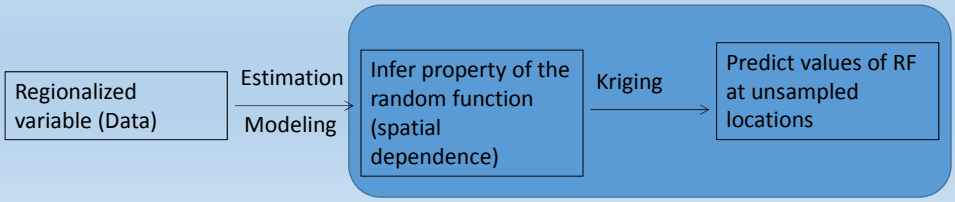
COA 616 Geostatistics in Environmental Sciences

# Lecture 10 – More kriging, stochastic simulation, and model diagnostics

Wei Wu

November 15, 2016




## Types of kriging (use of semivariances)

- Ordinary kriging
  - Intrinsic hypothesis
- Simple kriging: Kriging with known mean
- Kriging with external drift
  - Weaker form of stationarity is allowed, in which the expected values of both Z(x) and [Z(x)-Z(x+h)] may vary regularly with location.
- Block kriging
- Indicator kriging
- Cokriging – Multiple variables may be cross correlated
- Stratified kriging – When a categorical variable is available that splits the area of interest in a number of disjoint areas. No correlation is assumed between residuals from different disjoint areas.

## Simple kriging and universal kriging

- Simple kriging: Kriging with known mean
- Universal kriging - Weaker form of stationarity is allowed, in which the expected values of both Z(x) and [Z(x)-Z(x+h)] may vary regularly with location.

Kriging with external drift: one covariate and it does not include coordinates

Regression kriging:
1) The trend is unbiased estimates but does not guarantee estimates of minimum variance.
2) Estimates of the semivariances obtained from residuals from the trend are biased. They depend in a nonlinear way on the trend parameters, which are estimated with errors
3) It does not allow us to combine trend and kriging from the residuals into a valid prediction variance for the kriging estimates.

## Cokriging

A data set often contains not only the primary variable of interest, but also one or more secondary variables. These secondary variables are usually spatially cross-correlated with the primary variable and thus contain useful information about the primary variable.

Cokriging – a method for estimation that minimizes the variance of the estimation error by exploiting the cross-correlation between several variables. The estimates are derived using secondary variables as well as the primary variable.

The usefulness of the secondary is often enhanced by the fact that the primary variable of interest in undersampled.

Multivariate variogram modelling

$$\gamma_{ij}(h) = E[(Z_i(s) - Z_i(s+h))(Z_j(s) - Z_j(s+h))]$$

$$\gamma_{ij}(h) = E[(Z_i(s) - m_i)(Z_j(s) - m_j)]$$

## Stratified kriging

- Stratified kriging – When a categorical variable is available that splits the area of interest in a number of disjoint areas. No correlation is assumed between residuals from different disjoint areas.

## Model diagnostics – cross validation

It is important to make sure that the variogram model is chosen correctly and it fits well.

Splits the data set into two sets: a modeling set and a validation set

1) The modeling set is used for variogram modeling
2) Krigging on the location of validation set, get $\hat{V}_i^*$ and $\sigma_i^*$
3) Compare measurements and predictions in validation set, Compute standardized residuals, $Z(x_i) = \frac{(V(x_i) - \hat{V}_i^*)}{\sigma_i}$
they should be standard Normal, with mean 0 and variance 1.
4) Examine the residuals for distribution, spatial patterns, outliers etc.

## Stochastic simulation

A kriged map shows best estimate but does not represent the variation well. To obtain a statistical surface that retains the variation we know or believe to be present. Many realizations are needed but they are often needed when the uncertainty of kriging predictions is input to a next level of analysis and spatial correlation plays a role.

Conditional simulation – retain observed data at data locations.
Unconditional simulation – ignore observations and only reproduce means and prescribed variability.

Sequential simulation

1) Compute the conditional distribution given data and previously simulated values using simple kriging – most computationally expensive
2) Draw a value from this conditional distribution
3) Add this value to the dataset
4) Go to the next unvisited location and go back to 1 until all locations are visited

## Review

Data exploration

↓

Variogram

↓

Variogram model

↓

Predictions

↓

Model diagnostics

↓

Simulation

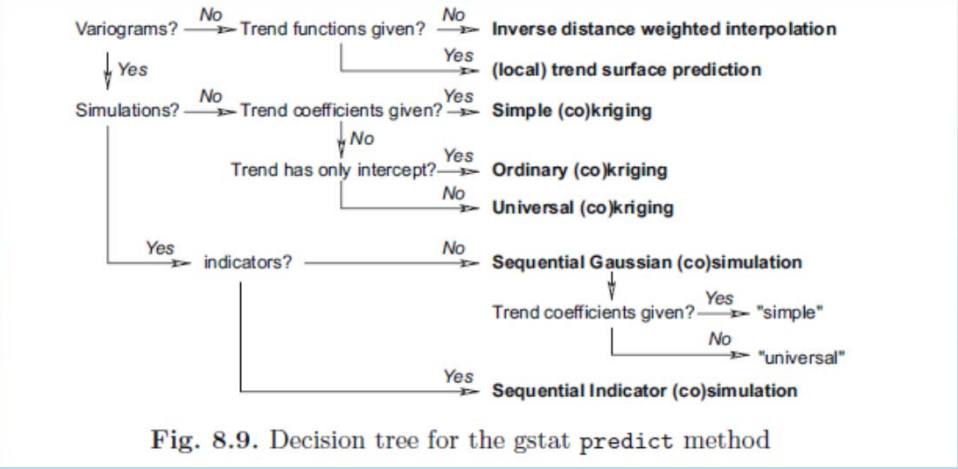



Fig. 8.9. Decision tree for the gstat `predict` method