University of California, Los Angeles
Department of Statistics

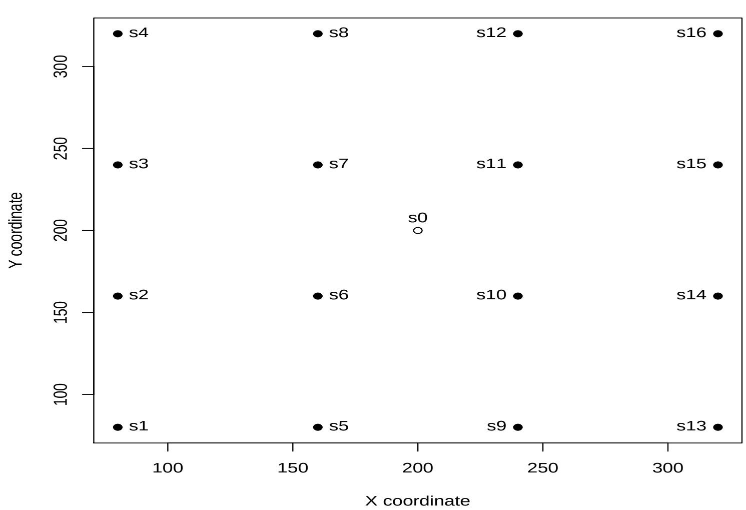
Statistics C173/C273 **Instructor: Nicolas Christou**

Effect of variogram parameters on kriging weights

We will explore in this document how the kriging weights are affected by the variogram parameters. We will use the following data which are measurements of soil pH.

```
> b
     x   y  pH
1   80  80 7.0
2   80 160 6.9
3   80 240 7.9
4   80 320 8.0
5  160  80 6.0
6  160 160 6.2
7  160 240 8.0
8  160 320 8.0
9  240  80 5.8
10 240 160 6.2
11 240 240 7.8
12 240 320 8.0
13 320  80 6.0
14 320 160 6.2
15 320 240 7.8
16 320 320 8.0
```

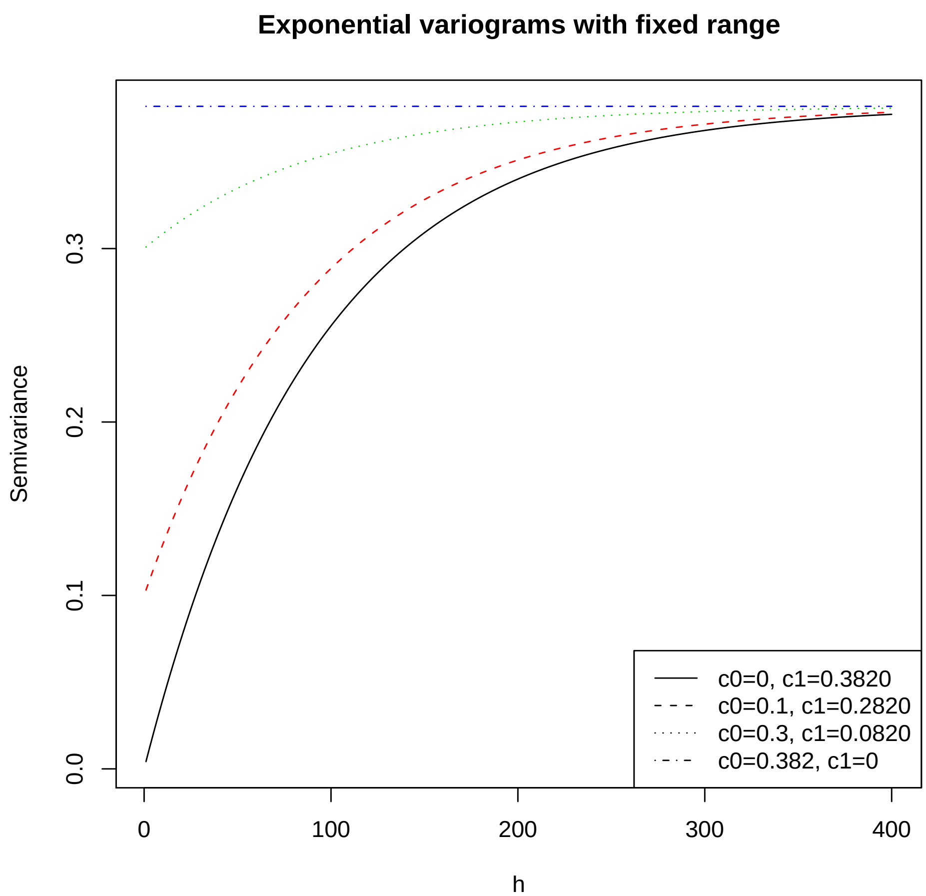
And we will predict the point at location (`x=200, y=200`) using different variogram parameters. Here is the plot of the data.



A. We use the exponential variogram and we keep the range constant at $\alpha = 90.53$ and vary the values of the nugget and partial sill as follows.

