University of California, Los Angeles
Department of Statistics

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

Ordinary kriging

Kriging (Matheron 1963) owes its name to D. G. Krige a South African mining engineer and it was first applied in mining data. Kriging assumes a random field expressed through a variogram or covariance function. It is a very popular method to solve the spatial prediction problem. Let $\mathbf{Z} = (Z(s_1), Z(s_2), \ldots, Z(s_n))'$ be the vector of the observed data at known spatial locations $s_1, s_2, \ldots, s_n$. The objective is to predict the unobserved value $Z(s_0)$ at location s_0.

The model:

The model assumption is:

$$Z(s) = \mu + \delta(s)$$

where $\delta(s)$ is a zero mean stochastic term with variogram $2\gamma(\cdot)$. The variogram was discussed in previous handouts in detail.

The Kriging System

The predictor assumption is

$$\hat{Z}(s_0) = \sum_{i=1}^{n} w_i Z(s_i)$$

i.e. it is a weighted average of the sample values, and $\sum_{i=1}^{n} w_i = 1$ to ensure unbiasedness. The w_i's are the weights that will be estimated.

Kriging minimizes the mean squared error of prediction

$$min \quad \sigma_e^2 = E[Z(s_0) - \hat{Z}(s_0)]^2$$

or

$$min \quad \sigma_e^2 = E\left[Z(s_0) - \sum_{i=1}^{n} w_i Z(s_i)\right]^2$$

For intrinsically stationary process the last equation can be written as:

$$\sigma_e^2 = 2\sum_{i=1}^{n} w_i \gamma(s_0 - s_i) \; - \; \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \gamma(s_i - s_j) \qquad (1)$$

See next page for the proof:

Let's examine $(Z(s_0) - \sum_{i=1}^n w_i Z(s_i))^2$:

$$\begin{aligned}
\left(z(s_0) - \sum_{i=1}^n w_i z(s_i)\right)^2 &= \\
z^2(s_0) - 2z(s_0)\sum_{i=1}^n w_i z(s_i) + \sum_{i=1}^n\sum_{j=1}^n w_i w_j z(s_i)z(s_j) &= \\
\sum_{i=1}^n w_i z^2(s_0) - 2\sum_{i=1}^n w_i z(s_0)z(s_i) + \sum_{i=1}^n\sum_{j=1}^n w_i w_j z(s_i)z(s_j) & \\
-\frac{1}{2}\sum_{i=1}^n w_i z^2(s_i) - \frac{1}{2}\sum_{j=1}^n w_j z^2(s_j) + \sum_{i=1}^n w_i z^2(s_i) &= \\
-\frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n w_i w_j \left[z(s_i) - z(s_j)\right]^2 + \sum_{i=1}^n w_i \left[z(s_0) - z(s_i)\right]^2 &
\end{aligned}$$

If we take expectations on the last expression we have

$$-\frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n w_i w_j E\left[z(s_i) - z(s_j)\right]^2 + \sum_{i=1}^n w_i E\left[z(s_0) - z(s_i)\right]^2 =$$

$$-\frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n w_i w_j \text{var}\left[z(s_i) - z(s_j)\right] + \sum_{i=1}^n w_i \text{var}\left[z(s_0) - z(s_i)\right]$$

But $\text{var}\left[z(s_i) - z(s_j)\right] = 2\gamma(\cdot)$ is the definition of the variogram, and therefore the previous expression is written as: $2\sum_{i=1}^n w_i\gamma(s_0 - s_i) \ - \ \sum_{i=1}^n\sum_{j=1}^n w_i w_j \gamma(s_i - s_j)$
Therefore kriging minimizes

$$\begin{aligned}
\sigma_e^2 = E[(Z(s_0) - \sum_{i=1}^n w_i Z(s_i)]^2 &= \\
2\sum_{i=1}^n w_i\gamma(s_0 - s_i) \ - \ \sum_{i=1}^n\sum_{j=1}^n w_i w_j \gamma(s_i - s_j) &
\end{aligned}$$

subject to

$$\sum_{i=1}^n w_i = 1$$

The minimization is carried out over $(w_1, w_2, ..., w_n)$, subject to the constraint $\sum_{i=1}^n w_i = 1$. Therefore the minimization problem can be written as:

$$\min \ \ 2\sum_{i=1}^n w_i\gamma(s_0 - s_i) \ - \ \sum_{i=1}^n\sum_{j=1}^n w_i w_j \gamma(s_i - s_j) \ - \ 2\lambda(\sum_{i=1}^n w_i - 1) \tag{2}$$

where λ is the Lagrange multiplier. After differentiating (2) with respect to $w_1, w_2, ..., w_n$, and λ and set the derivatives equal to zero we find that

