)$ denote the deterministic (not random) trend function evaluated at $\boldsymbol{x}_i$.

$$z(\boldsymbol{x}_i) = f(\boldsymbol{x}_i) + \nu(\boldsymbol{x}_i) + \epsilon_i$$

where $\nu()$ is a **random function** and $\epsilon_i, \quad i = 1, 2, \ldots$ are i.i.d. $N(0, \sigma^2)$. [More on random functions coming up; this is a fairly deep and slippery concept!]

The Model for Kriging

$$\underbrace{z(\boldsymbol{x}_i)}_{\text{Outcome}} = \underbrace{f(\boldsymbol{x}_i)}_{\text{True Function}} + \underbrace{\nu(\boldsymbol{x}_i)}_{\text{Random Function}} + \underbrace{\epsilon_i}_{N(0,\sigma^2)\text{ error}}$$

Multivariate Normal Distribution

$\boldsymbol{X} = [X_1, X_2, \ldots, X_p]'$. $\boldsymbol{X} \sim \text{MVN}(\boldsymbol{\mu}, \Sigma)$ means that $\boldsymbol{X}$ has a multivariate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix Σ with PDF

$$f(\boldsymbol{x}) = \frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})'\Sigma^{-1}(\boldsymbol{x} - \boldsymbol{\mu})\right)$$

Mean, Variance, Covariance, and Correlation of Random Variables

$$E(X) = \int_{-\infty}^{\infty} x \; p(x) \; dx = \mu_X$$

$$E(X^2) = \int_{-\infty}^{\infty} x^2 \; p(x) \; dx$$

$$V(X) \stackrel{\text{def}}{=} E\left[(X - \mu_X)^2\right] = E(X^2) - \mu_X^2 = \sigma_X^2$$

$$\text{covariance } = \sigma_{X,Y} = \text{cov}(X, Y) = E\left[(X - \mu_X)(Y - \mu_Y)\right]$$

$$\text{correlation } = \rho_{X,Y} = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y}$$

Mean and Covariance of Multivariate Normal Distribution

If $\boldsymbol{X} \sim \mathrm{MVN}(\boldsymbol{\mu}, \Sigma)$ then

$$E(\boldsymbol{X}) = \boldsymbol{\mu}$$

and

$$\mathrm{cov}(\boldsymbol{X}) \stackrel{\mathrm{def}}{=} \begin{bmatrix} V(X_1) & \mathrm{cov}(X_1, X_2) & \mathrm{cov}(X_1, X_3) \\ \mathrm{cov}(X_1, X_2) & V(X_2) & \mathrm{cov}(X_2, X_3) \\ \mathrm{cov}(X_1, X_3) & \mathrm{cov}(X_2, X_3) & V(X_3) \end{bmatrix} = \Sigma$$

Random Function in One Variable

Spatial data analysis requires us to consider random functions in two dimensions, because the plane is two dimensional. Let's begin with random functions in one dimension to get the concept down.

Suppose x_1, x_2, x_3 are three points on a (one-dimensional) number line, and let $\nu(x_1), \nu(x_2), \nu(x_3)$ denote the outcome at these three points.

Suppose

$$\begin{bmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{bmatrix} \sim \text{MVN} \left(\begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \sigma_{1,2} & \sigma_{1,3} \\ \sigma_{1,2} & \sigma_2^2 & \sigma_{2,3} \\ \sigma_{1,3} & \sigma_{2,3} & \sigma_3^2 \end{bmatrix} \right)$$

What about the covariance (or correlation) between the random function at two nearby x-points?

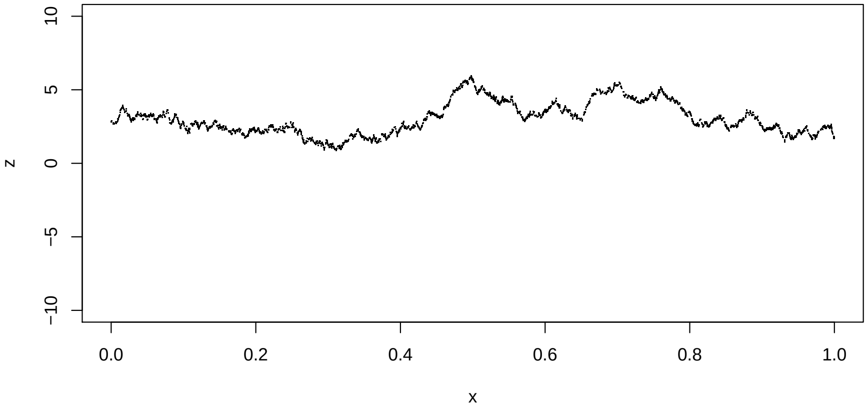
We make the assumption that the correlation between the outcomes ν_1 and ν_2 at points x_1 and x_2, respectively, is a function of the distance between them:

$$\rho_{ij} = \rho(d)$$

where d is the distance between x_1 and x_2.

