Exercise 3 - Variogram fitting

Office hours: Friday 09:00-12:00

A. Berne alexis.berne@epfl.ch
G. Ghiggi gionata.ghiggi@epfl.ch
M. Guidicelli matteo.guidicelli@epfl.ch

The objectives of this third exercise are (1) to investigate different data transformation techniques and (2) to fit theoretical variogram models on isotropic sample variograms.

A R code (ex3.R) showing how to read and analyze the data is already provided. Information about the syntax and arguments of R functions is available via the *help* within R. A short but useful tutorial is also available on moodle.

1. Plot and show the empirical probability density function (pdf) of the cadmium and lead concentrations. Hint: use the code from exercise 1.

2. Apply a square root transform on the values of cadmium and lead. Show the empirical probability density function of the transformed concentration values. Do the same using a log transform.

3. Apply the Shapiro-Wilk test to assess which transformation (sqrt or log) provides the best results, i.e., the most Gaussian like distributions. Hint: use the function "shapiro.test()" included in the stats package.

4. Compute the isotropic sample variogram of log(cadmium) using the provided boundaries. Fit a spherical variogram model on it. Show the sample variogram and the fitted model on the same graph. What are the fitted values of the nugget, the sill and the range? Hint: use the functions vgm() and fit.variogram().

5. Compute the isotropic sample variogram of log(lead) using the provided boundaries. Fit both a spherical and an exponential variogram model on it. Which fit is the best? Hint: use the function "attributes()" to access the different attributes of your variogram model. Have a look at the attribute named "SSErr".

6. Show that $\text{Var}\left[\sum_{i=1}^{n} \lambda_i X_i\right] = \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j \text{Cov}[X_i, X_j]$