Spatial Prediction using Co-Kriging

Coregionalization and Cokriging

These two concepts are related in the same way as the theory of regionalized variables and (univariate) kriging:

1. Coregionalization is a theoretical model of how several variables spatially co-vary: this is used for . . .
2. Cokriging (CK), which is a method of using
   • supplementary information on a co-variable . . .
   • . . . to improve the prediction of a target variable.

That is, we extend the theory of ordinary kriging of regionalized variables to several variables which we show to have a multivariate spatial cross-correlation as well as the univariate spatial auto-correlation; these are then called co-regionalized variables and can be used to predict.

Cokriging: Motivation

• We want to predict the values of a target variable
  * often hard or expensive to measure . . .
  * . . . so we have relatively few observations of it.
• We have, however, more values of a second variable, which we believe is co-related with the target variable; this is the co-variable.
  * this is often easy and cheap to measure . . .
  * . . . so we have many observations of it.

Strategies for using a co-variable

Co-kriging may be applicable when:
• Values of the co-variable are only used at its sample points, not needed over the whole field.
• Co-located observations (with the target variable) are used for modelling the spatial structure
• Typically there are more observations of the co-variable, i.e. at some points where the target variable was not measured
• Requires a model of coregionalization.

Two sampling intensities, some co-located points

Variograms

Two types of variograms:
1. direct: single regionalized variables, one variogram per variable
2. cross: per pair of regionalised variables
Estimating the cross-variogram

For the two variables \( u \) and \( v \):
\[
\gamma_u = \frac{1}{2m(h)} \sum \left( \frac{z_i(x) - z_i(x + h)}{z_i(x) - z_i(x + h)} \right) \left( \frac{z_j(y) - z_j(y + h)}{z_j(y) - z_j(y + h)} \right)
\]
where \( m(h) \) is the number of point-pairs separated by vector \( h \); for an omnidirectional variogram this is the distance class.

In words: if high differences between point-pairs of one variable are positively associated with high differences between point-pairs of the other variable, they will have a high positive cross-correlation. This can also be a negative relation. If the differences are randomly associated, there will be on average no spatial cross-correlation in this distance class.

Modeling the variograms

The direct and cross-variograms must be modeled together, with some restrictions to ensure that the resulting CK system can be solved.

The simplest way to ensure this is to assume a linear model of co-regionalisation: all variograms are linearly related

- Same model, same range (spatial structure)
- Different sill (overall variability); nugget (uncertainty at sample point)

Others are possible but much more complicated to estimate.

Sample size

- In general, 30 or more points are needed to generate a reasonable sample variogram.
- The most important part of a variogram is its shape near the origin as closest points are given more weight in the interpolation process.

Variogram models

- **Pure nugget** \( \gamma(h) = \begin{cases} 0 & h = 0 \\ C & h > 0 \end{cases} \)
- **Linear** \( \gamma(h) = \begin{cases} C h / a & 0 \leq h \leq a \\ C & h > a \end{cases} \)
- **Spherical** \( \gamma(h) = \begin{cases} \frac{3h}{2a^2} - \frac{1}{2} \left( \frac{h}{a} \right) & 0 \leq h \leq a \\ C & h > a \end{cases} \)
- **Exponential** \( \gamma(h) = Ce^{-h/a} \)
- **Gaussian** \( \gamma(h) = Ce^{-h^2/4a^2} \)
- **Linear** \( \gamma(h) = mh \)
- **Power** \( \gamma(h) = mh^\beta \)

Co-Kriging: Prerequisites

1. Two point data sets, usually with some observations co-located:
   - (a) the target variable \( z \) at locations \( x_1 \ldots x_n \)
   - (b) the co-variable \( w \) at locations \( y_1 \ldots y_m \)
2. Spatial structure in both variables separately (i.e. non-nugget variograms);
3. Spatial structure between the variables (cross-variograms); this can be either a positive or negative spatial correlation;
4. Certain restrictions on the joint spatial structure.
The Co-kriging predictor
Prediction of the target variable at an unknown point \( x_0 \) is computed as the sum of two weighted averages:
1. one of the \( n_z \) sample values of the target variable \( z \); and
2. one of the \( n_w \) sample values of the co-variable \( w \)

\[
z(\mathbf{x}_0) = \sum_{i=1}^{n_z} \lambda_i z(\mathbf{x}_i) + \sum_{j=1}^{n_w} \mu_j w(\mathbf{y}_j)
\]

We want to find the weights \( \lambda \) (for the target variable) and \( \mu \) (for the co-variable) which minimize the prediction variance.

The Co-kriging variance
Suppose there are \( V \) variables (target and co-variables), which are indexed from \( 1 \ldots l \) each has \( n_l \) observations. Variable \( u \) is the target, one of the \( V \). We want to minimize the variance.

\[
\sigma_u^2 = \sum_{j=1}^{V} \sum_{j=1}^{n_j} \lambda_j \gamma_{iu}(x_j, x_0) + \psi_g
\]

So the semivariance of all observations of all variables with the point to be estimated is minimized.

Note there is one of these equations for each variable.

The Co-Kriging system
Solve: \( \mathbf{A} \lambda = \mathbf{b} \), where \( \mathbf{A} \) is built up from the direct and cross-semi-variances of the sample points (both target and co-variable):

\[
\mathbf{A} = \begin{bmatrix}
\Gamma_{zz} & \Gamma_{zw} & 1 & 0 \\
\Gamma_{zw} & \Gamma_{ww} & 0 & 1 \\
1 & 0 & \mathbf{1}^T & 0 \\
0 & 1 & \mathbf{1}^T & 0
\end{bmatrix}
\]

- The vectors of 1’s and 0’s control which of the semi-variances are included in each equation.
- The \( \mathbf{1}^T \) are the matrices of semi-variances (next slide)

Example of a cross-variable matrix
\[
\Gamma_{zw} = \begin{bmatrix}
\gamma_{u_1}(\mathbf{x}_1, \mathbf{y}_1) & \gamma_{u_1}(\mathbf{x}_1, \mathbf{y}_2) & \ldots & \gamma_{u_1}(\mathbf{x}_1, \mathbf{y}_{n_w}) \\
\gamma_{u_2}(\mathbf{x}_2, \mathbf{y}_1) & \gamma_{u_2}(\mathbf{x}_2, \mathbf{y}_2) & \ldots & \gamma_{u_2}(\mathbf{x}_2, \mathbf{y}_{n_w}) \\
\vdots & \vdots & \ddots & \vdots \\
\gamma_{u_l}(\mathbf{x}_l, \mathbf{y}_1) & \gamma_{u_l}(\mathbf{x}_l, \mathbf{y}_2) & \ldots & \gamma_{u_l}(\mathbf{x}_l, \mathbf{y}_{n_w})
\end{bmatrix}
\]

The Co-Kriging system
\( \lambda \) of the target variable is:

\[
\lambda = \begin{bmatrix}
\lambda_1 \\
\vdots \\
\lambda_{n_z} \\
\mu_1 \\
\vdots \\
\mu_{n_w} \\
\psi_z \\
\psi_w
\end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix}
\gamma_{u}(\mathbf{x}_1, \mathbf{y}_0) \\
\gamma_{u}(\mathbf{x}_2, \mathbf{y}_0) \\
\vdots \\
\gamma_{u}(\mathbf{x}_l, \mathbf{y}_0)
\end{bmatrix}
\]

The co-kriging prediction variance is then:

\[
\hat{\sigma}_e^2(x_0) = \mathbf{b}^T \mathbf{C} \lambda_C
\]
Exercises

- Computer program: Geostatistics +