University of California, Los Angeles
Department of Statistics

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

Co-kriging

Suppose that at each spatial location $s_i, i = 1, \ldots, n$ we observe k variables as follows:

$$
\begin{array}{cccc}
Z_1(s_1) & Z_1(s_2) & \ldots & Z_1(s_n) \\
Z_2(s_1) & Z_2(s_2) & \ldots & Z_2(s_n) \\
\vdots & \ldots & \ldots & \vdots \\
\vdots & \ldots & \ldots & \vdots \\
Z_k(s_1) & Z_k(s_2) & \ldots & Z_k(s_n)
\end{array}
$$

We want to predict $Z_1(s_0)$, i.e. the value of variable Z_1 at location s_0.

This situation that the variable under consideration (the target variable) occurs with other variables (co-located variables) arises many times in practice and we want to explore the possibility of improving the prediction of variable Z_1 by taking into account the correlation of Z_1 with these other variables. For example, the prediction of lead can be improved if we also know the values of zinc at each spatial location. The value of lead will be predicted by the observed values of lead but also by the observed values of zinc.

The predictor assumption:

$$
\begin{aligned}
\hat{Z}_1(s_0) &= \sum_{j=1}^{k}\sum_{i=1}^{n} w_{ji} Z_j(s_i) \\
&= w_{11}z_1(s_1) + w_{12}z_1(s_2) + \ldots + w_{1n}z_1(s_n) \\
&+ w_{21}z_2(s_1) + w_{22}z_2(s_2) + \ldots + w_{2n}z_2(s_n) \\
&+ \quad \vdots \quad + \quad \vdots \quad + \quad + \quad \vdots \\
&+ w_{k1}z_k(s_1) + w_{k2}z_k(s_2) + \ldots + w_{kn}z_k(s_n)
\end{aligned}
$$

We see that there are weights associated with variable Z_1 but also with each one of the other variables. We will examine ordinary co-kriging, which means that $E(Z_j(s_i)) = \mu_j$ for all j and i. In vector form:

$$
E(\mathbf{Z}(\mathbf{s})) = \begin{pmatrix} E(Z_1(s)) \\ E(Z_2(s)) \\ \vdots \\ \vdots \\ E(Z_k(s) \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \vdots \\ \mu_k \end{pmatrix} = \boldsymbol{\mu}.
$$

We want the predictor $\hat{Z}_1(s_0)$ to be unbiased, that is $E\left(\hat{Z}_1(s_0)\right) = \mu_1$:

$$\begin{aligned}
E\left(\hat{Z}_1(s_0)\right) &= \sum_{j=1}^{k}\sum_{i=1}^{n} w_{ji} E\left(Z_j(s_i)\right) \\
&= w_{11}E(z_1(s_1)) + w_{12}E(z_1(s_2)) + \ldots + w_{1n}E(z_1(s_n)) \\
&+ w_{21}E(z_2(s_1)) + w_{22}E(z_2(s_2)) + \ldots + w_{2n}E(z_2(s_n)) \\
&+ \quad \vdots \quad + \quad \vdots \quad + \ldots + \quad \vdots \\
&+ w_{k1}E(z_k(s_1)) + w_{k2}E(z_k(s_2)) + \ldots + w_{kn}E(z_k(s_n)) \\
&= \sum_{i=1}^{n} w_{1i}\mu_1 + \sum_{i=1}^{n} w_{2i}\mu_2 + \ldots + \sum_{i=1}^{n} w_{ki}\mu_k = \mu_1.
\end{aligned}$$

Therefore, we must have the following set of constraints:

$$\begin{aligned}
\sum_{i=1}^{n} w_{1i} &= 1 \\
\sum_{i=1}^{n} w_{2i} &= 0 \\
\vdots \quad & \vdots \quad \vdots \\
\sum_{i=1}^{n} w_{ki} &= 0
\end{aligned}$$

As with the other forms of kriging, co-kriging minimizes the mean squared error of prediction (MSE):