Note, in the case of a single variable (as we've assumed here), $d = |x_1 - x_2|$. In the case of points in the plane, $d = \|\boldsymbol{x}_1 - \boldsymbol{x}_2\|$, the distance between the two points.

Using R to Visualize a Random Function



Back to Vectors in the Plane

The correlation between the outcomes ν_1 and ν_2 at points $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, respectively, is a function of the distance between them:

$$\rho_{ij} = \rho(\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|)$$

If the outcomes $[\nu(x_1), \nu(x_2), \ldots, \nu(x_p)]'$ are multivariate normal, then the limit as $n \to \infty$ of this process is said to be a **Gaussian random field**.

A Gaussian random field is **isotropic**, which means that the covariance function between points $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ depends only on the distance between $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ and not on the direction.

A function that does not have this property is said to be **anisotropic**.

Assumptions Behind Kriging

$$\underbrace{z(\boldsymbol{x}_i)}_{\text{Outcome}} = \underbrace{f(\boldsymbol{x}_i)}_{\text{True Function}} + \underbrace{\nu(\boldsymbol{x}_i)}_{\text{Random Function}} + \underbrace{\epsilon_i}_{N(0,\sigma^2)\text{ error}}$$

Theory behind the variogram and the semi-variogram

A few facts to recall ...

$$V(cX) = c^2 V(X)$$

$$V(X+Y) = V(X) + V(Y) + 2\text{cov}(X, Y)$$

$$V(X-Y) = V(X) + V(Y) - 2\text{cov}(X, Y)$$

The Variogram

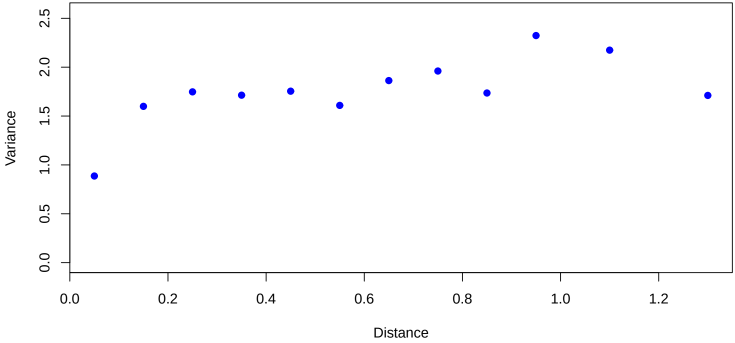
Let $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ be distinct points in the plane.

$$V(\nu(\boldsymbol{x_1}) - \nu(\boldsymbol{x_2})) =$$

$$V(\nu(\boldsymbol{x_1}) - \nu(\boldsymbol{x_2})) = 2\,\underbrace{\sigma^2(1 - \rho(\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|))}_{\text{Semi-variogram: } \gamma(\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|)}$$

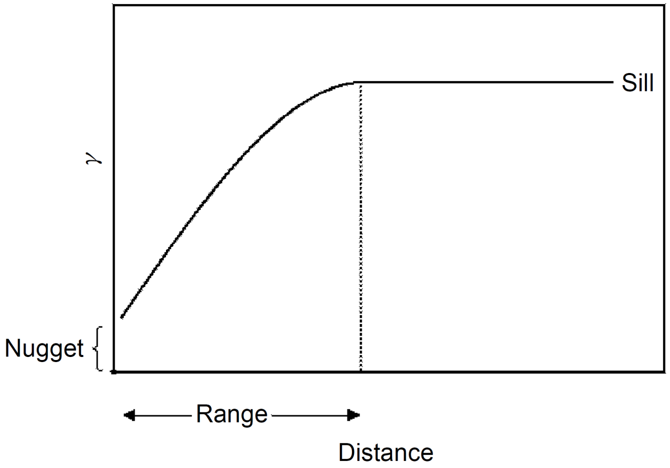
Use R to estimate the semi-variogram

Fitting a (Smooth) Model to the Semivariogram

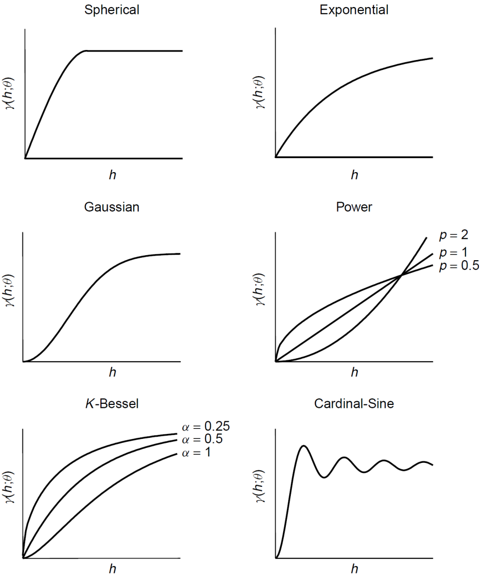


Terminology for Semivariogram

Sill: the z-value where the semivariogram first becomes flat
Range: the d value where the semivariogram first becomes flat
Nugget: the z-value at a distance of 0



