Topic 1: Kriging

This topic applies the previously considered methods for spatial extrapolation. In particular, we consider

- Introduction to **kriging methods**.
- **Three types** of kriging.
- **Kriging with geoR package**.
- Example of **kriging for Meuse data.**

Kriging methods

The origin of the word **kriging** is by D.G. Krige, a South African mining engineer, who in the 1950's developed empirical methods for predicting ore grades at unsampled locations using the known grades of ore sampled at nearby sites.

Krige's original method is what is now called ordinary kriging. There have been several modifications and extensions (e.g., universal kriging, indicator kriging, disjunctive kriging, and others) which are based on very similar ideas.

- **Simple kriging** assumes a known constant trend: $m(x) = 0$.
- **Ordinary kriging** assumes an unknown constant trend: $m(x) = m$.
- **Universal kriging** assumes a general linear trend model $m(x) = \sum_{k=0}^{p} \beta_k f_k(x)$.

Let $\mathbf{s}_0$ denote an arbitrary location in an region D. Usually this is an unsampled location but it need not be.

Goal of kriging: to predict the value of $X(\mathbf{s}_0)$ at $\mathbf{s}_0$.
The kriging predictor must satisfy the next properties:

1. It is a **linear** combination of the data values, i.e.,

$$\hat{X}(\mathbf{s}_0) = \sum_{i=1}^{n} \lambda_i X(\mathbf{s}_i).$$

2. It is **unbiased**, i.e., it satisfies

$$E[\hat{X}(\mathbf{s}_0)] = E[X(\mathbf{s}_0)].$$

3. Among all functions of the data that satisfy the first 2 properties, it is the **best** in the sense that minimizes the variance of prediction error, $\text{Var}[\hat{X}(\mathbf{s}_0) - X(\mathbf{s}_0)]$.

Let us consider the ordinary kriging which is based on the following model:

1. The mean $m(\mathbf{x})$ is assumed to be an unknown constant m.
2. The semivariogram $\gamma(\mathbf{h})$ is assumed to be known.

The kriging's properties above imply

$$\sum_{i=1}^{n} \lambda_i = 1.$$

The properties we have imposed on our predictor lead us to minimize

$$\text{var}[\hat{X}(\mathbf{s}_0) - X(\mathbf{s}_0)]$$

subject to the restriction $\sum_{i=1}^{n} \lambda_i = 1$.

The method of Lagrange multipliers (from calculus) is applicable here. It can be used to show that the optimal coefficients $\lambda_1, \ldots, \lambda_n$ are the first n elements of the vector $\boldsymbol{\lambda}_O$ that satisfies **the ordinary kriging equations**:

$$\boldsymbol{\Gamma}_O \boldsymbol{\lambda}_O = \boldsymbol{\gamma}_O,$$

where

$$\boldsymbol{\lambda}_O = (\lambda_1, \ldots, \lambda_n, m)'$$

$$\boldsymbol{\gamma}_O = [\gamma(\mathbf{s}_1 - \mathbf{s}_0), \ldots, \gamma(\mathbf{s}_n - \mathbf{s}_0), 1]'$$

$$\boldsymbol{\Gamma}_O = \begin{cases} \gamma(\mathbf{s}_i - \mathbf{s}_j), & \text{for } i = 1, \ldots, n; j = 1, \ldots, n; \\ 1, & \text{for } i = n+1; j = 1, \ldots, n; \\ 0, & \text{for } i = n+1; j = n+1. \end{cases}$$

The minimized variance, called the **kriging variance,** is

$$\sigma^2_{OK}(\mathbf{s}_0) = \sum_{i=1}^{n} \lambda_i \gamma(\mathbf{s}_i - \mathbf{s}_0) + m = \boldsymbol{\lambda}'_O \boldsymbol{\gamma}_O.$$

- Ordinary kriging is derived under the assumption that the semivariogram is known. In practice, the semivariogram is unknown and must be estimated. The estimator $\hat{\gamma}(\cdot)$ replaces $\gamma(\cdot)$ in the kriging equations and in the expression for the kriging variance.

- Ordinary kriging is a linear combination of *all* the observations. In practice, often only the observations within a **moving window** or **kriging neighborhood** are used.

- **Environmental monitoring programs.** Note that the kriging variance $\sigma^2_{OK}(\mathbf{s}_0)$ at any given site $\mathbf{s}_0$ does not depend on the data. Thus, it can be used to answer sampling design questions, such as where to take one more observation to maximize the reduction in σ^2_{OK} at a certain point, or where to take one more observation to minimize the maximum (or average) value of σ^2_{OK} over the entire spatial domain (R example will be presented later).