$$-\sum_{j=1}^n w_j\gamma(s_i - s_j) + \gamma(s_0 - s_i) - \lambda = 0, \quad i = 1, ..., n$$

and

$$\sum_{i=1}^n w_i = 1$$

Using matrix notation the previous system of equations can be written as

$$\mathbf{\Gamma W} = \boldsymbol{\gamma}$$

Therefore the weights $w_1, w_2, ..., w_n$ and the Lagrange multiplier λ can be obtained by

$$\mathbf{W} = \mathbf{\Gamma^{-1}}\boldsymbol{\gamma}$$

where

$$\mathbf{W} = (w_1, w_2, ..., w_n, \lambda)$$

$$\boldsymbol{\gamma} = (\gamma(s_0 - s_1), \gamma(s_0 - s_2), ..., \gamma(s_0 - s_n), 1)'$$

$$\mathbf{\Gamma} = \begin{cases} \gamma(s_i - s_j), & i = 1, 2, ..., n, \quad j = 1, 2, ..., n, \\ 1, & i = n+1, \quad j = 1, ..., n, \\ 1, & j = n+1, \quad i = 1, ..., n, \\ 0, & i = n+1, \quad j = n+1. \end{cases}$$

The variance of the error of prediction:
So far, we found the weights and therefore we can compute the estimator: $\hat{Z}(s_0) = \sum_{i=1}^n w_i Z(s_i)$. How about the variance of the error of prediction, namely σ_e^2?

We multiply

$$-\sum_{j=1}^n w_j \gamma(s_i - s_j) + \gamma(s_0 - s_i) - \lambda = 0$$

by w_i and we sum over all $i = 1, \cdots, n$ to get:

$$-\sum_{i=1}^n w_i \sum_{j=1}^n w_j \gamma(s_i - s_j) + \sum_{i=1}^n w_i \gamma(s_0 - s_i) - \sum_{i=1}^n w_i \lambda = 0$$

Or

$$-\sum_{i=1}^n \sum_{j=1}^n w_i w_j \gamma(s_i - s_j) + \sum_{i=1}^n w_i \gamma(s_0 - s_i) - \sum_{i=1}^n w_i \lambda = 0$$

Therefore,

$$\sum_{i=1}^n \sum_{j=1}^n w_i w_j \gamma(s_i - s_j) = \sum_{i=1}^n w_i \gamma(s_0 - s_i) - \sum_{i=1}^n w_i \lambda$$

If we substitute this result into equation (1) we finally get:

$$\sigma_e^2 = \sum_{i=1}^n w_i \gamma(s_i - s_0) + \lambda \tag{3}$$

The Kriging System

$$\begin{pmatrix} \gamma(s_1-s_1) & \gamma(s_1-s_2) & \gamma(s_1-s_3) & \cdots & \gamma(s_1-s_n) & 1 \\ \gamma(s_2-s_1) & \gamma(s_2-s_2) & \gamma(s_2-s_3) & \cdots & \gamma(s_2-s_n) & 1 \\ \cdots & \cdots & \ddots & \cdots & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \cdots & 1 \\ \gamma(s_n-s_1) & \gamma(s_n-s_2) & \gamma(s_n-s_3) & \cdots & \gamma(s_n-s_n) & 1 \\ 1 & 1 & \cdots & \cdots & 1 & 0 \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ \vdots \\ w_n \\ \lambda \end{pmatrix} = \begin{pmatrix} \gamma(s_0-s_1) \\ \gamma(s_0-s_2) \\ \vdots \\ \vdots \\ \gamma(s_0-s_n) \\ 1 \end{pmatrix}$$

