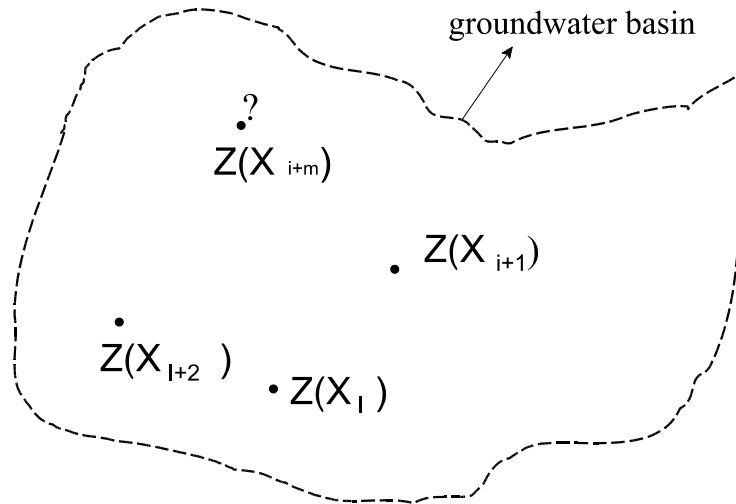


L-9. GEOSTATISTICS



Geostatistics is a parameter estimation technique that allows one to estimate, for example, the water level at locations where no measurement is available, using water levels at adjacent observation wells.

Suppose the hydraulic head of a groundwater basin at a given time or under a steady state flow condition is represented by a field variable, $Z(x)$. We want to estimate the piezometric surface $Z(x_{i+m})$ at locations x_{i+m} using water levels, $Z(x_i)$, $Z(x_{i+1})$, ... and $Z(x_{i+m-1})$ measured at locations x_i , x_{i+1} , ... and x_{i+m-1} . Notice that i is a location index and m is the total number of sample locations. A mathematical statement for this problem becomes

$$\hat{Z}(x_{i+m}) = F[Z(x_i), Z(x_{i+1}), \dots, Z(x_{i+m-1})]$$

where $\hat{}$ denotes the estimate which is a function of the measured values at the sample locations, x_i , x_{i+1} , ... and x_{i+m-1} . The problem then becomes what the function, F , is. Hydrologists often manually construct a contour map of a groundwater piezometric surface, based on water level data at several wells. By doing this, hydrologists essentially define a function for the problem, based on their intuition and experience. The function defined by them thus is subjective and qualitative and one cannot assess the uncertainty or probability that the contoured piezometric head field is correct or incorrect. On the other hand, geostatistics is a mathematical tool that defines this function in a quantitative manner that allows us to quantify the uncertainty associated with the estimate.

Before we develop the function, we have to ask ourselves a question: what is the objective of the function or what do we want the function do? Since we cannot expect the function to provide an exact estimate of the water level, the only hope we have is to develop a function that gives us a "best" estimate at the point where there is no sample. What do we mean by a "best estimate"? To define it, one needs a "criterion" to measure the accuracy of $Z(x_i)$ as a predictor of $Z(x_{i+m})$. The following are the criteria used in geostatistics:

Bias

$$B = E[\hat{Z}] - E[Z] = E[\hat{Z} - Z] \quad (2)$$

In other words, if we estimate Z at a location many times, the bias is defined as a measure of the difference between the average of the estimates and the mean of the random variable Z , the property at the location. A good estimator should possess small bias. If the estimator yields zero difference, it is called an unbiased estimator.

Variance of an estimator

In addition to the minimal bias, one would also like to have an estimator that has a small variance. That is, we want an estimator that produces a very small variation between estimates since the estimator may give different estimates. Mathematically, that is,

$$\text{var}[\hat{Z}] = E[(\hat{Z} - E[\hat{Z}])^2] \quad (3)$$

Mean Square Error (MSE)

Mean square error is defined, mathematically, as the expected value of the square of the difference between our estimate and the true value:

$$\begin{aligned} E[(\hat{Z} - Z)^2] &= E[\{(\hat{Z} - E(\hat{Z})) + (E(\hat{Z}) - Z)\}^2] \\ &= E[(\hat{Z} - E(\hat{Z}))^2] + E[E(\hat{Z}) - Z]^2 + 2E[(E(\hat{Z}) - Z)E(\hat{Z} - E(\hat{Z}))] \\ &= \text{var}[\hat{Z}] + (E[\hat{Z} - Z])^2 \\ &= \text{var}[\hat{Z}] + B^2 \end{aligned} \quad (4)$$