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

or

$$min \;\; \sigma_e^2 = E\left[Z(s_0) - \sum_{j=1}^{k}\sum_{i=1}^{n} w_{ji} Z_j(s_i)\right]^2$$

subject to the constraints:

$$\begin{aligned}
\sum_{i=1}^{n} w_{1i} &= 1 \\
\sum_{i=1}^{n} w_{2i} &= 0 \\
\vdots \quad & \vdots \quad \vdots \\
\sum_{i=1}^{n} w_{ki} &= 0
\end{aligned}$$

For smplicity, lets assume $k = 2$, in other words, we observe variables Z_1 and Z_2 and we want to predict Z_1. Therefore,

$$min \quad \sigma_e^2 = E\left[Z(s_0) - \sum_{i=1}^n w_{1i}Z_1(s_i) - \sum_{i=1}^n w_{2i}Z_2(s_i)\right]^2$$

Let's add the following quantities: $-\mu_1 + \mu_1 + \sum_{i=1}^n w_{2i}\mu_2$:

$$min \quad \sigma_e^2 = E\left[Z(s_0) - \sum_{i=1}^n w_{1i}Z_1(s_i) - \sum_{i=1}^n w_{2i}Z_2(s_i) - \mu_1 + \mu_1 + \sum_{i=1}^n w_{2i}\mu_2\right]^2$$

or

$$min \quad \sigma_e^2 = E\left[(Z(s_0) - \mu_1) - \sum_{i=1}^n w_{1i}[Z_1(s_i) - \mu_1] - \sum_{i=1}^n w_{2i}[Z_2(s_i) - \mu_2]\right]^2$$

We complete the square above to get:

$$\begin{aligned}
[Z(s_0) - \mu_1]^2 \quad & - \quad 2\sum_{i=1}^n w_{1i}[Z_1(s_0) - \mu_1][Z_1(s_i) - \mu_1] \\
& - \quad 2\sum_{i=1}^n w_{2i}[Z_1(s_0) - \mu_1][Z_2(s_i) - \mu_2] \\
& + \quad \sum_{i=1}^n\sum_{j=1}^n w_{1i}w_{1j}[Z_1(s_i) - \mu_1][Z_1(s_j) - \mu_1] \\
& + \quad \sum_{i=1}^n\sum_{j=1}^n w_{2i}w_{2j}[Z_2(s_i) - \mu_2][Z_2(s_j) - \mu_2] \\
& + \quad 2\left[\sum_{i=1}^n w_{1i}[Z_1(s_i) - \mu_1]\right]\left[\sum_{i=1}^n w_{2i}[Z_2(s_i) - \mu_2]\right]
\end{aligned}$$

It can be shown that the last term of the expression above is equal to:

$$2\left[\sum_{i=1}^n w_{1i}[Z_1(s_i) - \mu_1]\right]\left[\sum_{i=1}^n w_{2i}[Z_2(s_i) - \mu_2]\right] = 2\sum_{i=1}^n\sum_{j=1}^n w_{1i}w_{2j}[Z_1(s_i) - \mu_1][Z_2(s_j) - \mu_2]$$

Find now the expected value of the above expression:

$$\begin{aligned}
min \quad E\left[Z(s_0) - \mu_1\right]^2 \quad & - \quad 2\sum_{i=1}^n w_{1i}E[Z_1(s_0) - \mu_1][Z_1(s_i) - \mu_1] \\
& - \quad 2\sum_{i=1}^n w_{2i}E[Z_1(s_0) - \mu_1][Z_2(s_i) - \mu_2] \\
& + \quad \sum_{i=1}^n\sum_{j=1}^n w_{1i}w_{1j}E[Z_1(s_i) - \mu_1][Z_1(s_j) - \mu_1] \\
& + \quad \sum_{i=1}^n\sum_{j=1}^n w_{2i}w_{2j}E[Z_2(s_i) - \mu_2][Z_2(s_j) - \mu_2] \\
& + \quad 2\sum_{i=1}^n\sum_{j=1}^n w_{1i}w_{2j}E[Z_1(s_i) - \mu_1][Z_2(s_j) - \mu_2]
\end{aligned}$$

