Applied Spatial Statistics: Homework 2

The goal of this homework is to find some variorums for spatial data and also to explore the sampling properties of the variogram estimator. Keep in mind that the R code to reproduce all the figures from lecture 2 are in the commented source code file `SpStatsLecture2.R`. Please keep your answers concise and email your assigment to me (`nychka@ucar.edu`).

But to get started be sure to load `fields` and also the CO dat set.

```
library(fields)
data(COmonthlyMet)
```

Introduction

Following the example in the lectures simulate a random field with an exponential covariance. e. g.

```
  M <- 25
  N <- 25
  nobs<- M*N
  grid.list<- list( x= seq( 0,10,,M), y=seq(0,10,,N) )
  x.grid<- make.surface.grid( grid.list)
  theta<- 1.0
  Sigma<- exp( - rdist( x.grid, x.grid)/theta)
  B<- chol( Sigma)
  U<- rnorm(M*N)
  Z<- t(B)%*% U
```

Simulate 100 different fields and use the `vgram` function to find the means when the variogram is binned by distance into 25 bins. e.g.

```
 #set the random seed so rnorm gives the same "random" numbers.
  set.seed(243)
  Nbins<- 25
  brk<- seq( 0, 10,,Nbins+1)
  vgmeans<- matrix( NA, ncol=100, nrow=Nbins)
# simulate 100 spatial fields and find the variorums
  for( j in 1:100){
     cat( j, " ") # I call these idiot numbers
     U<- rnorm(M*N)
     Z<- t(B)%*% U
     hold<- vgram( x.grid, Z, breaks=brk)
     vgmeans[,j]<- hold$stats[2,]
   }
```

```
  centers<- hold$centers
# spaghetti plot
  matplot( centers, vgmeans, type="l", lty=1, col=1)
  lines( centers,  1- exp( -centers/theta), col="red", lwd=3)
```

Problems

1. Loop over the rows of `vgmeans` *(or use* `apply`*) to find the 5 and 95 quartiles of these estimates. Plot to get an envelope that contains the middle 95% of the curves for each distance bin.*

2. Redo your analysis from 1. using a choice of randomly distributed locations

```
  set.seed(445) # so we all get the same x.grid
  x.grid<- cbind( runif(nobs)*10, runif(nobs)*10)
  theta<- 1.0
  Sigma<- exp( - rdist( x.grid, x.grid)/theta)
  B<- chol( Sigma)
```

Now simulate `Z` as above.

3. Redo the analysis from 2. using a larger range parameter (e.g. `theta = 4.0`*). Is the accuracy of the variogram better or worse.*

4. Fit a variogram estimator using the `nls` *function to the* `CO.tmin.MAM.climate` (Including a possible nugget variance in your fit.)

5. Extra Credit Using the case from problem 1., fit an exponential function to the variogram bin means using `nls` and make a scatter plot of the fitted parameters `rho` and `theta` for the 100 simulated cases. Use the true values as starting points – how many times does the nonlinear fitting function fail?