Functional Forms for the Semi-Variogram



Functional Forms for the Semi-Variogram

Name	Abbrev.	Functional Form
Exponential	"Exp"	$\gamma(d) = a\left(1 - \exp(-d/b)\right)$
Gaussian	"Gau"	$\gamma(d) = a\left(1 - \exp(-\frac{1}{2}d^2/b^2)\right)$
Matern	"Mat"	$\gamma(d) = a\left(1 - \frac{(d/b)^\kappa}{2^{\kappa-1}\Gamma(\kappa)}\ K_\kappa\left(\frac{d}{b}\right)\right)$
Spherical	"Sph"	$\gamma(d) = \begin{cases} a\left(\frac{3}{2}\frac{d}{b} - \frac{1}{2}\left(\frac{d}{b}\right)^3\right) & \text{if } d \leq b \\ a & \text{if } d > b \end{cases}$

In the definition of the Matern function, $K_\kappa()$ is the modified Bessel function of the second kind.

St. Louis Trivia: Bessel's niece married Gauss's son and they lived in St. Louis. Both are buried in Bellefontaine Cemetary in north St. Louis city.

CHARITON COURIER.

Keytesville, Mo. Aug. 30, 1879

C. W. Gauss senior member of the boot and shoe house of Gauss & Son, St. Louis, died in that city last Saturday morning of dropsy. He was a son of Prof. C. F. Gauss, the distinguished astronomer and mathematician. He was himself a graduate of the celebrated University of Gottingen, in Germany, and was a man of considerable attainment. Arriving in America forty-two years ago, he came to Missouri and settled in this county, as a farmer, and remained here until 1856, when he moved to St. Louis and entered the wholesale boot and shoe business, in which he continued up to the time of his death. He was a relative of Mrs. Montague, now living in this place, and father of Dr. G. W. Gauss, who formerly practiced medicine and conducted the drug business here in connection with Mr. J. C. Crawley.

Kriging (Finally!): How Do You Use the Semi-Variogram to Predict Outcome at Points other than the Collection Points?

Predict the outcome at $\boldsymbol{x}_{\text{new}}$ by taking a weighted average of all other data points:

$$\hat{z}(\boldsymbol{x}_{\text{new}}) = \sum_{i=1}^{n} w_i z(\boldsymbol{x}_i)$$

This is similar to inverse distance weighting, where the weights were selected to be the distances raised to a negative power.

Here we use the estimated semivariogram to select optimal values for the weights w_i.

Determining the Weights

Select $w_1, \ w_2, \ \ldots, \ w_n$ to minimize the **mean squared prediction error** (MSPE)

$$E\left[\left(\sum_{i=1}^{n} w_i z(\boldsymbol{x}_i) - z(\boldsymbol{x}_{\text{new}})\right)^2\right]$$

subject to the constraint that

$$\sum_{i=1}^{n} w_i = 1.$$

Apply some calculus, including Lagrange multipliers, to find the constrained minimum.

Optimal Weights

After lots of messy calculus and algebra, we would find that the weights must satisfy

$$\begin{bmatrix} \gamma(d_{11}) & \gamma(d_{12}) & \cdots & \gamma(d_{1n}) & 1 \\ \gamma(d_{12}) & \gamma(d_{22}) & \cdots & \gamma(d_{2n}) & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ \gamma(d_{1n}) & \gamma(d_{2n}) & \cdots & \gamma(d_{nn}) & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \\ \lambda \end{bmatrix} = \begin{bmatrix} \gamma(d_1) \\ \gamma(d_2) \\ \vdots \\ \gamma(d_n) \\ 1 \end{bmatrix}$$

where $\gamma()$ is the smooth estimate of the semivariogram using one of the functional forms (Exponential, Matern, etc.)

$d_{ij} =$ distance between points $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$

$d_i =$ distance between points $\boldsymbol{x}_i$ and $\boldsymbol{x}_{\text{new}}$

$\lambda =$ value of the Lagrange multiplier

Ordinary Kriging

The method described here is called **ordinary kriging**. It assumes a stationary process with unknown mean function.

Steps in Applying Ordinary Kriging

1. Determine points for the semivariogram
2. Fit a continuous model to the points from step 1
3. Determine the weights by solving the equations on the previous slide
4. Make predictions at each point $\boldsymbol{x}_{\text{new}}$ in a grid and plot the surface or contour.

Using R

Use R to calculate the kriging predictor and plot it.