University of California, Los Angeles
Department of Statistics

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

Homework 5

EXERCISE 1
Access the soil data:

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

Answer the following questions:

a. Convert `a` into a `geodata` object using `geoR`.
b. Compute and plot the empirical semivariogram up to a maximum distance of 1800.
c. Fit by eye a spherical theoretical semivariogram and add the plot on the empirical variogram.
d. Fit the spherical semivariogram using the default weights and add it on the plot of (c).
e. Fit the spherical semivariogram using Cressie's weights and add it on the plot of (c).
f. Fit the spherical semivariogram using equal weights and add it on the plot of (c).
g. Fit the spherical semivariogram using MML and add it on the plot of (c).

EXERCISE 2
Consider the spatial locations for the ozone monitoring stations in California. You can access them here:

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

Answer the following questions:

a. Using the Cholesky decomposition method and assuming multivariate normal data with mean $\mu = 0.10$ and variance covariance matrix $\mathbf{\Sigma}$, simulate manually the following isotropic process: Exponential model with $c_0 = 0, c_1 = 0.05, \alpha = 2.8$.

b. Compute the sample semivariogrm and fit the exponential model to it. On the same graph plot the theoretical model used in (a).

EXERCISE 3
Access the coal ash data:

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

Answer the following questions:

a. Convert `a` into a `geodata` object using `geoR` and compute the maximum distance for the data.

b. Compute and plot the semivariogram up to distance 10 in the east-west direction. Use both the classical and the robust estimators.

c. Fit a model semivariogram to the empirical semivariograms.

d. Compute the semivariogram cloud for both estimators for direction east-west and then construct the box plot for each cloud.

EXERCISE 4
Access the Maas river data in R here:

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

Answer the following questions:

a. Use `gstat` to compute the and plot the sample semivariogram of `log(lead)`.

b. Fit the spherical semivariogram model to the sample semivariogram of part (a) using Cressie's weights and OLS (fit.method=2 or 6).

EXERCISE 5
Access the Wolfcamp data in R here:

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

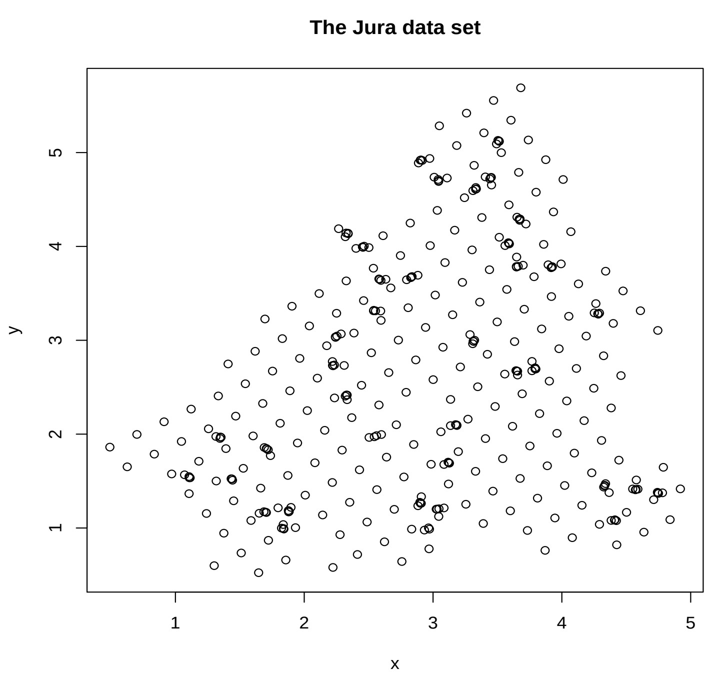
Answer the following questions:

a. Use `gstat` to compute and plot the sample semivariogram of `level)`. Assume constant mean (use formula = level $\sim$ 1). What do you observe.

b. Remove the trend by using: formula = level $\sim x + y$. Compute and plot the sample semivariogram using the de-trended data.

c. Fit the spherical semivariogram model to the sample semivariogram of part (b) using fit.method=1,2,6,7).

EXERCISE 6
These Jura data were collected by the Swiss Federal Institute of Technology at Lausanne. See Goovaerts, P. 1997, "Geostatistics for Natural Resources Evaluation", Oxford University Press, New-York, 483 p. for more details. Data were recorded at 359 locations scattered in space (see figure below).



Concentrations of seven heavy metals (cadmium, cobalt, chromium, copper, nickel, lead, and zinc) in the topsoil were measured at each location. The type of land use and rock type was also recorded for each location. The data can be accessed here:

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

```
> names(a)
 [1] "x"       "y"       "Landuse" "Rock"    "Cd"
 [6] "Co"      "Cr"      "Cu"      "Ni"      "Pb"
[11] "Zn"
```

The variables x, y are the coordinates. Landuse and Rock represent type of land use (forest, pasture, meadow, tillage) and rock type (Argovian, Kimmeridgian, Sequanina, Portlandian, and Quaternary). The other variables are concerntrations in ppm of the following chemical elements:
`Cd`: Cadmium
`Co`: Cobalt
`Cr`: Chromium
`Cu`: Copper
`Ni`: Nickel
`Pb`: Lead
`Zn`: Zinc

Compute and plot the semivariogram for each of the variables `Cd`, `Co`, `Cr`, `Cu`, `Ni`, `Pb`, `Zn`, and fit a model semivariogram to them.