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

Homework 8

Exercise 1
Load the `data(parana)` data (you need to load `geoR` first) and check the details of the data set using `help(parana)`. Answer the following questions:

   a. Fit a model variogram of your choice to the sample variogram.

   b. Use cross validation to choose between ordinary and universal kriging.

   c. Make universal kriging predictions at a dense grid of your choice. Construct a raster map of the predicted values and their standard errors and add contours. On the raster map of the predicted values also add the observed data points as a bubble plot.

Exercise 2
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. Also add the observed data points as a bubble plot.

   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.

   h. Suppose that another possibility is to fit the exponential semivariogram. How would you choose between the spherical (used above) and the exponential semivariogram models?

Exercise 3
Develop the universal kriging system in terms of semivariogram. You solution should involve the minimization procedure that gives the universal kriging system in order to obtain the vector of the weights and the Lagrange multipliers.

Exercise 4
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=T)
```

   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).