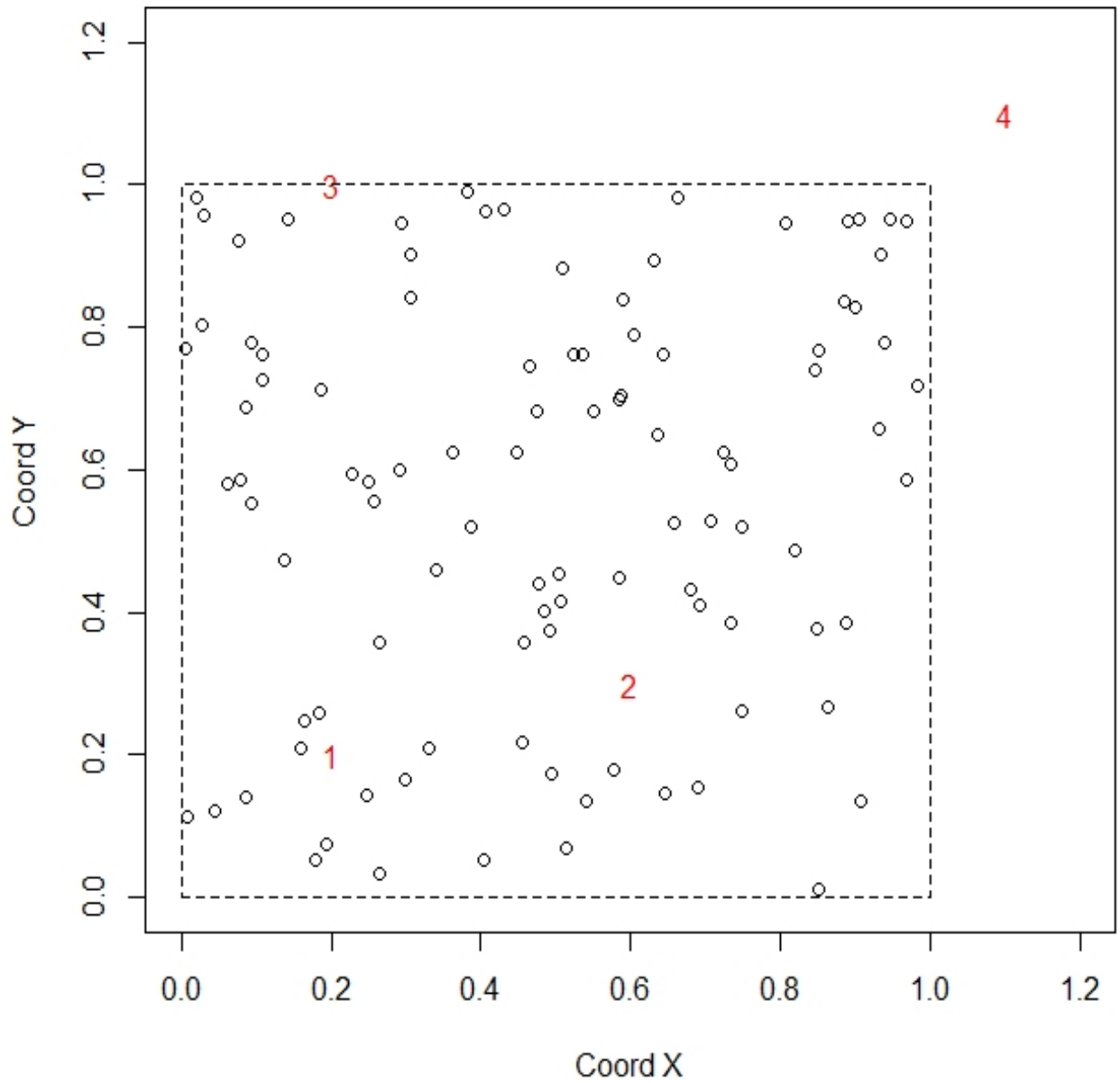
geoR Kriging

In GEOR conventional geostatistical spatial interpolation (kriging) can be performed with options for:

1. Simple kriging.
2. Ordinary kriging.
3. Universal kriging.

Let us consider the prediction at the four locations labeled 1, 2, 3, 4 and indicated in red in the figure below.

```
> library(geoR)
> data(s100)
> plot(s100$coords, xlim=c(0,1.2), ylim=c(0,1.2),
+ xlab="Coord X", ylab="Coord Y")
> loci <- matrix(c(0.2, 0.6, 0.2, 1.1, 0.2, 0.3, 1.0, 1.1),
+ ncol=2)
> text(loci, as.character(1:4), col="red")
> polygon(x=c(0,1,1,0), y=c(0,0,1,1), lty=2)
```