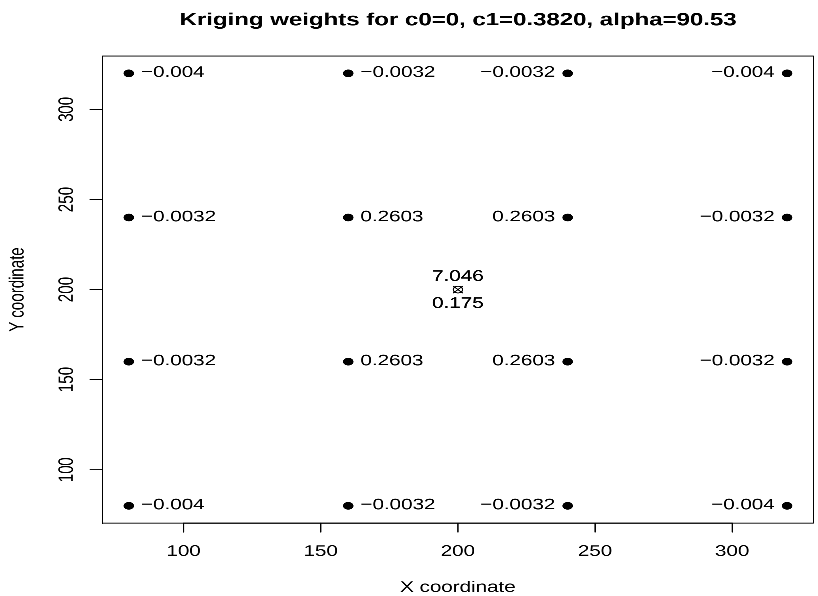
c_0	c_1	α
0.0000	0.3820	90.53
0.1000	0.2820	90.53
0.3000	0.0820	90.53
0.3820	0.0000	90.53

The corresponding model variograms are shown below:

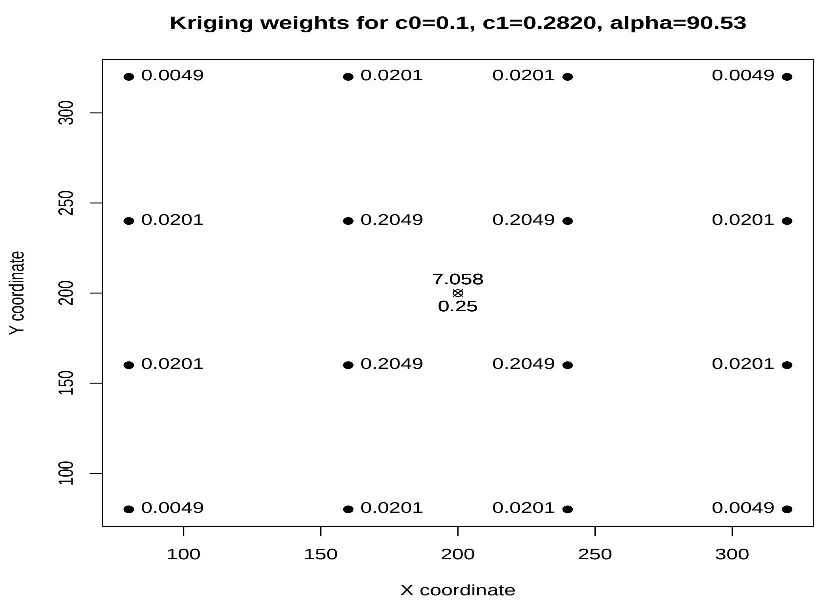


The last model is the so called pure nugget model where no spatial correlation exists. There is no difference in terms of "closeness" between data that are near the point to be estimated or data that are far from the point. The next four plots show the kriging weights estimated using each one of the models above. We can also see the predicted value of the point s_0 and its variance (value below the point).

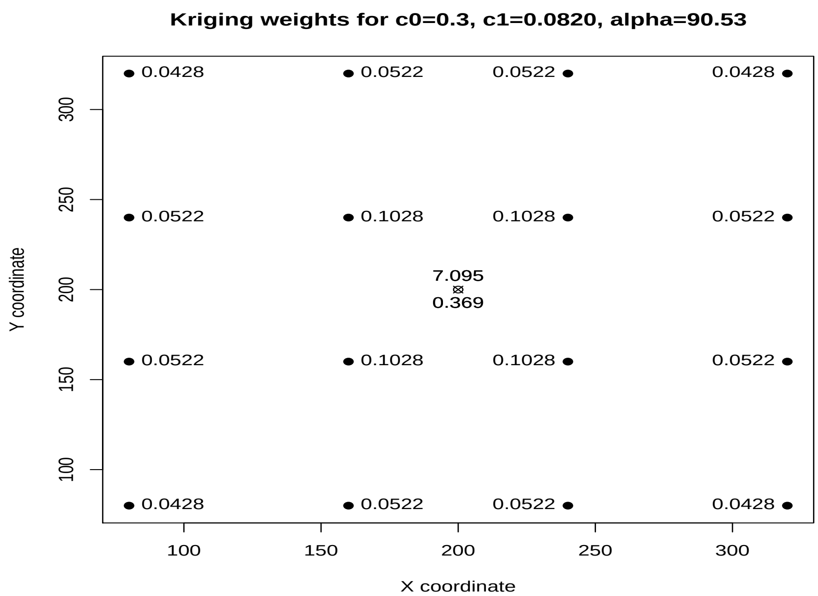
Case 1: $c_0 = 0, c_1 = 0.3820, \alpha = 90.53.$