It is essentially the sum of the variance of our estimate and the square of the bias of our estimator. Of course, we want an estimator that has minimal bias and variance. Therefore, the criterion of an estimator is that it must have a minimal mean square error, i.e.,

$$E[(\hat{Z} - Z)^2] = \text{minimum}$$

Now, the objective for the function has been defined. However, recall that in addition to the minimal mean square error requirement, during the estimation the function must use information that we already have. That is,

$$\hat{Z}(x_{i+m}) = F[Z(x_i), Z(x_{i+1}), \dots, Z(x_{i+m})] \quad (5)$$

These two requirements lead us to the conditional expectation operator in statistics.

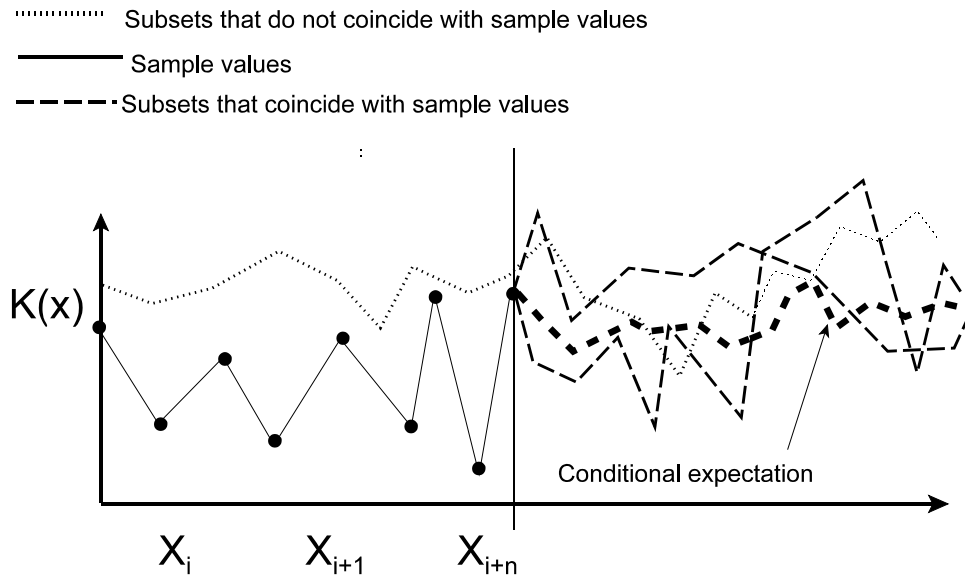
Conditional Expectation

Conditional expectation can be written in a mathematic format as:

$$E[Z(x_{i+m})|Z(x_i), Z(x_{i+1}), \dots, Z(x_{i+m})] \quad (6)$$

It is the expected value of Z under the circumstance that if we have already known $Z(x_i)$, $Z(x_{i+1})$, ... and $Z(x_{i+m-1})$. The concept of the conditional expectation is illustrated in the figure below. If we assume that the hydraulic conductivity is a stochastic process, there are an infinite number of possible realizations of hydraulic conductivity values (ensemble) along a transect that has the same variance and correlation structure. It is also expected that within this ensemble there are realizations (a subset of the ensemble) that will agree with the conductivity values at the sample locations. On the other hand, we also expect that there are many realizations (a subset of the ensemble) that do not agree with the sample values. Suppose we discard the subset that does not include the measurements. We then have a subset where the conductivity value will coincide with the measurements at sample locations and will be a stochastic variable (many possible values) at unsampled locations. The conditional expectation simply means that we take the ensemble average of this subset. As a result, this average yields a conductivity field that agrees with measurements at sample locations but is the average of many possible conductivity values at the nonsampled locations of the subset.

How can the concept of the conditional expectation translate into mathematics?



