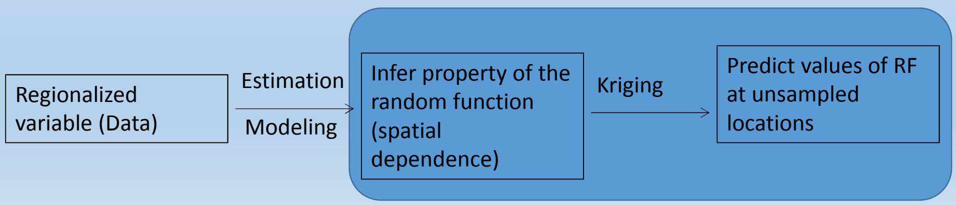
COA 616 Geostatistics in Environmental Sciences

Lecture 7 – Ordinary Kriging

Wei Wu

October 18, 2016



Ordinary kriging (OK)

It is a family of methods to predict a random variable based on observed structure of spatial variability.

It is popular because

1. It is intuitively appealing
2. Estimation variance can be quantified
3. Kriging methods minimize estimation variance

OK – B.L.U.E.

B.L.U.E. – Best Linear Unbiased Estimator

Best – Minimize the variance of errors σ_R^2

Linear – Estimates are weighted linear combination of available data

Unbiased – The mean residuals or errors is equal to 0 $m_R = 0$

$$\hat{v}(x_0) = \sum_{i=1}^{n} \lambda_i v(x_i)$$

$$R(x_0) = \hat{v}(x_0) - v(x_0) = \sum_{i=1}^{n} \lambda_i v(x_i) - v(x_0)$$

$$E\{R(x_0)\} = E\{\sum_{i=1}^{n} \lambda_i v(x_i) - v(x_0)\}$$

$$= \sum_{i=1}^{n} \lambda_i E\{v(x_i)\} - E\{v(x_0)\}$$

OK – Unbiased

$$\hat{v}(x_0) = \sum_{i=1}^{n} \lambda_i v(x_i)$$

$$R(x_0) = \hat{v}(x_0) - v(x_0) = \sum_{i=1}^{n} \lambda_i v(x_i) - v(x_0)$$

$$E\{R(x_0)\} = E\{\sum_{i=1}^{n} \lambda_i v(x_i) - v(x_0)\}$$

$$= \sum_{i=1}^{n} \lambda_i E\{v(x_i)\} - E\{v(x_0)\}$$

$$Stationarity: \mathrm{E\{v(x_i)\}} = E\{v(x_0)\} = E\{v\}$$

$$Unbiased: \mathrm{E\{R(x_0)\}} = 0 = E\{v\}\sum_{i=1}^{n} \lambda_i - E\{v\}$$

$$\Rightarrow E\{v\}\sum_{i=1}^{n} \lambda_i = E\{v\}$$

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

OK – Minimize var{R(x_0)}

$$\mathrm{var}(aX \pm bY) = a^2\,\mathrm{var}(X) + b^2\,\mathrm{var}(Y) \pm 2ab\,\mathrm{cov}(X,Y)$$

$$\mathrm{var}\{R(x_0)\} = \mathrm{var}\{\hat{v}(x_0) - v(x_0)\}$$

$$= \mathrm{var}\{\sum_{i=1}^{n} \lambda_i v(x_i)\} + \mathrm{var}\{v(x_0)\} - 2\,\mathrm{cov}\{\hat{v}(x_0), v(x_0)\}$$

$$= 2\sum_{i=1}^{n} \lambda_i \gamma(x_i, x_0) - \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j \gamma(x_i, x_j)$$

The minimization of a function of n variables is usually by setting the n partial first derivatives to 0. Setting the derivative to 0, however we have a constraint: $\sum_{i=1}^{n} \lambda_i = 1$

OK – Lagrange parameter

A procedure for converting a constrained minimization problem into an unconstrained one.