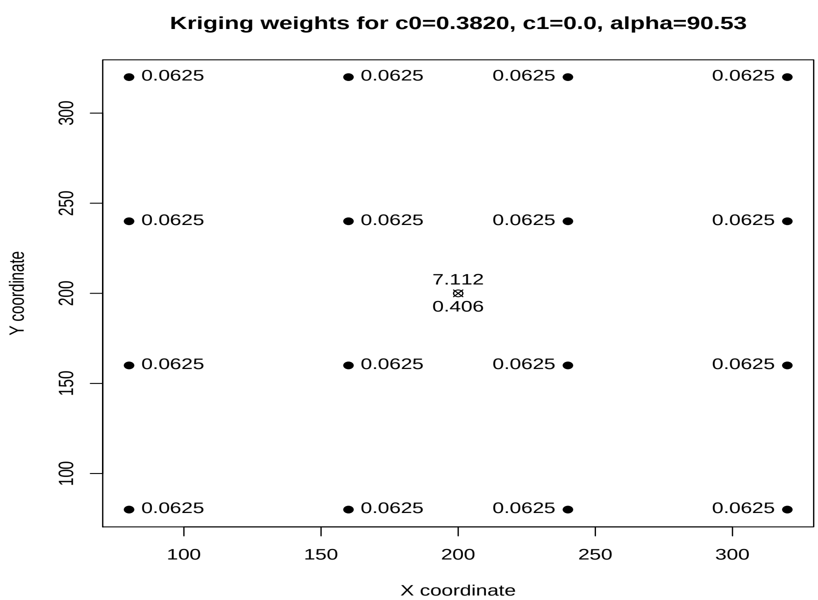
Case 2: $c_0 = 0.1, c_1 = 0.2820, \alpha = 90.53.$



Case 3: $c_0 = 0.3, c_1 = 0.0820, \alpha = 90.53.$



Case 4: $c_0 = 0.3820, c_1 = 0, \alpha = 90.53.$

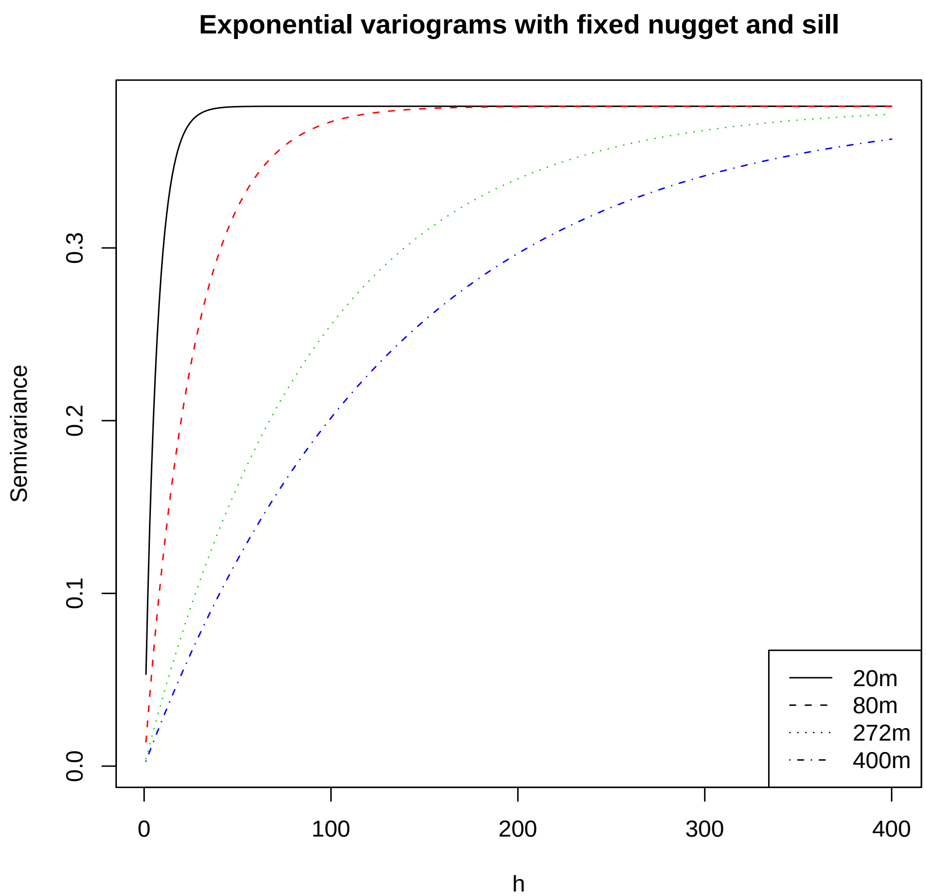


What do you observe?