If Z is multivariate normal (jointly normal), the conditional expectation is a linear sum of all the weighted sample values (See p. 728, Priestley, 1981, Spectral Analysis and Time Series).

$$\begin{aligned}
 Z(x_{i+m}) &= \sum_i \lambda_i Z(x_i) \\
 &= \lambda_1 Z(x_1) + \lambda_2 Z(x_2) + \dots + \lambda_m Z(x_m)
 \end{aligned}
 \tag{7}$$

In other words, if the stochastic process has a normal joint probability distribution, we can obtain the conditional expectation by using a linear sum of weighted measurements of conductivity at the sampled locations. This basically defines the function that we are looking for and this function is often called the linear predictor.

Although we have selected the linear predictor and we know the values of Z at sample locations, we still have to determine the values of λ_i so that we can use the linear predictor to estimate the unknown. Determining the weights and the estimate becomes the task of kriging. In this class, we will focus on the discussion of simple kriging and ordinary kriging.

Kriging in Stationary Case. (Simple Kriging)

Recall $Z(x)$ is a second-order stationary process if

- 1) $E[Z(x)] = m$ (constant mean)
- 2) the covariance of Z depends on the separation distance only, i.e.,

$$\text{cov}(x_i, x_j) = E[(Z(x_i) - m)(Z(x_j) - m)] = c(h) \quad (8)$$

$$\begin{aligned} c(h) &= E[Z(x_i)Z(x_j)] - mE[Z(x_i)] - mE[Z(x_j)] + m^2 \\ &= E[Z(x_i)Z(x_j)] - m^2 \end{aligned} \quad (9)$$

since $E[Z(x_i)] = m$. Now we define a new process with mean zero:

$$Y(x) = Z(x) - m \quad (10)$$

and

$$E[Y(x)] = 0 \quad (11)$$

Then, the estimation of Y at the point x_0 (if $Y_i = Y(x_i)$, $i = 1 \dots n$ at locations x_i which are known) becomes:

$$\hat{Y}_0 = \hat{Y}_0(x_0) = \lambda_{01}Y_1 + \lambda_{02}Y_2 + \dots + \lambda_{0n}Y_n \quad (12)$$

or

$$\hat{Y}_0 = \sum_{i=1}^n \lambda_{0i}Y_i \quad (13)$$

where \hat{Y}_0 is the estimate of Y at location x_0 , Y_i represents the observation of Y at sample locations x_i , and λ_{0i} is the weight associated with Y_0 and Y_i . It represents the contribution of the observed Y value at location x_i to the estimate at location x_0 . As discussed before, an optimal estimator requires that the error of estimation be minimal.

$$E[\hat{Y}_0 - Y_0]^2 = \text{minimum}$$

\hat{Y}_0 and Y_0 are random variables. We define a new random variable, $U = (\hat{Y}_0 - Y_0)$. The above expression says that we are seeking the minimum of the variance of U . Before the minimization, we will evaluate the variance:

$$\begin{aligned} E[(\hat{Y}_0 - Y_0)^2] &= E\left[\left(\sum_{i=1}^n \lambda_{0i} Y_i - Y_0\right)^2\right] \\ &= E\left[\left(\sum_{i=1}^n \lambda_{0i} Y_i\right)\left(\sum_{j=1}^n \lambda_{0j} Y_j\right)\right] - 2E\left[\sum_{i=1}^n \lambda_{0i} Y_i Y_0\right] + E[(Y_0)^2] \\ &= \sum_{i=1}^n \sum_{j=1}^n \lambda_{0i} \lambda_{0j} E[Y_i, Y_j] - 2 \sum_{i=1}^n \lambda_{0i} E[Y_i, Y_0] + E(Y_0^2) \end{aligned} \quad (14)$$

Below gives an example showing how the product of two sums can be written in the double sums as used in the above equation.

$$\begin{aligned} &E[(\lambda_{01} Y_1 + \lambda_{02} Y_2)(\lambda_{01} Y_1 + \lambda_{02} Y_2)] \\ &= \lambda_{01} \lambda_{01} E[Y_1 Y_1] + \lambda_{01} \lambda_{02} E[Y_1 Y_2] + \lambda_{02} \lambda_{01} E[Y_2 Y_1] + \lambda_{02} \lambda_{02} E[Y_2 Y_2] \\ &= \sum_{i=1}^2 \sum_{j=1}^2 \lambda_{0i} \lambda_{0j} E[Y_i Y_j] \end{aligned} \quad (15)$$

