Kriging Notes

December 13, 2020

1 Introduction

Information from ARCGIS

Kriging is a way to interpolate data and estimate a surface for geospatial data. There are two main ways of doing this:

- Inverse distance weighting (IDW) and spline methods

- Kriging

IDW are deterministic interpolation methods as they are directly based on the surrounding values or smoothed formulas. Kriging is different as it uses autocorrelation and takes position into account in the statistical models. Kriging uses a certain number of neighbouring points, or all points within a specified radius (cf kNN).
The general formula for IDW and kriging is:

$$\hat{Z}\left(s_0\right) = \sum_{i=1}^{N} \lambda_i Z\left(s_i\right)$$

where: $Z\left(s_i\right)$ = the measured value at the ith location
λ_i = an unknown weight for the measured value at the ith location
s_0 = the prediction location
N = the number of measured values

The difference between IDW and kriging is that in IDW, λ_i only depends on distance to prediction location. In kriging λ_i also depends on autocorrelation i.e. spatial relationship between prediction locations.

2 Creating a prediction map with kriging

There are two steps:

1. Create the variograms and covariance functions to estimate the spatial autocorrelation values that depend on the model of autocorrelation (fitting a model).

2. Predict the unknown values

2.1 Variography (spatial modelling/structural analysis)

There are often too many pairs of spatial points to calculate and plot the distance for each pair. Instead, spatial distances are put into lag bins i.e. all points in the range $40m < x \leq 50m$ of point A, and calculate the semivariance. The semivariance is equal to half the variance of the differences between all possible points spaced a constant distance apart.

$$\gamma(h) = \frac{1}{2}\left[z(x_i) - z(x_j)\right]^2$$

Plotting the distance vs semivariance produces an empirical semivariogram. Closer items should be more similar, therefore lower semivariance. The opposite is true for further points.
A model is fit to the empirical semivariogram (cf regression). Different types of models can be fit the the semivariogram:

- Spherical (most common)

- Circular

- Exponential

- Gaussian

- Linear