B. We use the exponential variogram and we keep the nugget and partial sill constant and vary the range as follows.

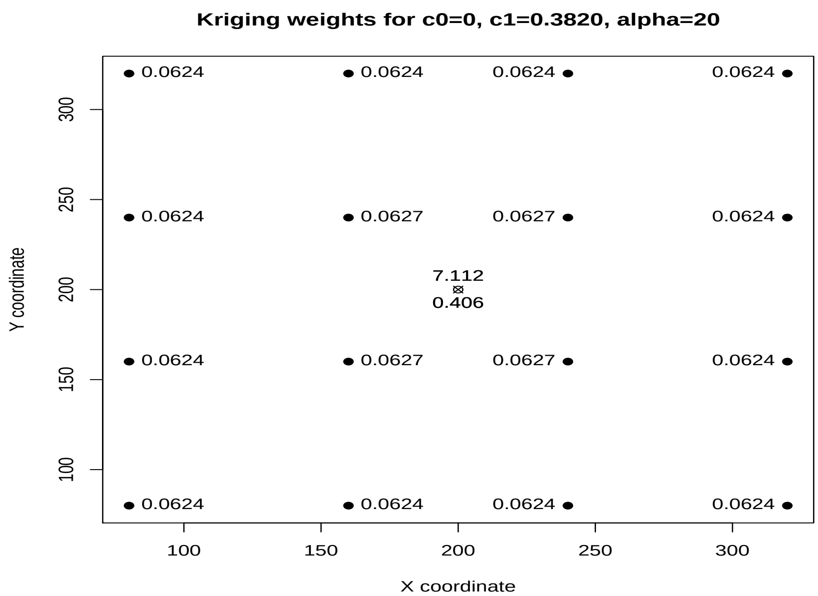
c_0	c_1	α
0.0000	0.3820	20.0
0.0000	0.3820	80.0
0.0000	0.3820	272.0
0.0000	0.3820	400.0

The corresponding model variograms are shown below:

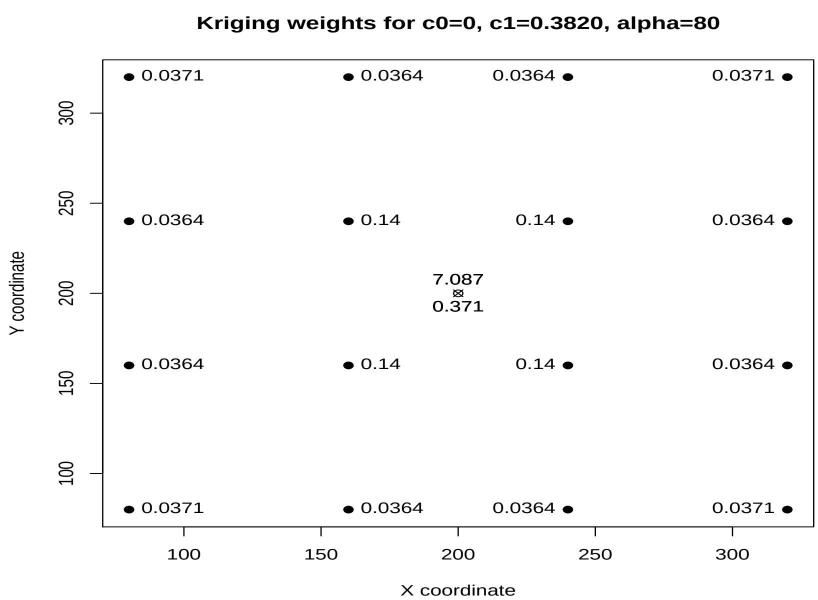


The next four plots show the kriging weights estimated using each one of the models above. We can also see the predicted value of the point s_0 and its variance (value below the point).

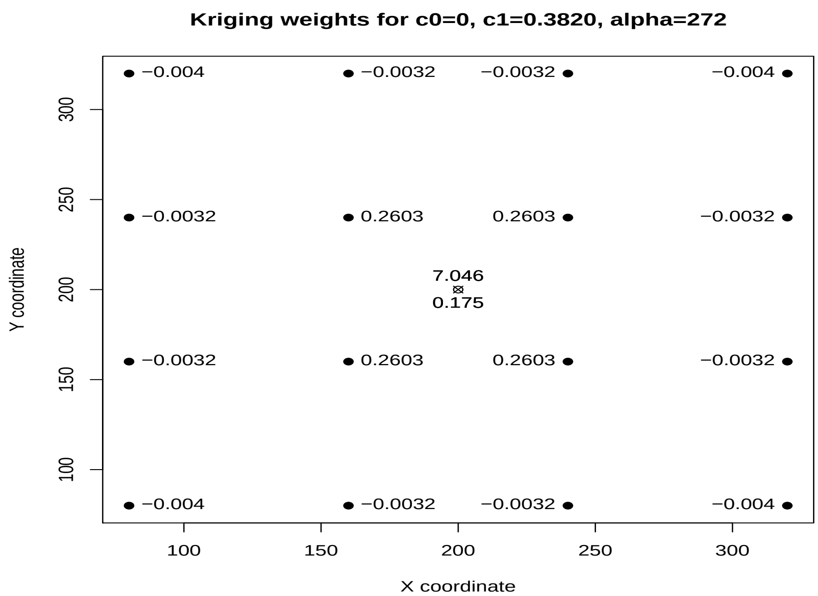
Case 1: $c_0 = 0, c_1 = 0.3820, \alpha = 20.0.$