Finally, with the Lagrange multipliers we get:

$$
\begin{aligned}
min \;\; \sigma_1^2 \;\; & - \;\; 2\sum_{i=1}^{n} w_{1i} C[Z_1(s_0), Z_1(s_i)] \\
& - \;\; 2\sum_{i=1}^{n} w_{2i} C[Z_1(s_0), Z_2(s_i)] \\
& + \;\; \sum_{i=1}^{n}\sum_{j=1}^{n} w_{1i} w_{1j} C[Z_1(s_i), Z_1(s_j)] \\
& + \;\; \sum_{i=1}^{n}\sum_{j=1}^{n} w_{2i} w_{2j} C[Z_2(s_i), Z_2(s_j)] \\
& + \;\; 2\sum_{i=1}^{n}\sum_{j=1}^{n} w_{1i} w_{2j} C[Z_1(s_i), Z_2(s_j)] \\
& - \;\; 2\lambda_1[\sum_{i=1}^{n} w_{1i} - 1] - 2\lambda_2[\sum_{i=1}^{n} w_{2i} - 0]
\end{aligned}
\tag{1}
$$

The unknowns are the weights $w_{11}, \ldots, w_{1n}$ and $w_{21}, \ldots, w_{2n}$ and the two Lagrange multipliers λ_1 and λ_2. We take the derivatives with respect to these unknowns and set them equal to zero.

$$
-2C[Z_1(s_0), Z_1(s_i)] + 2\sum_{j=1}^{n} w_{1j} C[Z_1(s_i), Z_1(s_j)] + 2\sum_{j=1}^{n} w_{2j} C[Z_1(s_i), Z_2(s_j)] - 2\lambda_1 = 0, \; i = 1, \ldots n \tag{2}
$$

$$
-2C[Z_1(s_0), Z_2(s_i)] + 2\sum_{j=1}^{n} w_{2i} C[Z_2(s_i), Z_2(s_j)] + 2\sum_{j=1}^{n} w_{1j} C[Z_2(s_i), Z_1(s_j)] - 2\lambda_2 = 0, \; i = 1, \ldots n \tag{3}
$$

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

$$
\sum_{i=1}^{n} w_{2i} = 0
$$

Note: for simplicity we will denote the covariances involving Z_1 with C_{11}, the covariances involving Z_2 with C_{22}, and the cross-covariance between Z_1 and Z_2 with C_{12}. For example,

$$
\begin{aligned}
\sigma_1^2 & \equiv C_{11}(0) \\
C[Z_1(s_0), Z_1(s_i)] & \equiv C_{11}(s_0, s_i) \\
C[Z_1(s_i), Z_1(s_j)] & \equiv C_{11}(s_i, s_j) \\
C[Z_1(s_i), Z_2(s_j)] & \equiv C_{12}(s_i, s_j) \\
C[Z_1(s_0), Z_2(s_i)] & \equiv C_{12}(s_0, s_i) \\
C[Z_2(s_i), Z_2(s_j)] & \equiv C_{22}(s_i, s_j) \\
C[Z_2(s_i), Z_1(s_j)] & = C_{21}(s_i, s_j)
\end{aligned}
$$

We get the following co-kriging system in matrix form:

The co-kriging system in terms of covariance

$$\begin{pmatrix}
C_{11}(s_1,s_1) & C_{11}(s_1,s_2) & \dots & C_{11}(s_1,s_n) & C_{12}(s_1,s_1) & C_{12}(s_1,s_2) & \dots & C_{12}(s_1,s_n) & 1 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
C_{11}(s_n,s_1) & C_{11}(s_n,s_2) & \dots & C_{11}(s_n,s_n) & C_{12}(s_n,s_1) & C_{12}(s_n,s_2) & \dots & C_{12}(s_n,s_n) & 1 & 0 \\
C_{21}(s_1,s_1) & C_{21}(s_1,s_2) & \dots & C_{21}(s_1,s_n) & C_{22}(s_1,s_1) & C_{22}(s_1,s_2) & \dots & C_{22}(s_1,s_n) & 0 & 1 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
C_{21}(s_n,s_1) & C_{21}(s_n,s_2) & \dots & C_{21}(s_n,s_n) & C_{22}(s_n,s_1) & C_{22}(s_n,s_2) & \dots & C_{22}(s_n,s_n) & 0 & 1 \\
1 & 1 & \dots & 1 & 0 & 0 & \dots & 0 & 0 & 0 \\
0 & 0 & \dots & 0 & 1 & 1 & \dots & 1 & 0 & 0
\end{pmatrix}
\begin{pmatrix}
w_{11} \\ \vdots \\ \vdots \\ w_{1n} \\ w_{21} \\ \vdots \\ \vdots \\ w_{2n} \\ -\lambda_1 \\ -\lambda_2
\end{pmatrix}
=
\begin{pmatrix}
C_{11}(s_0,s_1) \\ \vdots \\ \vdots \\ C_{11}(s_0,s_n) \\ C_{12}(s_0,s_1) \\ \vdots \\ \vdots \\ C_{12}(s_0,s_n) \\ 1 \\ 0
\end{pmatrix}$$

The co-kriging system is written as $\boldsymbol{\Sigma w} = \boldsymbol{c}$, where the vectors $\boldsymbol{w}, \boldsymbol{c}$ have dimensions $(2n+2) \times 1$ and the matrix $\boldsymbol{\Sigma}$ has dimensions $(2n+2) \times (2n+2)$. The weights will be obtained by $\boldsymbol{w} = \boldsymbol{\Sigma^{-1} c}$.

The co-kriging system can be written in terms of the variogram only when $C_{12}(s_i, s_j) = C_{21}(s_i, s_j)$. In general this is not true (see next pages), but if it is, then we simply substitute $C(\cdot)$ with $-\gamma(\cdot)$.

Definition of cross-covariance:

$$C_{uv}(h) = E\left[Z_U(s) - \mu_U\right]\left[Z_V(s+h) - \mu_V\right], \; C_{UV}(h) \neq C_{VU}(h).$$

Definition of cross-variogram:

$$2\gamma_{uv}(h) = E\left[Z_U(s) - Z_U(s+h)\right]\left[Z_V(s) - Z_V(s+h)\right], \quad \gamma_{UV}(h) = \gamma_{VU}(h).$$

Variance of the predicted value:
We multiply equation (2) by w_{1i} and equation (3) w_{2i} and sum over $i = 1, \ldots, n$. This is what we get:

$$-\sum_{i=1}^{n} w_{1i}C_{11}(s_0, s_i) + \sum_{i=1}^{n}\sum_{l=1}^{n} w_{1i}w_{1l}C_{11}(s_i, s_l) + \sum_{i=1}^{n}\sum_{j=1}^{n} w_{1i}w_{2j}C_{12}(s_i, s_j) - \sum_{i=1}^{n} w_{1i}\lambda_1 \quad = \quad 0 \tag{4}$$

$$-\sum_{i=1}^{n} w_{2i}C_{12}(s_0, s_i) + \sum_{i=1}^{n}\sum_{l=1}^{n} w_{2i}w_{2l}C_{22}(s_i, s_l) + \sum_{i=1}^{n}\sum_{j=1}^{n} w_{1i}w_{2j}C_{12}(s_i, s_j) - \sum_{i=1}^{n} w_{2i}\lambda_2 = 0 \quad = \quad 0 \tag{5}$$

To simplify the expression for the variance of the predicted value we substitute (4) and (5) into equation (1):

$$\sigma^2(s_0) = C_{11}(0) - \sum_{i=1}^{n} w_{1i}C_{11}(s_0, s_i) - \sum_{i=1}^{n} w_{2i}C_{12}(s_0, s_i) + \lambda_1$$