Use the fact that

$$E(Y_i Y_j) = C(x_i - x_j) \quad (16)$$

and $E[Y_i Y_i] = C(0) = \sigma_y^2$. Thus, one can express the variance as:

$$E[(\hat{Y} - Y_0)^2] = \sum_{i=1}^n \sum_{j=1}^n \lambda_{0i} \lambda_{0j} C(x_i - x_j) - 2 \sum_{i=1}^n \lambda_{0i} C(x_i - x_0) + C(0) \quad (17)$$

To find the values of λ_{0i} that minimize the variance, we differentiate the variance with respect to λ_{0i} and set the resultant equal to zero. That is,

$$\frac{\partial}{\partial \lambda_{0i}} E[(\hat{Y}_0 - Y_0)^2] = 0 \quad (18)$$

The result is

$$2 \sum_j \lambda_{0j} C(x_i - x_j) - 2C(x_i - x_0) = 0 \quad (19)$$

or

$$\sum_j \lambda_{0j} C(x_i - x_j) = C(x_i - x_0) \quad (20)$$

for a given subscript i . Therefore, we have a system of n equations in a matrix form:

$$\begin{bmatrix} c(x_1 - x_1) & c(x_1 - x_2) & \dots & c(x_1 - x_n) \\ \vdots & \vdots & \ddots & \vdots \\ c(x_n - x_1) & c(x_n - x_2) & \dots & c(x_n - x_n) \end{bmatrix} \begin{bmatrix} \lambda_{01} \\ \lambda_{02} \\ \vdots \\ \lambda_{0n} \end{bmatrix} = \begin{bmatrix} c(x_1 - x_0) \\ c(x_2 - x_0) \\ \vdots \\ c(x_n - x_0) \end{bmatrix} \quad (21)$$

We then solve for λ_{0i} , $i = 1, \dots, n$, which are the weights to be used in the linear predictor. Afterwards, our estimate can be obtained by using

$$\hat{Y}_0 = \sum_{i=1}^n \lambda_{0i} Y_i \quad (22)$$

Now check if the estimate derived satisfies the minimum bias and variance criterion we specified previously. Bias:

$$\begin{aligned} E(\hat{Y}_0 - Y_0) &= E[Y_0] \\ &= E\left[\sum \lambda_{0i} Y_i\right] - E[Y_0] \\ &= \sum \lambda_{0i} E[Y_i] - E[Y_0] = 0 \end{aligned} \quad (23)$$

Therefore, we can say our estimator is an unbiased estimator. We also can check the variance.

Kriging Variance:

$$\begin{aligned}\text{var}(\hat{Y}_0 - Y_0) &= E\left[\left(\hat{Y}_0 - Y_0\right)^2\right] - E\left[\left(\hat{Y}_0 - Y_0\right)\right]^2 \\ &= E\left[\left(\hat{Y}_0 - Y_0\right)^2\right] - 0\end{aligned}\quad (24)$$

= minimal

From (20) we know

$$\sum_i \sum_j \lambda_{0i} \lambda_{0j} C(x_i - x_j) = \sum_i \lambda_{0i} C(x_i - x_0)$$

From (17), the kriging variance becomes

$$\text{var}(\hat{Y}_0 - Y_0) = E\left[\left(\hat{Y}_0 - Y_0\right)^2\right] = C(0) - \sum_i \lambda_{0i} (x_i - x_0)$$

The kriging variance is smaller than the variance of y process itself because of the measurements at x_i . Also, notice that at the measurement location, the kriging variance is zero.

Now, we can change \hat{Y}_0 back to Z . Recall that $Y(x) = Z(x) - m$, then

$$\hat{Z}_0 = m + \sum_i \lambda_{0i} (Z_i - m)$$

This is our best, linear, and unbiased estimate of Z at x_0 . It should be pointed out that the estimate is simply a conditional expectation which does not provide us the exact value of Z but the most likely value for Z at the location.