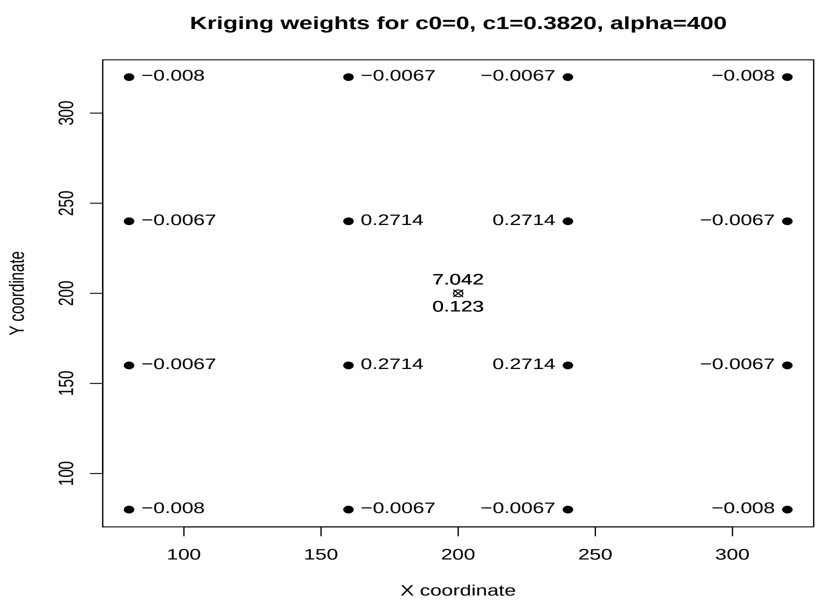
Case 2: $c_0 = 0, c_1 = 0.3820, \alpha = 80.0.$



Case 3: $c_0 = 0, c_1 = 0.3820, \alpha = 272.0.$



Case 4: $c_0 = 0, c_1 = 0.3820, \alpha = 400.0.$



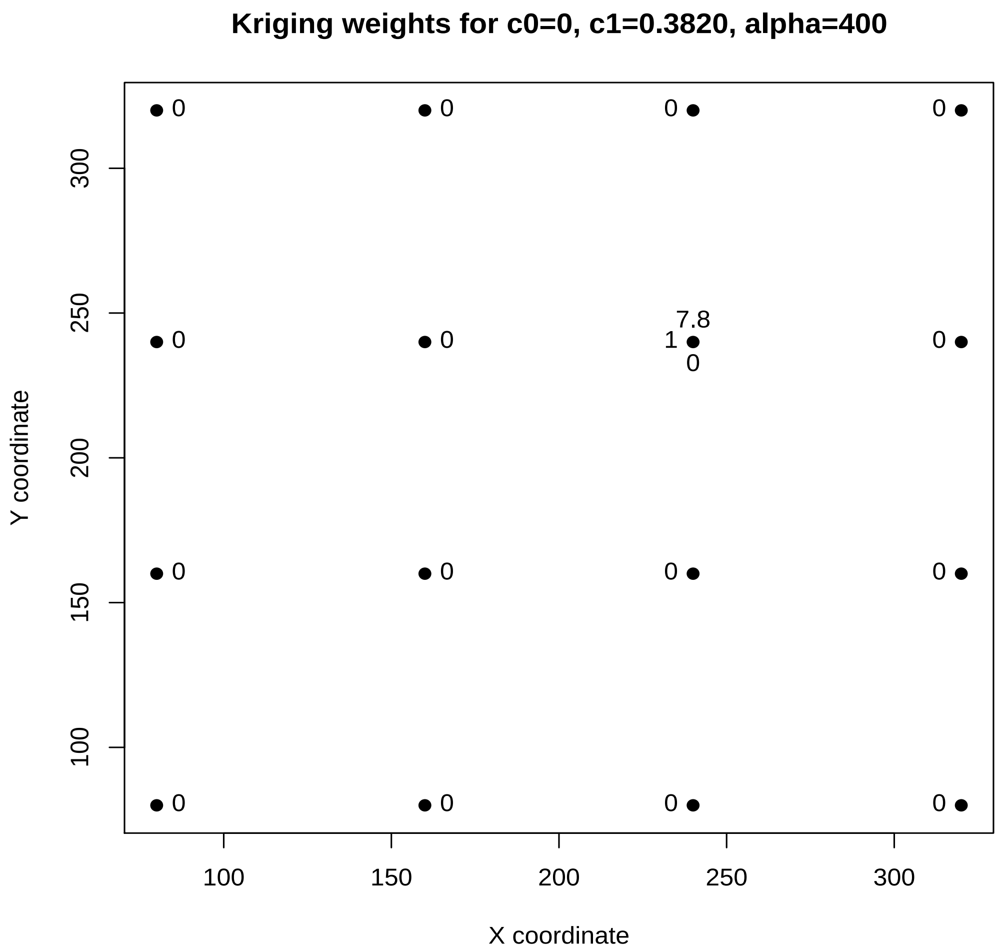
What do you observe?

The previous calculations were done with the following short `R` code:

```
#Read the data.  Note:  The first file contains only the coordinates
#of the 16 data points plus the point to be estimated.  The second
#file contains the coordinates and the data values of the 16 points.
a <- read.table("kr_effect_coord.txt", header=TRUE)
b <- read.table("kr_effect_data.txt", header=TRUE)
#Or from the course website:
a <- read.table("http://www.stat.ucla.edu/~nchristo/statistics_c173_c273/
     kr_effect_coord.txt", header=TRUE)
b <- read.table("http://www.stat.ucla.edu/~nchristo/statistics_c173_c273/
     kr_effect_data.txt", header=TRUE)

x <- as.matrix(cbind(a$x, a$y))
#Initialize the distance matrix:
x1 <- rep(rep(0,17),17)
dist <- matrix(x1,nrow=17,ncol=17)

#Compute the distance matrix:
for (i in 1:17){
for (j in 1:17){
dist[i,j]=((x[i,1]-x[j,1])^2+(x[i,2]-x[j,2])^2)^.5
}
}

#Variogram parameters:
c0 <- 0
c1 <- 0.3820
alpha <- 6.67

#Initialize the C1 matrix:
x1 <- rep(rep(0,17),17)
C1 <- matrix(x1,nrow=17,ncol=17)

#Compute the C1 matrix:
for(i in 1:17){
for (j in 1:17){
C1[i,j]=c1*exp(-dist[i,j]/alpha)
                if(i==j){C1[i,j]=c0+c1}
if(i==17){C1[i,j]=1}
if(j==17){C1[i,j]=1}
if(i==17 & j==17) {C1[i,j]=0}
}
}

#Initialize the D1 vector:
D1 <- rep(0,17)
```