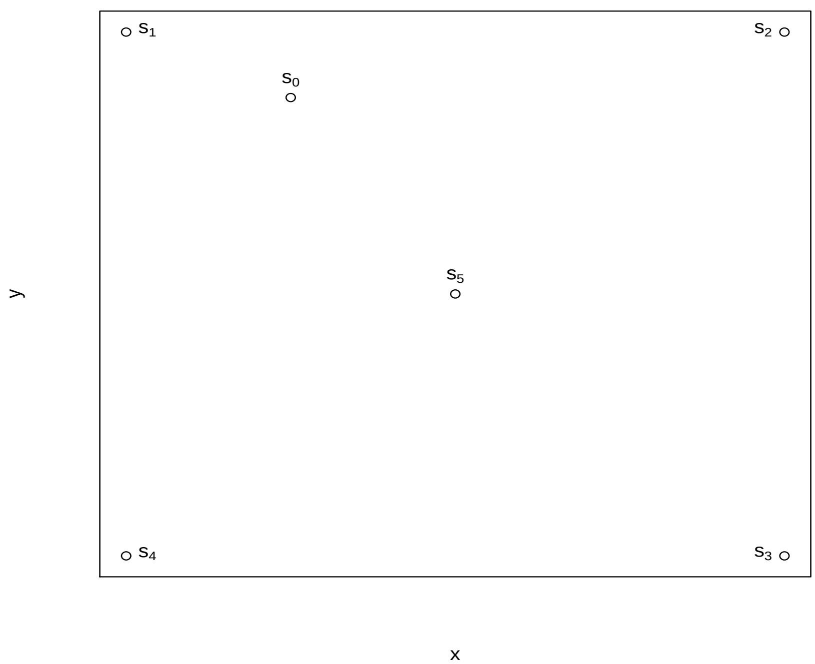
or

$$\sigma^2(s_0) = C_{11}(0) - \sum_{k=1}^{2}\sum_{i=1}^{n} w_{ki}C_{ki}(s_0, s_i) + \lambda_1.$$

Other benefits of co-kriging:
Consider the situation where the target variable is under-sampled. In the figure below, we want to predict $Z(s_0)$ from observed data at locations s_1, s_2, s_3, s_4, s_5. However, the data available at these locations are: $Z_1(s_1), Z_2(s_1), Z_2(s_2), Z_2(s_3), Z_2(s_4), Z_2(s_5)$. The predictor will be:

$$\hat{Z}_1(s_0) = w_{11}Z_1(s_1) + \sum_{i=1}^{5} w_{2i}Z_2(s_i)$$

with the constraints, $w_{11} = 1$, $\sum_{i=1}^{5} w_{2i} = 0$.



Co-kriging using gstat

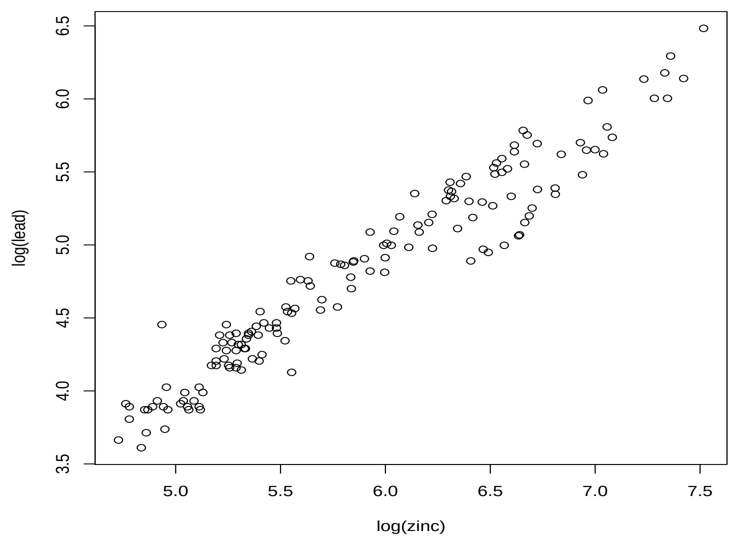
We continue with the same data (Maas river data).