Kriging Covariance

$$\begin{aligned}E[(\hat{Y}_p - Y_p)(\hat{Y}_q - Y_q)] &= E\left[\left(\sum_{i=1}^n \lambda_{pi} Y_i - Y_p\right)\left(\sum_{i=1}^n \lambda_{qi} Y_i - Y_q\right)\right] \\ &= E\left[\left(\sum_{i=1}^n \lambda_{pi} Y_i\right)\left(\sum_{j=1}^n \lambda_{qj} Y_j\right)\right] - E\left[\sum_{i=1}^n \lambda_{pi} Y_i Y_q\right] - E\left[\sum_{i=1}^n \lambda_{qi} Y_i Y_p\right] + E[Y_p Y_q] \\ &= \sum_{i=1}^n \sum_{j=1}^n \lambda_{pi} \lambda_{qj} E[Y_i, Y_j] - \sum_{i=1}^n \lambda_{pi} E[Y_i, Y_q] - \sum_{i=1}^n \lambda_{qi} E[Y_i, Y_p] + E(Y_p Y_q)\end{aligned}\quad (26)$$

The Intrinsic Hypothesis

It is a less stringent hypothesis than the second-order stationary assumption. It assumes that even if the variance of Z is not finite, the variance of the first-order increments of Z is finite and the increments are second-order stationary, i.e., that $[Z(x+h) - Z(x)]$ satisfies

- 1) $E [Z(x+h) - Z(x)] = m(h)$
- 2) $\text{var} [Z(x+h) - Z(x)] = 2\gamma(h)$

In other words, the difference between Z s at different locations will have a constant mean and variance as long as the Z s are separated by the same distance h . The mean and the variance of the difference are functions of h , separation distance, not x , and

$$\gamma(h) = \frac{1}{2} \text{VAR}[Z(x+h) - Z(x)] \quad \text{variogram (semi-variogram)}$$

The properties of the variogram are: 1) the value of the variogram at the origin is zero, $\gamma(0)=0$; 2) the values of the variogram are positive, $\gamma(h)\geq 0$, and 3) the variogram is an even function, $\gamma(h)=\gamma(-h)$.

Now, if we assume

$$E[Z(x+h) - Z(x)] = m(h) = 0$$

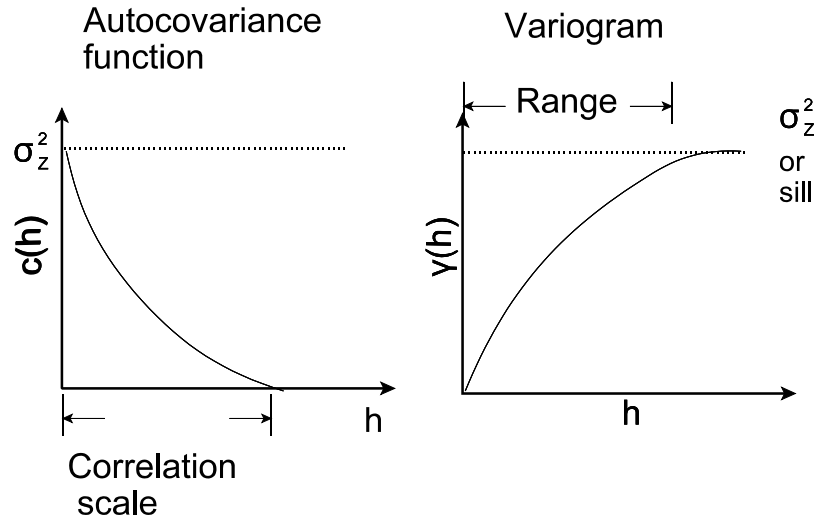
then

$$\begin{aligned} \gamma(h) &= \frac{1}{2} E \left\{ [Z(x+h) - Z(x)]^2 \right\} \\ &= \frac{1}{2} \left\{ E[Z(x+h)^2] - 2E[Z(x+h)Z(x)] + E[Z(x)^2] \right\} \end{aligned}$$

If Z is a second-order stationary process, the variogram can be related to the covariance.

$$\gamma(h) = C(0) - C(h)$$

The variogram, in this case, is a mirror image of the autocovariance function as illustrated in the figure below.