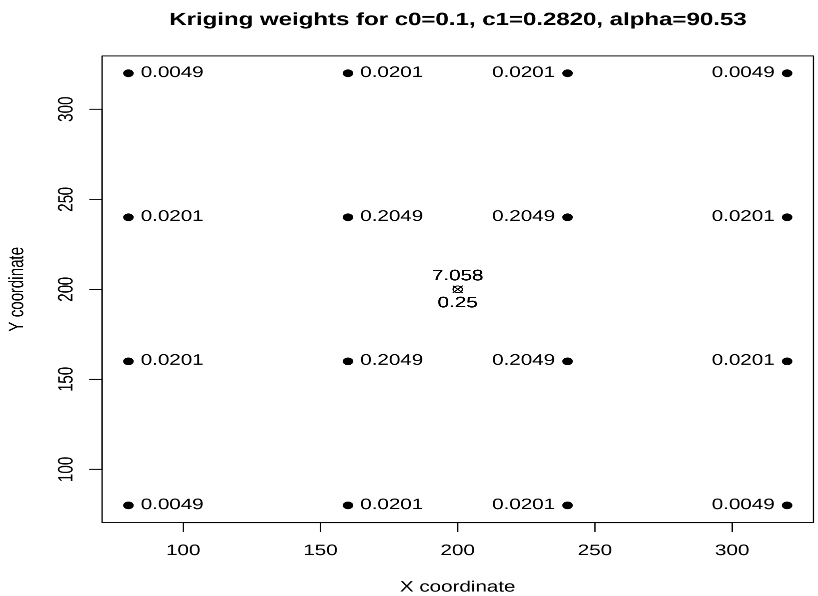
```
#Compute the D1 vector:
for(j in 1:17){
D1[j]=c1*exp(-dist[17,j]/alpha)
if(j == 17) {D1[j]=1}
}

#Obtain the weights and the Lagrange parameter:
w <- solve(C1) %*% D1

#Obtain the weights and the Lagrange parameter (keep 4 decimal points):
w1 <- round(solve(C1) %*% D1, digits=4)

#Compute the estimate:
z_hat <- round(w[-17] %*% b$pH, digits=3)

#Compute the variance of the estimate:
var_z_hat <- round(c0+c1 - t(w) %*% D1, digits=3)

#Save the plot into a file:
pdf("range1.pdf")

#Plot the 16 data points:
plot(b$x,b$y, xlab="X coordinate", ylab="Y coordinate",
main="Kriging weights for c0=0, c1=0.3820, alpha=20", pch=16)

#Add to the plot the point to be estimated:
points(a$x[17], a$y[17],pch=13)
#Add text to the plot:
text(200, 200, label=z_hat, pos=3)
text(200, 200, label=var_z_hat, pos=1)
text(80, 80, label=w1[1], pos=4)
text(80, 160, label=w1[2], pos=4)
text(80, 240, label=w1[3], pos=4)
text(80, 320, label=w1[4], pos=4)
text(160, 80, label=w1[5], pos=4)
text(160, 160, label=w1[6], pos=4)
text(160, 240, label=w1[7], pos=4)
text(160, 320, label=w1[8], pos=4)
text(240, 80, label=w1[9], pos=2)
text(240, 160, label=w1[10], pos=2)
text(240, 240, label=w1[11], pos=2)
text(240, 320, label=w1[12], pos=2)
text(320, 80, label=w1[13], pos=2)
text(320, 160, label=w1[14], pos=2)
text(320, 240, label=w1[15], pos=2)
text(320, 320, label=w1[16], pos=2)
dev.off()
```

Estimating one of the data points:
Kriging not only produces an unbiased estimator of the value at s_0, but also is an exact interpolator. In other words, if we attempt to estimate one of the data points, we will find exactly the observed value of that point and with variance of the estimator equal to zero. The following example will verify this statement. We will use ordinary kriging to estimate the value at location $(x = 240, y = 240)$. The value of the soil pH at this location is equal to 7.8. Here is the figure with the corresponding weights, estimated value, and its variance.



We observe that all the weights are equal to zero, except the weight for the point $(x = 240, y = 240)$ which is equal to 1. Therefore the estimated value will be equal to itself (here 7.8). Also, the variance of the estimate is zero. The kriged surface (the surface that is constructed with many points on a grid) will pass through the observed data points.

An intuitive look at ordinary kriging:
The matrix $\mathbf{\Sigma}$ and the vector $\boldsymbol{c}$ play an important role for the values of the kriging weights. The matrix $\mathbf{\Sigma}$ is always the same regardless of which point we want to estimate. But the entries of the vector $\boldsymbol{c}$ change and this depends on the location of the point that we want to estimate. For example, let's examine the earlier case where $c_0 = 0.1, c_1 = 0.2820, \alpha = 90.53$. The figure below shows the kriging weights, the estimate at location $(x = 200, y = 200)$, and its variance.