Coord Y
Coord X
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1
2
3
4

One can apply the commands below to perform the ordinary kriging by using the parameters estimated by ordinary least squares with nugget as

```
> bin1 <- variog(s100, uvec=seq(0,1,l=11))
> ols.n <- variofit(bin1, ini = c(1,0.5), nugget=0.5,
+  weights="equal", cov.model = "exponential")
> kc4 <- krige.conv(s100, locations = loci,
+ krige = krige.control(obj.m = ols.n))
```

The predicted values at the four specified locations can be found as

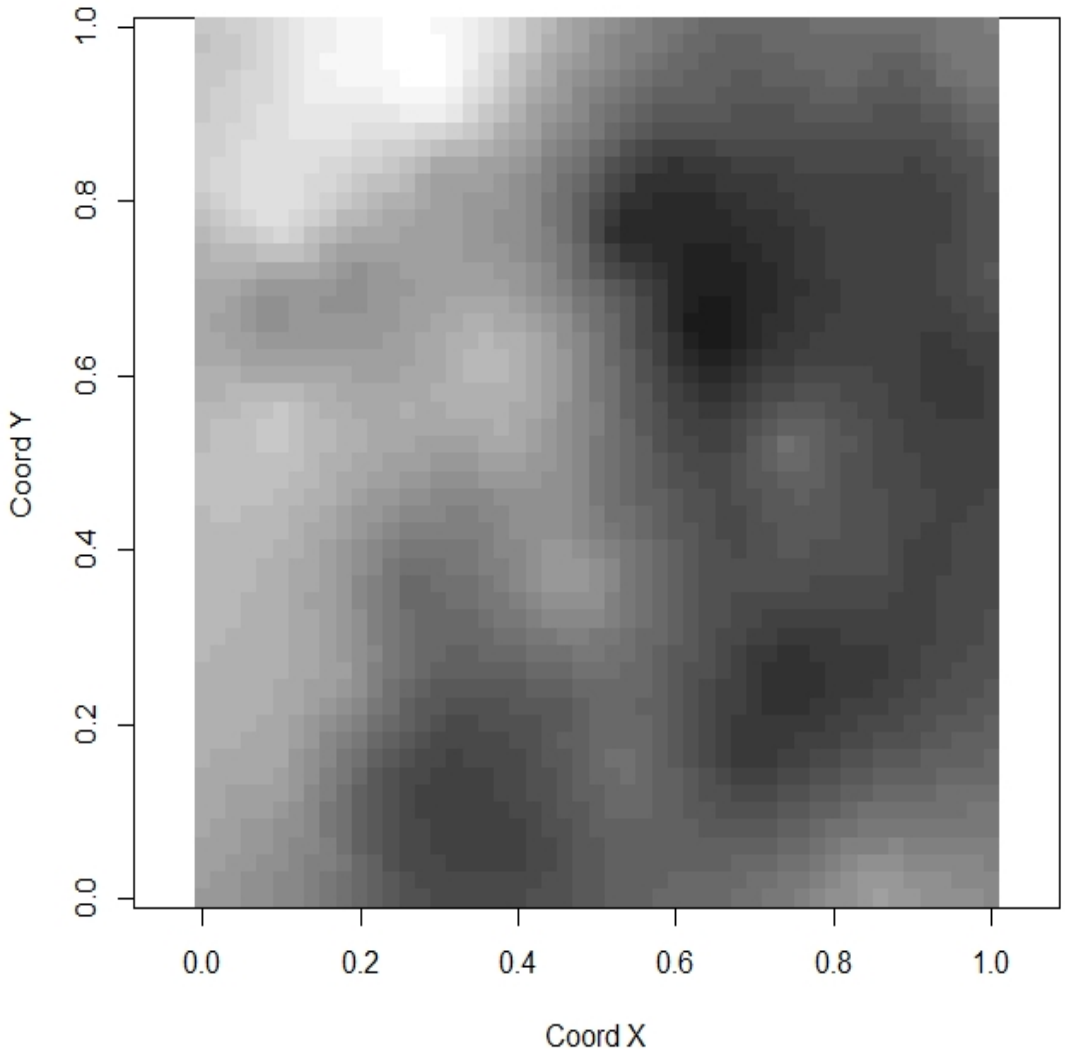
```
> kc4
$predict
[1]  0.8728067  1.2611669 -0.6888013  0.7924067
$krige.var
[1] 0.2854335 0.3709802 0.3839016 0.7356138
$beta.est
beta 0.5968567 ...
```

The output is a list including the predicted values (kc4$predict) and the kriging variances (kc4$krige.var).

