Anniversary College on Soil Physics

Lecture Notes

Optimal Estimations of Random Fields using Kriging

by
Gautam Barua
Department of Civil Engineering
Indian Institute of Technology, Guwahati
India

Date: 11th March 2003 **Time: 14.00-14.50**

Kriging is a statistical procedure of estimating the 'best weights' of a linear estimator. Suppose there is a point or an area or a volume of ground over which we do not know a hydrological variable and wish to estimate it. In order to produce an estimator, we need some information to work on, usually available in the form of samples. Let there be a set S of n samples with values $g_1, g_2, g_3, \dots g_n$. The estimator, T^*, can then be calculated as

$$T^* = w_1 g_1 + w_2 g_2 + w_3 g_3 + \dots + w_n g_n \tag{1}$$

where $w_1, w_2, w_3 \dots w_n$ are the weights assigned to each sample. If all the weights are equal and sum up to one, T^* is just the arithmetic mean of the sample values.

The estimation variance for the general 'unbiased linear' estimator can be determined as

$$\sigma_{\in}^2 = 2\sum_{i=1}^{n} w_i \bar{\gamma}(S_i, A) - \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j \bar{\gamma}(S_i, S_j) - \bar{\gamma}(A, A) \tag{2}$$

where $w_i \bar{\gamma}(S_i, A)$ is the weighted average semi-variogram value between each point in the 'sample set' S and each point in the panel A, $w_i w_j \bar{\gamma}(S_i, S_j)$ is the weighted average semi-variogram value between each point in the sample set and each point in the sample set, and $\bar{\gamma}(A, A)$ is the average semi-variogram value between each point in the panel and each point in the panel.

Now for the linear estimator to be unbiased, we require that

$$E(T^*) = E(T) = m \text{ (say)} \tag{3}$$

$$E\left[\sum_{i=1}^{n} w_i g_i\right] = E[g_i]\sum_{i=1}^{n} w_i = m\sum_{i=1}^{n} w_i = m \tag{4}$$

Thus

$$\sum_{i=1}^{n} w_i = 1 \tag{5}$$

There can, however, be infinite number of linear unbiased estimators for which the weights sum upto one. The problem is how to determine the 'best' weights for which the estimation variance is the least? This can be done by minimizing σ_{ϵ}^{2} subject to equation (5). This calls for the introduction of a Langrange's multiplier, λ, in the minimizing function $F(w_1, w_2, w_3, \ldots w_n, \lambda)$, say, such that

$$F(w_1, w_2, w_3 \ldots w_n, \lambda) = 2\sum_{i=1}^{n} w_i \bar{\gamma}(S_i, A) - \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j \bar{\gamma}(S_i, S_j) - \bar{\gamma}(A, A) - \lambda\left(\sum_{i=1}^{n} w_i - 1\right) \tag{6}$$

Minimizing (6) with respect to $w_1, w_2, w_3, \ldots w_n$ and λ, we obtain the following set of n+1 equations

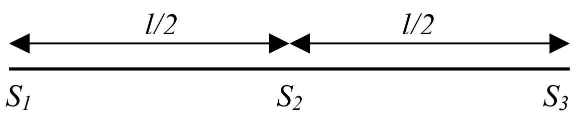
$$
\begin{aligned}
&w_1\bar{\gamma}(S_1,S_1)+w_2\bar{\gamma}(S_1,S_2)+w_3\bar{\gamma}(S_1,S_3)+\ldots+w_n\bar{\gamma}(S_1,S_n)+\lambda=\bar{\gamma}(S_1,A)\\
&w_1\bar{\gamma}(S_2,S_1)+w_2\bar{\gamma}(S_2,S_2)+w_3\bar{\gamma}(S_2,S_3)+\ldots+w_n\bar{\gamma}(S_2,S_n)+\lambda=\bar{\gamma}(S_2,A)\\
&w_1\bar{\gamma}(S_3,S_1)+w_2\bar{\gamma}(S_3,S_2)+w_3\bar{\gamma}(S_3,S_3)+\ldots+w_n\bar{\gamma}(S_3,S_n)+\lambda=\bar{\gamma}(S_3,A)\\
&---\quad ---\quad ---\quad ---\quad ---\quad ---\quad ---\\
&---\quad ---\quad ---\quad ---\quad ---\quad ---\quad ---\\
&w_1\bar{\gamma}(S_n,S_1)+w_2\bar{\gamma}(S_n,S_2)+w_3\bar{\gamma}(S_n,S_3)+\ldots+w_n\bar{\gamma}(S_n,S_n)+\lambda=\bar{\gamma}(S_n,A)\\
&w_1+w_2+w_3+\ldots\ldots\ \ldots+w_n=1
\end{aligned}
\tag{7}
$$

The system of equations as shown above is generally known as the kriging system and the estimator produced is the kriging estimator. The variance of the kriging estimator can be found by substitution of the weights in the general estimation variance equation (2). However, it can be shown that the kriging variance, σ_k^2, can be expressed as

$$
\sigma_k^2=\sum_{i=1}^{n} w_i\bar{\gamma}(S_i,A)\ +\lambda-\bar{\gamma}(A,A) \tag{8}
$$

Example

Let us now show an application of the kriging equations to a one-dimension problem. Let there be a line segment of length *l* and let it contain three samples, S_1, S_2, S_3, one at the beginning, the other at the middle, and the last at the end of the line as shown in the figure below



We assume here a linear model for the semi-variogram of the form $\gamma(h) = ph$, where h is the separation and p is the slope of the semi-variogram. For the model considered, $\lambda(A, A)$ works out to be *pl/3* and $\lambda(S, A)$ becomes equal to *pl/2*. Applying the model to equation (7), we obtain the following set of kriging equations

$$
\begin{aligned}
& w_2 pl/2 + w_3 pl + \lambda = pl/2 \\
& w_1 pl/2 + w_3 pl/2 + \lambda = pl/4 \\
& w_2 pl/2 + w_3 pl + \lambda = pl/2 \\
& w_1 + w_2 + w_3 = 1
\end{aligned}
$$

By solving these equations, we obtain $\lambda = 0$, $w_1 = w_3 = 1/4$ and $w_2 = 1/2$. Substituting these values in (8), we obtain the kriging variance, σ_k^2, to be $pl/24$. If we evaluate the estimation variance by treating all the weights to be same (arithmetic mean), we find σ_ϵ^2 to be $pl/18$. Thus, for the one-dimensional problem considered, kriging definitely gives a better estimation variance than the extension variance.

References:

These lecture notes have been compiled from the following texts.

1. Clark, I., 2001. Practical Geostatistics. Geostokos Limited, Alloa Business Centre, Whins Road, Alloa, Scotland. This text is freely available in the net.

2. Gelhar, LW, 1993. Stochastic Subsurface Hydrology, Prentice Hall, Englewood Cliffs, New Jersey.

****