University of California, Los Angeles
Department of Statistics

Statistics C173/C273 **Instructor: Nicolas Christou**

Homework 9

Exercise 1
Use the variable zinc of the soil data to perform cross validation with `gstat` and `geoR`.

```
a <- read.table("http://www.stat.ucla.edu/~nchristo/statistics_c173_c273/soil.txt",
     header=TRUE)
```

a. Using `gstat`:

1. Split the data into two parts (one for modeling and one for cross validation). Create a gstat object and use it to compute the sample variogram and fit the spherical and exponential variograms to it. Predict the points of the cross validation part of the data set and compare the prediction sum of squares (`PRESS`) for each variogram.
2. Delete one point at a time and use the remaining $n-1$ points to predict it. Compare the `PRESS` for the two variograms.

b. Using `geoR`: Create a data frame with `x, y, zinc`. Convert the data frame into a geodata object and use it to compute the sample variogram. Fit the spherical and exponential variograms to it and finally use the `xvalid` function with `reest=TRUE` to compare the two `PRESS`.

c. Based on your analysis above, which model variogram will choose?

d. Please submit all the `R` commands and results from questions (a) and (b).

Exercise 2
Access the elevation data (see also homework 3, exercise 3):

```
a <- read.table("http://www.stat.ucla.edu/~nchristo/statistics_c173_c273/
     elevation_data.txt", header=TRUE)
```

From *Statistics and Data Analysis in Geology* (Second edition), Davis, J. C. (1972).

a. Use `geoR` or `gstat` to perform ordinary kriging on a dense grid of your choice.

b. Use `geoR` or `gstat` to perform universal kriging on the same grid of part (a). Would you choose ordinary or universal kriging?

c. Plot the raster maps of the predicted values and their kriging variances.

Exercise 3
The following data give the location (x, y coordinates) and the calcium content at depth 0-20 cm ($ca20$), for each data point. There are 178 data points. Please access the data at:

```
a <- read.table("http://www.stat.ucla.edu/~nchristo/statistics_c173_c273/
     soil_ca_data.txt", header=TRUE)
```

a. Create a grid for spatial predictions (`by=10`).

b. Create a `gstat` object assuming that there is a linear trend in the data (on the coordinates `x, y`).

c. Plot the semivariogram up to a maximum distance of 510 m.

d. Fit the spherical semivariogram to the sample semivariogram above using Cressie's weights.

e. Perform universal kriging (linear trend on the coordinates).

f. Collapse the vector of the predicted values into a matrix and use the `image` function to create a raster map. Add contours to the raster map.

g. Collapse the vector of the variances of the predicted values into a matrix and use the `image` function to create a raster map. Add contours to the raster map.

Exercise 4
Consider universal kriging. In matrix/vector form universal kriging minimizes

$$C(0) - 2\mathbf{c}'\mathbf{w} + \mathbf{w}'\mathbf{\Sigma}\mathbf{w},$$

subject to the set of constraints $\mathbf{X}'\mathbf{w} = \mathbf{x}$. Find explicit solutions for $\mathbf{w}$ and $\boldsymbol{\lambda}$, where $\mathbf{w} = (w_1, w_2, \ldots, w_n)'$ and $\boldsymbol{\lambda} = (\lambda_0, \lambda_1, \ldots, \lambda_k)'$ is the vector of the Lagrange multipliers. Show that $\hat{Z}(s_0) = \mathbf{x_0}'\hat{\boldsymbol{\beta}}_{gls} + \mathbf{c}'\mathbf{\Sigma^{-1}}\left(\mathbf{Z} - \mathbf{X}\hat{\boldsymbol{\beta}}_{gls}\right)$.

Exercise 5
Show that using the simple kriging weights and the generalized least squares estimate of $\boldsymbol{\beta}$ we obtain the universal kriging weights. The generalized least squares estimate of $\boldsymbol{\beta}$ is given by $\hat{\boldsymbol{\beta}}_{gls} = (\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{Z}$.