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

image.orig <- image

#Load the library:
library(gstat)
```

Some background:
A co-located variable used in co-kriging must be correlated with the target variable and therefore have some predicting power. A simple way to choose a co-located variable is to compute the correlations between the target variable and co-located variables. Besides the predicting power that co-located variables may add to the kriging system, they may also be cheaper and faster to sample.

In our data set we will treat *log_lead* as our target variable and we will use *log_zinc* (together with *log_lead*) to make co-kriging predictions. Here is the plot of the two variables and their correlation:

```
> plot( log(a$zinc), log(a$lead), xlab="log(zinc)", ylab="log(lead)")
```




```
> cor(log(a$zinc), log(a$lead))
[1] 0.967162
```

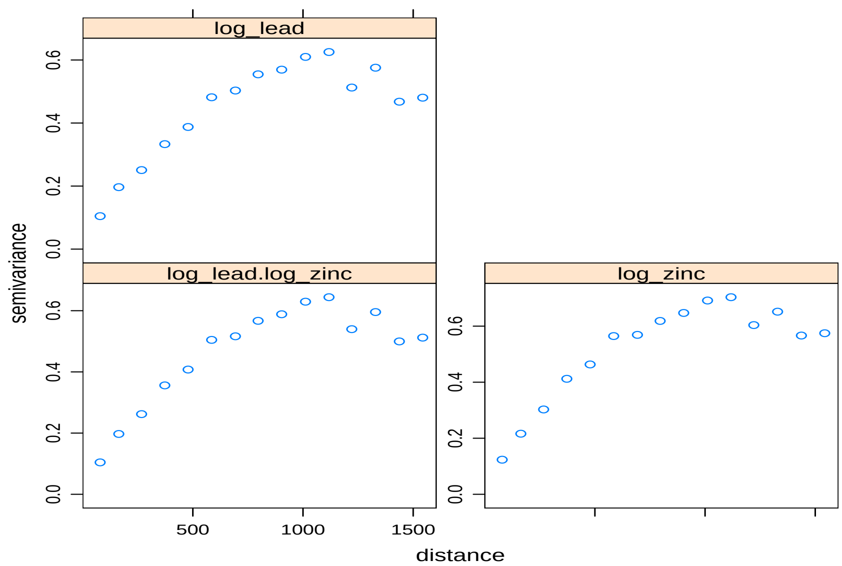
We observe high correlation between the two variables. This is not a surprise since activities associated with lead pollution also produce zinc. In addition, zinc is transported in a similar way with lead.

We begin now the co-kriging procedure with gstat. First we create a gstat object including both variables one at a time.

```
> g1 <- gstat(id="log_lead", formula = log(lead)~1, locations = ~x+y, data = a)
> g1 <- gstat(g1,id="log_zinc", formula = log(zinc)~1, locations = ~x+y, data = a)
```

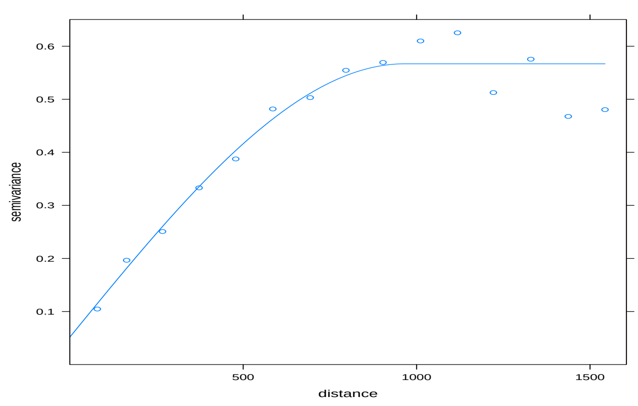
Note: First we include the variable *log_lead* to create a gstat object and then we append to it the variable *log_zinc*. We can now plot the variogram of the object g to get the following plot:

```
> plot(variogram(g1))
```



We obtained variograms for each one of the two variables as well as the cross-variogram. First we will fit a model variogram to the sample variogram of the target variable *log_lead*. Here are the commands and the plot:

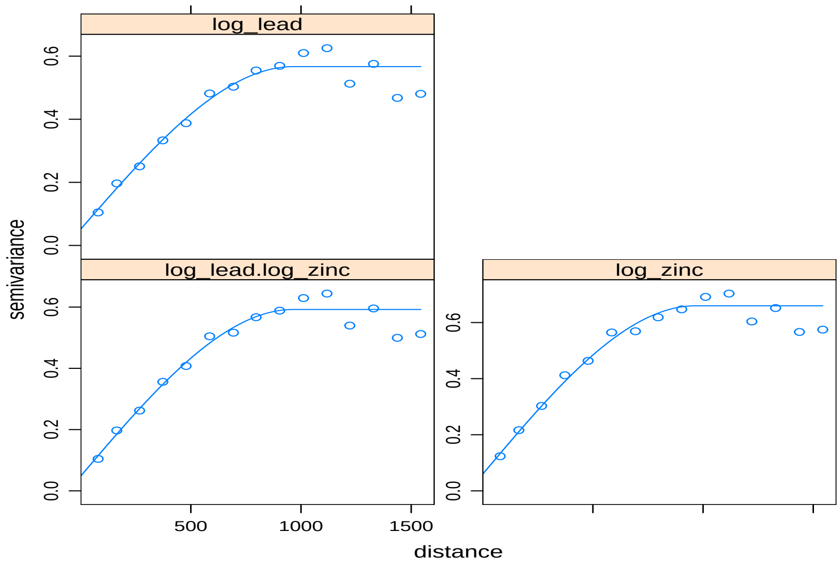
```
> g <- gstat(id="log_lead", formula = log(lead)~1, locations = ~x+y, data = a)
> v.fit <- fit.variogram(variogram(g), vgm(0.5,"Sph",1000,0.1))
> plot(variogram(g),v.fit)
```



We are ready now to fit a model variogram to all the sample variograms. We use the function fit.lmc (linear model of co-regionalization) to fit a model variogram to the sample variogram of the co-located variable and to the sample cross-variogram between the target and co-located variables. We begin with the gstat object g1 that was created earlier:

```
> g1 <- gstat(id="log_lead", formula = log(lead)~1, locations = ~x+y, data = a)
> g1 <- gstat(g1,id="log_zinc", formula = log(zinc)~1, locations = ~x+y, data = a)
> vm <- variogram(g1)
> vm.fit <- fit.lmc(vm, g1, model=v.fit)
> plot(vm, vm.fit)
```

Here are the plots of the three fitted variograms:



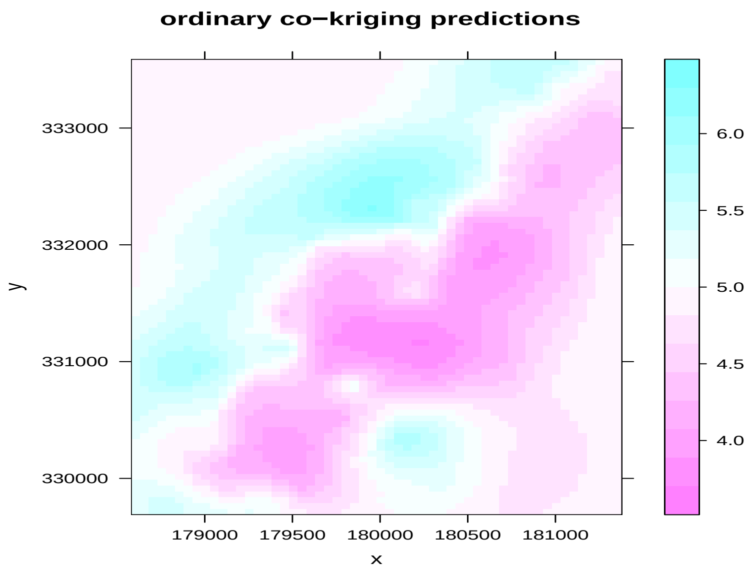
We create the grid for predictions and perform co-kriging. The function that performs co-kriging in gstat is predict.gstat and its arguments in our example are vm.fit and grd.

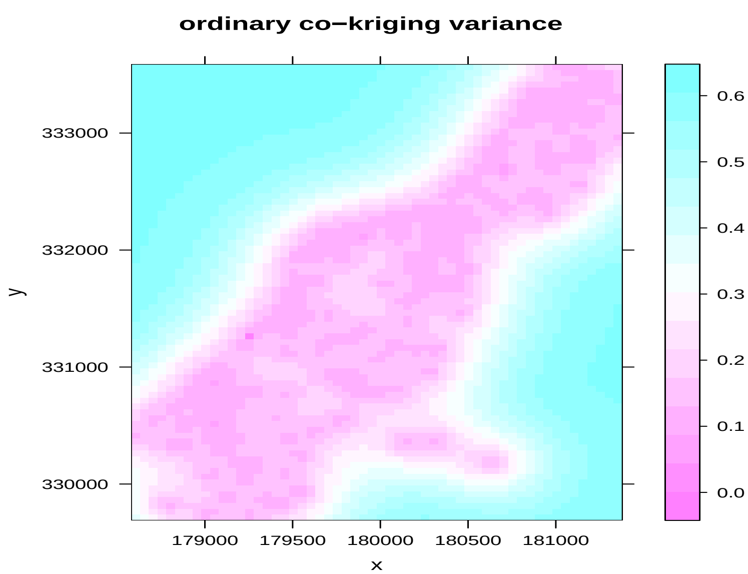
```
> x.range <- as.integer(range(a[,1]))
> y.range <- as.integer(range(a[,2]))
> grd <- expand.grid(x=seq(from=x.range[1], to=x.range[2], by=50),
  y=seq(from=y.range[1], to=y.range[2], by=50))

> ck <- predict.gstat(vm.fit, grd)
```

As always, the predictions are under ck$log_lead.pred and the variances under ck$log_lead.var. The following commands will produce the corresponding raster maps.

```
> library(lattice)
> levelplot(ck$log_lead.pred~x+y, ck, aspect = "iso",
main = "ordinary co-kriging predictions")

> levelplot(ck$log_lead.var~x+y, ck, aspect = "iso",
main = "ordinary co-kriging variance")
```

Here are the co-kriging raster maps:


ordinary co−kriging predictions
333000
332000
331000
330000
y
179000
179500
180000
180500
181000
x
6.0
5.5
5.0
4.5
4.0


ordinary co−kriging variance
333000
332000
331000
330000
y
179000
179500
180000
180500
181000
x
0.6
0.5
0.4
0.3
0.2
0.1
0.0

Or using the `image.orig` **function:**

```
#Collapse the vector of the predicted values into a matrix:
qqq <- matrix(ck$log_lead.pred, length(seq(from=x.range[1], to=x.range[2],
       by=50)), length(seq(from=y.range[1], to=y.range[2], by=50)))

#Construct the raster map of the predicted values:
> image.orig(seq(from=x.range[1], to=x.range[2], by=50),
  seq(from=y.range[1],to=y.range[2], by=50), qqq,
  xlab="West to East",ylab="South to North", main="Raster map of the
  predicted values")


#Collapse the vector of the variances into a matrix:
> qqq1 <- matrix(ck$log_lead.var, length(seq(from=x.range[1], to=x.range[2],
          by=50)), length(seq(from=y.range[1], to=y.range[2], by=50)))

#Construct the raster map of the variances:
> image.orig(seq(from=x.range[1], to=x.range[2], by=50),
  seq(from=y.range[1],to=y.range[2], by=50), qqq1,
  xlab="West to East",ylab="South to North", main="Raster map
  of the variances")
```