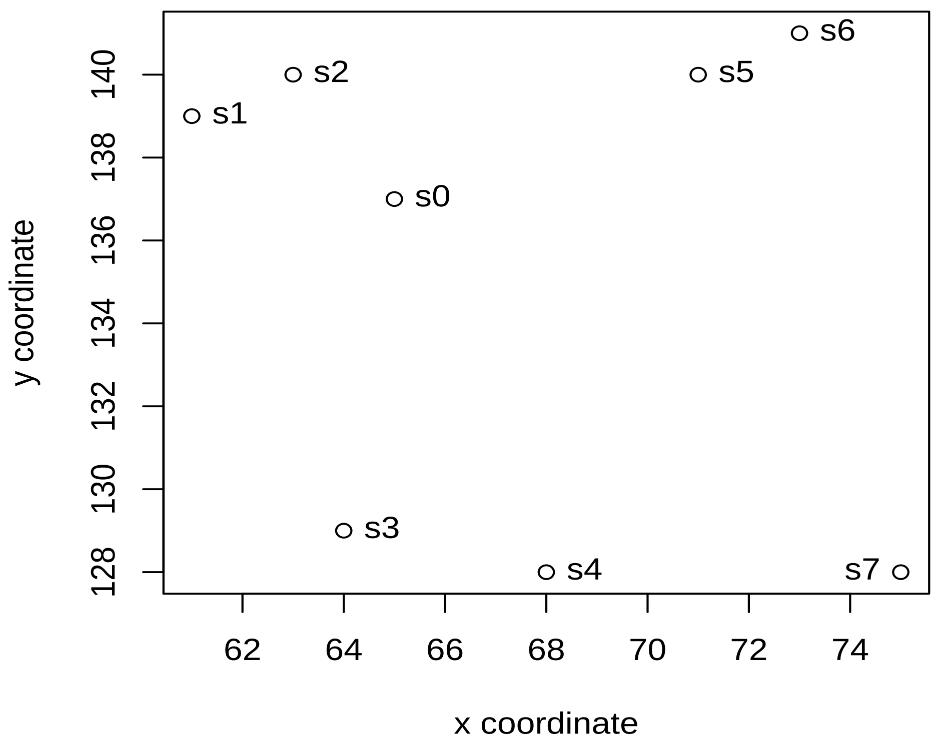
We understand now why the sample variogram cannot be used directly in the kriging system, and instead a theoretical variogram is used. First, the $\boldsymbol{\gamma}$ vector may call for variogram values for distances that are not available from sample data. There are situations where the distance from the point being estimated to a particular sample is smaller than the distance between pairs of available samples. Since the sample data set cannot provide any pairs for these small distances, we must rely on a function that provides variogram values for all distances. Second, the use of the sample variogram does not guarantee the existence and uniqueness of the solution to the kriging system. In other words the sample variogram version of $\mathbf{\Gamma}$ might not be positive definite. To be guaranteed of having one and only one solution, we must ensure that our system has the property of positive definiteness. This is ensured by the choice of one of the model variograms mentioned earlier.

A simple example:

Consider the following data

s_i	x	y	$z(s_i)$
s_1	61	139	477
s_2	63	140	696
s_3	64	129	227
s_4	68	128	646
s_5	71	140	606
s_6	73	141	791
s_7	75	128	783
$\mathbf{s_0}$	**65**	**137**	**???**

Our goal is to predict the unknown value at location s_0. Here is the $x-y$ plot:



For these data, let's assume that we use the exponential semivariogram model with parameters $c_0 = 0, c_1 = 10, \alpha = 3.33$.

$$\gamma(h) = 10(1 - e^{-\frac{h}{3.33}}).$$

We need to construct the matrix $\mathbf{\Gamma}$ and the vector $\boldsymbol{\gamma}$. First we calculate the distance matrix as shown below:

$$\text{Distance matrix} = \begin{array}{c} \\ s_0 \\ s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \\ s_6 \\ s_7 \end{array} \begin{array}{cccccccc} s_0 & s_1 & s_2 & s_3 & s_4 & s_5 & s_6 & s_7 \end{array} \hphantom{)}$$

$$\begin{pmatrix} 0.00 & 4.47 & 3.61 & 8.06 & 9.49 & 6.71 & 8.94 & 13.45 \\ 4.47 & 0.00 & 2.24 & 10.44 & 13.04 & 10.05 & 12.17 & 17.80 \\ 3.61 & 2.24 & 0.00 & 11.05 & 13.00 & 8.00 & 10.05 & 16.97 \\ 8.06 & 10.44 & 11.05 & 0.00 & 4.12 & 13.04 & 15.00 & 11.05 \\ 9.49 & 13.04 & 13.00 & 4.12 & 0.00 & 12.37 & 13.93 & 7.00 \\ 6.71 & 10.05 & 8.00 & 13.04 & 12.37 & 0.00 & 2.24 & 12.65 \\ 8.94 & 12.17 & 10.05 & 15.00 & 13.93 & 2.24 & 0.00 & 13.15 \\ 13.45 & 17.80 & 16.90 & 11.05 & 7.00 & 2.65 & 13.15 & 0.00 \end{pmatrix}$$

The ij_{th} entry in the matrix above was computed as follows:

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$$

Now, we can find the entries of the matrix $\mathbf{\Gamma}$ and the vector $\boldsymbol{\gamma}$. Each entry will be computed using the exponential semivariogram $\gamma(h) = 10(1 - e^{-\frac{h}{3.33}})$. Here they are:

$$\mathbf{\Gamma} = \begin{pmatrix} 0 & 4.893 & 9.564 & 9.800 & 9.510 & 9.740 & 9.952 & 1 \\ 4.893 & 0 & 9.637 & 9.798 & 9.093 & 9.510 & 9.938 & 1 \\ 9.564 & 9.637 & 0 & 7.095 & 9.800 & 9.889 & 9.637 & 1 \\ 9.800 & 9.798 & 7.095 & 0 & 9.755 & 9.847 & 8.775 & 1 \\ 9.510 & 9.093 & 9.800 & 9.755 & 0 & 4.893 & 9.775 & 1 \\ 9.740 & 9.510 & 9.889 & 9.847 & 4.893 & 0 & 9.806 & 1 \\ 9.952 & 9.938 & 9.637 & 8.775 & 9.775 & 9.806 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \end{pmatrix}$$

$$\boldsymbol{\gamma} = \begin{pmatrix} 7.384 \\ 6.614 \\ 9.109 \\ 9.420 \\ 8.664 \\ 9.316 \\ 9.823 \\ 1 \end{pmatrix}$$

The weights and the Lagrnange multiplier can be obtained as follows:

$$\boldsymbol{W} = \boldsymbol{\Gamma^{-1}\gamma} = \begin{pmatrix} 0 & 4.893 & 9.564 & 9.800 & 9.510 & 9.740 & 9.952 & 1 \\ 4.893 & 0 & 9.637 & 9.798 & 9.093 & 9.510 & 9.938 & 1 \\ 9.564 & 9.637 & 0 & 7.095 & 9.800 & 9.889 & 9.637 & 1 \\ 9.800 & 9.798 & 7.095 & 0 & 9.755 & 9.847 & 8.775 & 1 \\ 9.510 & 9.093 & 9.800 & 9.755 & 0 & 4.893 & 9.775 & 1 \\ 9.740 & 9.510 & 9.889 & 9.847 & 4.893 & 0 & 9.806 & 1 \\ 9.952 & 9.938 & 9.637 & 8.775 & 9.775 & 9.806 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \end{pmatrix}^{-1} \begin{pmatrix} 7.384 \\ 6.614 \\ 9.109 \\ 9.420 \\ 8.664 \\ 9.316 \\ 9.823 \\ 1 \end{pmatrix}.$$

The answer is:

$$\boldsymbol{W} = \begin{pmatrix} 0.174 \\ 0.317 \\ 0.129 \\ 0.086 \\ 0.151 \\ 0.057 \\ 0.086 \\ 0.906 \end{pmatrix}.$$

The last element of the **W** vector is the Lagrange multiplier, $\lambda = 0.906$. We can verify that the sum of the elements 1 through 7 is equal to 1, as it should be.

The predicted value at location s_0 is equal to:

$$\hat{z}(s_0) = \sum_{i=1}^{n} w_i z(s_i) = 0.174(477) + \cdots + 0.086(783) = 592.59.$$

And the variance:

$$\sigma_e^2 = \sum_{i=1}^{n} w_i \gamma(s_i - s_0) + \lambda = 0.174(7.384) + \cdots + 0.086(9.823) + 0.906 = 8.96.$$

Under the assumption that $Z(s)$ is Gaussian a 95% confidence interval can be computed as follows:

$$592.59 \pm 1.96\sqrt{8.96}$$

Or

$$577.09 \leq Z(s_0) \leq 588.83$$

Short code for ordinary kriging in terms of variogram:

```
a <- read.table("kriging_1.txt", header=TRUE)
b <- read.table("kriging_11.txt", header=TRUE)

x <- as.matrix(cbind(a$x, a$y))

x1 <- rep(rep(0,8),8)            #Initialize
dist <- matrix(x1,nrow=8,ncol=8)  #the distance matrix

for (i in 1:8){
for (j in 1:8){
dist[i,j]=((x[i,1]-x[j,1])^2+(x[i,2]-x[j,2])^2)^.5
}
}

c0 <- 0
c1 <- 10
alpha <- 3.33

x1 <- rep(rep(0,8),8)             #Initialize
G <- matrix(x1,nrow=8,ncol=8)     #the GAMMA matrix

for(i in 1:8){
for (j in 1:8){
G[i,j]=c1*(1-exp(-dist[i,j]/alpha))
                if(i==j){G[i,j]=0}
if(i==8){G[i,j]=1}
if(j==8){G[i,j]=1}
if(i==8 & j==8) {G[i,j]=0}
}
}

g <- rep(0,8)                    #Initialize
                                  #the g vector

for(j in 1:8){
g[j]=c1*(1-exp(-dist[8,j]/alpha))
if(j == 8) {g[j]=1}
}

w <- solve(G) %*% g          #Obtain the weights and the Lagrange parameter

z_hat <- w[-8] %*% b$z         #Compute the estimate
var_z_hat <- t(w) %*% g  #Compute the variance of the estimate
```