Kriging with Intrinsic Hypothesis (Ordinary Kriging)

The estimation criteria for this case are the same as those in the simple kriging:

- (1) $E(\hat{Z}_0 - Z_0) = 0$
- (2) $E[(\hat{Z}_0 - Z_0)^2] = \text{minimum}$

In the intrinsic hypothesis, we know the mean of the increment is a function of separation distance. However, we do not know the value of the mean of $Z(x)$, $N(x)$, in this case. From the first criterion, our estimator must have

$$E[\hat{Z}_0] = E[Z_0]$$

If we chose the linear estimator,

$$\hat{Z}_0 = \sum_{i=1}^n \lambda_{0i} Z_i$$

we have

$$E\left[\sum_{i=1}^n \lambda_{0i} Z_i\right] = E[Z_0] = N(x)$$

and

$$\sum_{i=1}^n \lambda_{0i} E(Z_i) = N$$

Therefore,

$$\sum_{i=1}^n \lambda_{0i} = 1$$

In essence, this is another condition that we must impose on the values of λ in addition to the condition that they must minimize the variance specified in (2).

To optimize the variance, we must derive the variance first. The variance can be derived as follows:

$$\begin{aligned} E\left[(\hat{Z}_0 - Z_0)^2\right] &= E\left[\left(\sum \lambda_{0i} Z_i - Z_0\right)^2\right] \\ &= E\left[\left(\sum \lambda_{0i} Z_i - \sum \lambda_{0i} Z_0\right)^2\right] \\ &= E\left[\left(\sum_i \lambda_{0i} (Z_i - Z_0)\right)^2\right] \\ &= E\left[\sum_i \lambda_{0i} (Z_i - Z_0) \sum_i \lambda_{0j} (Z_j - Z_0)\right] \\ &= \sum_i \sum_j \lambda_{0i} \lambda_{0j} E[(Z_i - Z_0)(Z_j - Z_0)] \end{aligned}$$

In the above equation, the term $E[(Z_i - Z_0)(Z_j - Z_0)]$ on the right-hand side of the equation can be related to the variogram defined earlier.

$$\begin{aligned}
\gamma(x_i - x_j) &= \frac{1}{2} E[(Z_i - Z_j)^2] \\
&= \frac{1}{2} E[(Z_i - Z_0) - (Z_j - Z_0)]^2 \\
&= \frac{1}{2} E[(Z_i - Z_0)^2] + \frac{1}{2} E[(Z_j - Z_0)^2] \\
&\quad - E[(Z_i - Z_0)(Z_j - Z_0)] \\
&= \gamma(x_i - x_0) + \gamma(x_j - x_0) - E[(Z_i - Z_0)(Z_j - Z_0)]
\end{aligned}$$

Therefore,

$$E[(Z_i - Z_0)(Z_j - Z_0)] = \gamma(x_i - x_0) + \gamma(x_j - x_0) - \gamma(x_i - x_j)$$

Thus,

$$\begin{aligned}
E[(\hat{Z}_0 - Z_0)^2] &= \sum \sum \lambda_{0i} \lambda_{0j} \gamma(x_i - x_0) \\
&\quad + \sum \sum \lambda_{0i} \lambda_{0j} \gamma(x_j - x_0) - \sum \sum \lambda_{0i} \lambda_{0j} \gamma(x_i - x_j)
\end{aligned}$$

The above expression can be simplified based on the following example. If we let i and j be 1 and 2, the first two terms of the right-hand-side can be rewritten as:

$$\begin{aligned}
&= \lambda_{01} \lambda_{01} \gamma(x_1 - x_0) + \lambda_{01} \lambda_{02} \gamma(x_1 - x_0) + \lambda_{02} \lambda_{01} \gamma(x_2 - x_0) + \lambda_{02} \lambda_{02} \gamma(x_2 - x_0) \\
&\quad + \lambda_{01} \lambda_{01} \gamma(x_1 - x_0) + \dots \\
&= 2[\lambda_{01} \lambda_{01} \gamma(x_1 - x_0) + \lambda_{01} \lambda_{02} \gamma(x_1 - x_0) + \lambda_{02} \lambda_{01} \gamma(x_2 - x_0) + \lambda_{02} \lambda_{02} \gamma(x_2 - x_0)] \\
&= 2[(\lambda_{01} + \lambda_{02}) \lambda_{01} \gamma(x_1 - x_0) + (\lambda_{01} + \lambda_{02}) \lambda_{02} \gamma(x_2 - x_0)] \\
&= 2[(\lambda_{01} + \lambda_{02})(\lambda_{01} \gamma(x_1 - x_0) + \lambda_{02} \gamma(x_2 - x_0))] \\
&= 2 \sum \lambda_{0j} \gamma(x_j - x_0)
\end{aligned}$$