Now, let's look at the entries of the vector $\boldsymbol{c}$. These entries were computed using the exponential covariogram $C(h) = 0.2820e^{-\frac{h}{93.53}}$. For example to compute the entries $c_{0,7}$ and $c_{0,4}$ we compute first the Euclidean distance

$$d_{07} = \sqrt{(200-160)^2 + (200-240)^2} = 56.569$$

and

$$d_{04} = \sqrt{(200-80)^2 + (200-320)^2} = 169.706$$

and then

$$c_{0,7} = .2820 \times exp(-56.569/90.53) = 0.15096$$

$$c_{0,4} = .2820 \times exp(-169.706/90.53) = 0.04326$$

Similarly, the rest of the entries of the vector $\boldsymbol{c}$ are given below:

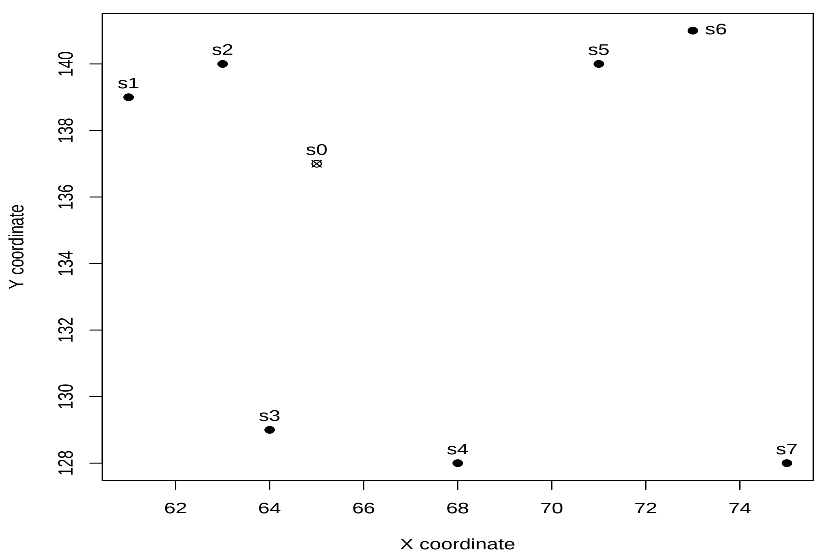
```
> c
 [1] 0.04326432 0.06973333 0.06973333 0.04326432 0.06973333 0.15096491
 [7] 0.15096491 0.06973333 0.06973333 0.15096491 0.15096491 0.06973333
[13] 0.04326432 0.06973333 0.06973333 0.04326432 1.00000000
```

The vector $\boldsymbol{c}$ therefore, can be viewed as a vector that weighs each of the observed points according to its distance to the point being estimated. This is shown from the figure above and from the entries of $\boldsymbol{c}$. Point 4 ($x = 80, y = 320$) is relatively far from the point being estimated ($x = 200, y = 200$) which means that the corresponding entry of $\boldsymbol{c}$ is small and therefore the kriging weight is also small $w_4 = 0.0049$. The opposite happens with point 7 ($x = 160, y = 240$) which is closer to the point being estimated ($x = 200, y = 200$) which means that the corresponding entry of $\boldsymbol{c}$ is large and therefore the kriging weight is large as well $w_7 = 0.2049$.

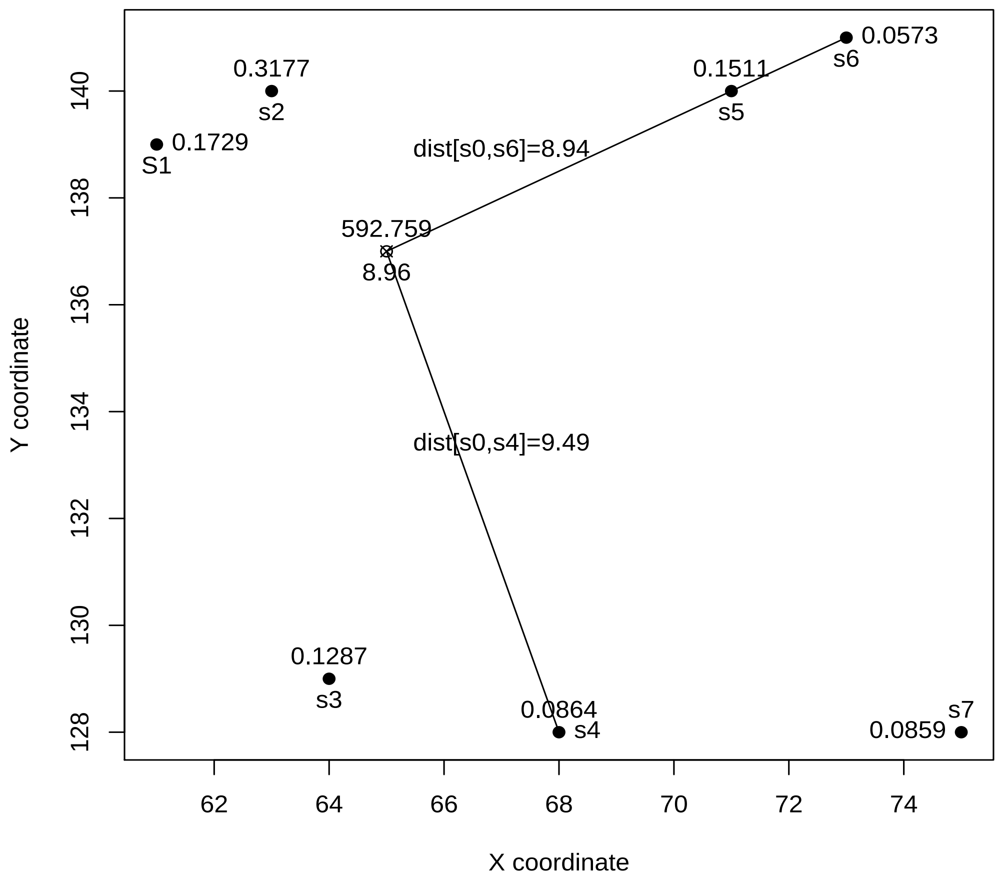
The role of the matrix $\boldsymbol{\Sigma}$ is more complex. In our example, our data points formed a regular grid and therefore we cannot see clearly the impact of $\boldsymbol{\Sigma}$. Instead we will use an example from earlier discussion. Here are the data:

```
> b
   x   y   z
1 61 139 477
2 63 140 696
3 64 129 227
4 68 128 646
5 71 140 606
6 73 141 791
7 75 128 783
```