We introduce an unknown into our equation for var{R(x_0)}: Ψ

$$f\{\lambda_i, \psi(\mathrm{x}_0)\}$$

$$= \mathrm{var}\{R(x_0)\} - 2\psi(x_0)[(\sum_{i=1}^{n} \lambda_i) - 1]$$

$$f(\lambda_i, \psi) = 2\sum_{i=1}^{n} \lambda_i \gamma(x_i, x_0) - \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j \gamma(x_i, x_j) - 2\psi(x_0)[(\sum_{j=1}^{n} \lambda_j) - 1]$$

OK – Minimize f{$\lambda_{i,}$ ψ}

$$\frac{\partial f\{\lambda_i,\psi(\mathrm{x}_0)\}}{\partial\lambda_i}=2\gamma(x_i,x_0)-2\sum_{j=1}^{n}\lambda_j\gamma(x_i,x_j)-2\psi(\mathrm{x}_0)=0$$

$$\Rightarrow\sum_{j=1}^{n}\lambda_j\gamma(x_i,x_j)+\psi(\mathrm{x}_0)=\gamma(x_i,x_0)$$

$$\frac{\partial f\{\lambda_i,\psi(\mathrm{x}_0)\}}{\partial\psi}=-2[(\sum_{j=1}^{n}\lambda_j)-1]$$

$$\Rightarrow\sum_{j=1}^{n}\lambda_j=1$$

$$\textit{matrix}\ \mathrm{form}:\vec{\mathrm{A}}\vec{\lambda}=\vec{b}$$

$$\begin{bmatrix}\gamma(x_1,x_1) & \gamma(x_1,x_2) & \gamma(x_1,x_3)\ldots & \gamma(x_1,x_n) & 1\\ \gamma(x_2,x_1) & \gamma(x_2,x_2) & \gamma(x_2,x_3)\ldots & \gamma(x_1,x_n) & 1\\ \gamma(x_3,x_1) & \gamma(x_3,x_2) & \gamma(x_3,x_3)\ldots & \gamma(x_3,x_n) & 1\\ \ldots & & & & \\ \gamma(x_n,x_1) & \gamma(x_n,x_2) & \gamma(x_n,x_3)\ldots & \gamma(x_n,x_n) & 1\\ 1 & 1 & 1 & 1 & 0\end{bmatrix}\begin{bmatrix}\lambda_1\\ \lambda_2\\ \lambda_3\\ \ldots\\ \lambda_n\\ \psi(x_0)\end{bmatrix}=\begin{bmatrix}\gamma(x_1,x_0)\\ \gamma(x_2,x_0)\\ \gamma(x_3,x_0)\\ \ldots\\ \gamma(x_n,x_0)\\ 1\end{bmatrix}\Rightarrow\lambda=A^{-1}b$$

Kriging variance

$$\mathrm{var}\{R(x_0)\}=2\sum_{i=1}^{n}\lambda_i\gamma(x_i,x_0)-\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j\gamma(x_i,x_j)$$

$$=2\sum_{i=1}^{n}\lambda_i(\sum_{j=1}^{n}\lambda_j\gamma(x_i,x_j)+\psi(x_0))-\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j\gamma(x_i,x_j)$$

$$=\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j\gamma(x_i,x_j)+2\psi(x_0)$$

$$=\sum_{i=1}^{n}\lambda_i(\sum_{j=1}^{n}\lambda_j\gamma(x_i,x_j)+\psi(x_0))+\psi(x_0)$$

$$=\sum_{i=1}^{n}\lambda_i\gamma(x_i,x_0)+\psi(x_0)$$

$$=b^T\lambda$$

Simple example

$$\gamma(\vec{h}) = 25 + 80Sph(0.35) + 15Sph(3.00)$$

An intuitive way to look at OK

1. The choice of a semivariogram model is prerequisite for OK. More time consuming but more flexible. It could also incorporate valuable qualitative insights such as the pattern of anisotropy.

   Again, why model

2. b matrix – provides a weighting scheme similar to that of the inverse distance methods $|h|^{-p}$. The semivarances calculated for our model can come from a much larger family of functions. Statistical distance

3. A matrix – Provides information on the clustering of the available sample data. Statistical distance

OK – some characteristics

1. Ordinary Kriging is exact: Estimates at sampling location = observation
2. $\hat{z}(x_0) \approx \bar{z}$ when x_0 is far away from all sampling points
3. Interpolation is smooth