Now we use the ordinary kriging to perform prediction on a grid covering the area and to display the results. The R commands are:

```
> pred.grid <-  expand.grid(seq(0,1, l=51), seq(0,1, l=51))
> kc <- krige.conv(s100, loc = pred.grid,
+ krige = krige.control(obj.m = ols.n))
> image(kc, loc = pred.grid, col=gray(seq(1,0.1,l=30)),
+ xlab="Coord X", ylab="Coord Y")
```

The grey level indicates values of s100 at the corresponding locations.

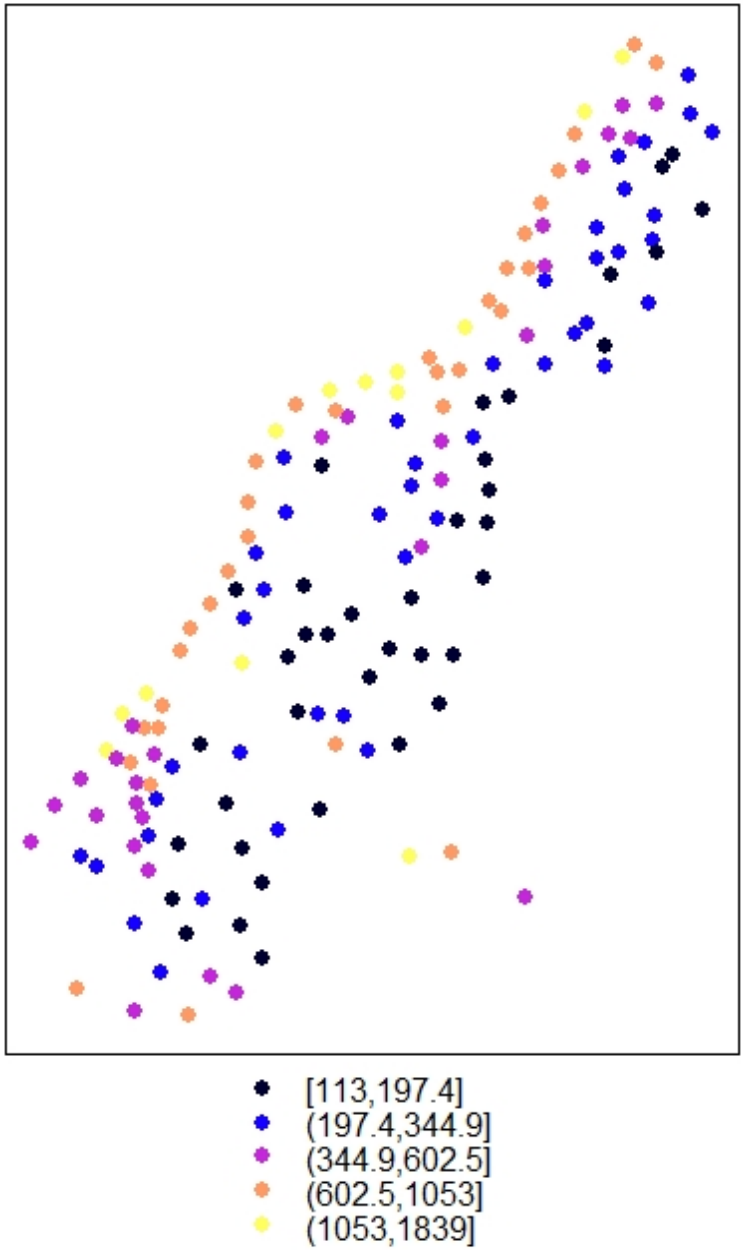

Coord Y
0.0
0.2
0.4
0.6
0.8
1.0
Coord X
0.0
0.2
0.4
0.6
0.8
1.0

Kriging for Meuse data

Now we perform a different type of kriging by using the fitted model.