Our goal is to estimate the value of z at location ($x = 65, y = 137$) as shown in the figure below:



On the next figure we can see the kriging weights, the estimate at s_0, and its variance.



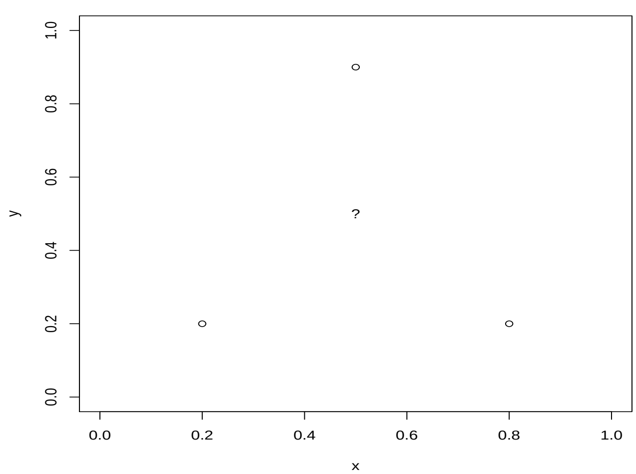
We observe that even though point s_4 is farther from point s_0 than point s_6 is, the kriging weight of point s_6 is smaller than the weight of point s_4. This is because the impact of point s_6 in estmating s_0 was reduced because of its proximity to point s_5. Therefore the matrix $\mathbf{\Sigma}$ gives information about clustering in the data.

The prediction weights

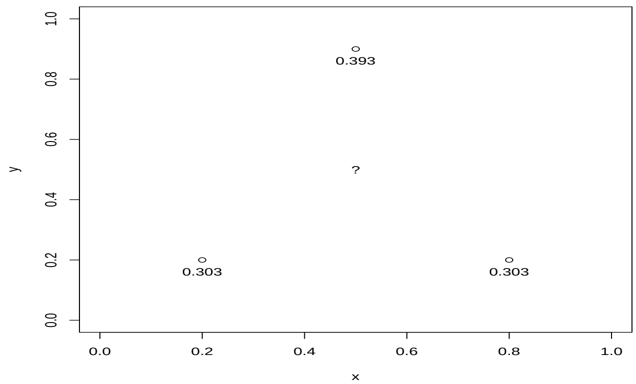
The predictions weights are shown for three sets of sample locations in the unit square as follows:

```
Set 1:
    x   y
1 0.2 0.2
2 0.5 0.9
3 0.8 0.2

Set 2:
    x   y
1 0.2 0.2
2 0.4 0.9
3 0.6 0.9
4 0.8 0.2

Set 3:
     x   y
1 0.10 0.5
2 0.35 0.5
3 0.50 0.1
4 0.80 0.8
```

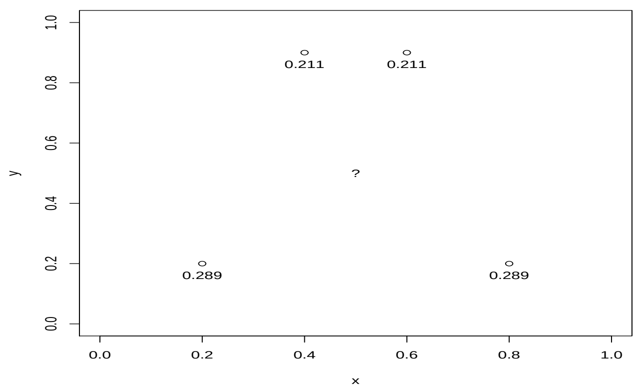
For all three examples we will use the exponential variogram with $c_0 = 0, c_1 = 10, \alpha = 3.33$, and the point to be predicted is at location $x = 0.5, y = 0.5$ (at the center of the unit square). See figure below for the first set.



The figure below shows that the uppermost of the three data points receives the largest weight because it is closer to the target point.



This figure shows a property of kriging known as de-clustering. The uppermost two points “share” the weight that the uppermost point in the figure above received.



This figure shows a property of kriging known as masking. This occurs when sample locations and the target point are collinear (or approximately so). The masked location is given here a negative weight. Depending on the assumed covariance structure the masked locations may receive positive, zero, or negative weights.