We will use Meuse zinc data and start by fitting a variogram model as in the previous week:

```
> library(gstat)
> library(sp)
> data(meuse)
> data(meuse.grid)
> coordinates(meuse) <- c("x", "y")
> coordinates(meuse.grid) <- ~ x + y
> spplot(meuse,"zinc", do.log = T)
> v  <- variogram(log(zinc) ~ 1, meuse)
> v.fit <- fit.variogram(v, vgm(1, "Sph", 800, 1))
```


[113,197.4]
(197.4,344.9]
(344.9,602.5]
(602.5,1053]
(1053,1839]

Then we will apply simple and ordinary kriging with the chosen spherical variogram model:

```
> lz.sk <- krige(log(zinc)~1, meuse, meuse.grid, v.fit,
+ beta = 5.9)
> lz.ok <- krige(log(zinc)~1, meuse, meuse.grid, v.fit)
```

To see the obtained results one can use the commands

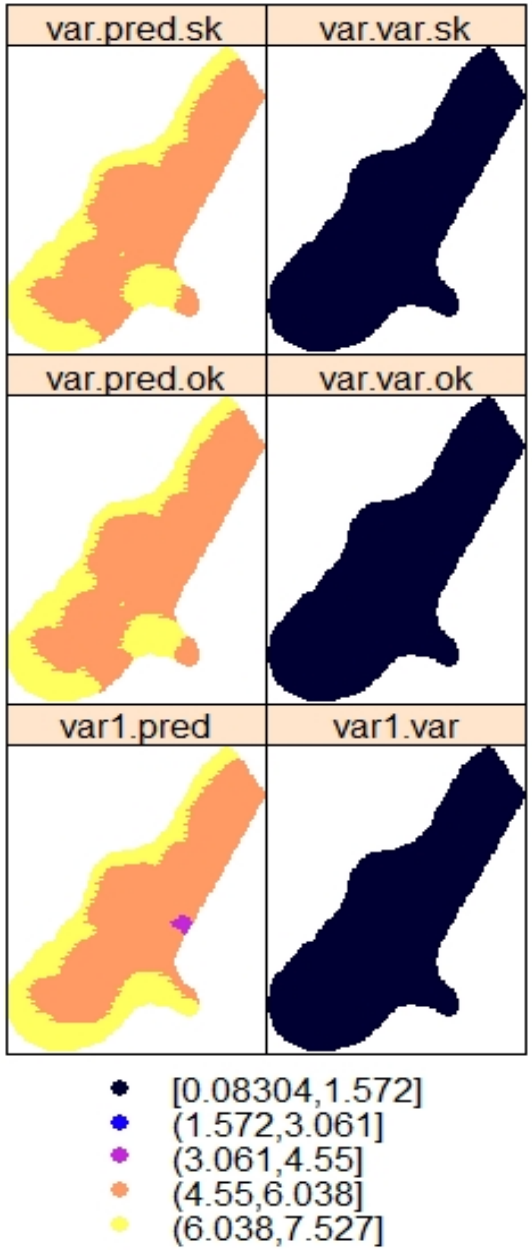
```
> lz.sk
coordinates var1.pred   var1.var
1   (181180, 333740)  6.452149 0.31600258
2   (181140, 333700)  6.588396 0.25007206
...
> class(lz.sk)
[1] "SpatialPointsDataFrame"
attr(,"package")
[1] "sp"
```

For the universal kriging we fit a variogram model to residuals first:

```
> vt <- variogram(log(zinc) ~ sqrt(dist), meuse)
> vt.fit <- fit.variogram(vt, vgm(1, "Exp", 300, 1))
> lz.uk <- krige(log(zinc)~sqrt(dist), meuse,
+ meuse.grid, vt.fit)
```

To plot all three kriging results on the same plot one can use the commands:

```
> lz.uk$var.pred.sk<-lz.sk$var1.pred
> lz.uk$var.var.sk<-lz.sk$var1.var
> lz.uk$var.pred.ok<-lz.ok$var1.pred
> lz.uk$var.var.ok<-lz.ok$var1.var
> spplot(lz.uk, c("var1.pred","var1.var","var.pred.ok",
+ "var.var.ok", "var.pred.sk","var.var.sk"),layout = c(2,3))
```


var.pred.sk
var.var.sk
var.pred.ok
var.var.ok
var1.pred
var1.var
[0.08304,1.572]
(1.572,3.061]
(3.061,4.55]
(4.55,6.038]
(6.038,7.527]

Summarising, we used

1. **lz.sk** for simple kriging.
2. **lz.ok** for ordinary kriging.
3. **lz.uk** for universal kriging.

The **krige** command chooses the kriging method itself, depending on the provided information (trend coefficients, constant or more complex trend). By default, all spatial predictions method provided by GSTAT use all available observations for each prediction. In many cases, it is more convenient to use only the data in the neighborhood of the prediction location. The reasons for this may be statistical or computational.
Statistical reasons: the hypothesis of constant mean or mean function is valid locally, the assumed model of the variogram is only valid up to a small distance.
Computational issues may involve both memory and speed: kriging for n data requires solving an $n \times n$ system. For large n this may be too slow.