Therefore, we can write

$$E[(\hat{Z}_0 - Z_0)^2] = 2 \sum \lambda_{0j} \gamma(x_j - x_0) - \sum \sum \lambda_{0i} \lambda_{0j} \gamma(x_i - x_j)$$

Now, we want to minimize this expression with the constraint,

$$\sum \lambda_{0j} = 1$$

Minimizing with a constraint, we must use the Lagrange multipliers (see your calculus book) and let

$$H = E\left[(\hat{Z}_0 - Z_0)^2\right] - \mu \left[\sum_j \lambda_{0j} - 1 \right]$$

where μ is the Lagrange multiplier. Taking the derivative of H with respect to λ s and μ and set the derivatives to zero. That is,

$$\frac{\partial H}{\partial \lambda_{0i}} = 0 \quad \text{and} \quad \frac{\partial H}{\partial \mu} = 0$$

We obtain the following system of equations:

$$\begin{aligned} \sum \lambda_{0i} \gamma(x_i - x_j) + \mu &= \gamma(x_i - x_0) & i = 1, \dots, n \\ \sum \lambda_{0i} &= 1 \end{aligned}$$

Expressing them in matrix form, we have

$$\begin{bmatrix} 0 & \gamma_{12} & \cdot & \cdot & \cdot & \gamma_{1n} & 1 \\ \gamma_{21} & & & & & & 1 \\ \cdot & & & & & & 1 \\ \cdot & & & & & & 1 \\ \cdot & & & & & & 1 \\ \gamma_{n1} & \cdot & \cdot & \cdot & \cdot & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_{01} \\ \lambda_{02} \\ \cdot \\ \cdot \\ \cdot \\ \lambda_{0n} \\ \mu \end{bmatrix} = \begin{bmatrix} \gamma_{01} \\ \gamma_{02} \\ \cdot \\ \cdot \\ \cdot \\ \gamma_{0n} \\ 1 \end{bmatrix}$$

Note that $\gamma_{ii} = \gamma(0) = 0$. Solve the system of equations for λ_{0i} and μ . Then we can estimate the unknown using the linear estimator. In addition, we can calculate the kriging variance which is

$$\text{var}(\hat{Z}_0 - Z_0) = \sum \lambda_{0i} \gamma(x_i - x_0) + \mu$$

Comments:

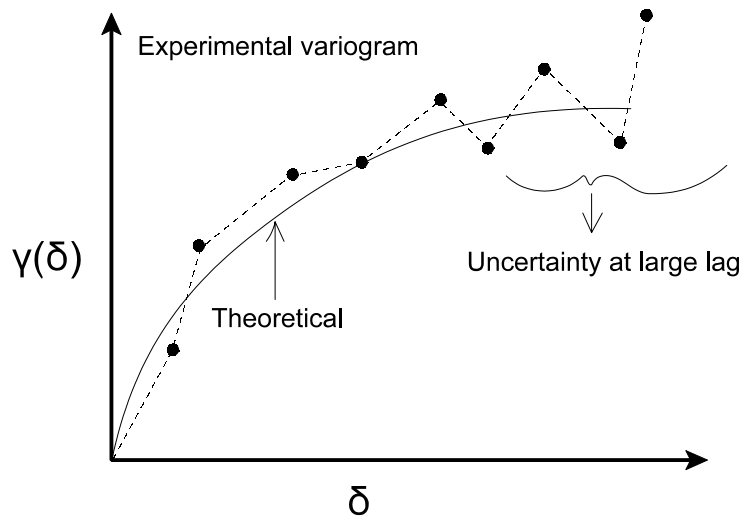
Kriging is often called BLUE, denoting the Best Linear Unbiased Estimator.

1. $B = E(\hat{Z}_0) - E(Z_0) = 0$ Unbiased
2. $\text{var}[\hat{Z}_0 - Z_0] = \min.$
3. No assumption about the type of distribution for $Z(x)$
4. Coincides with the conditional expectation if the joint normality condition is satisfied.
5. Unlike regression, kriging preserves data points.

§ Variogram

$$\gamma(\delta) = \frac{1}{2} E\{[Z(x + \delta) - Z(x)]^2\}$$

In general, variograms of hydraulic properties are estimated from observed data using the above formula or using software packages such Geo-EAS.

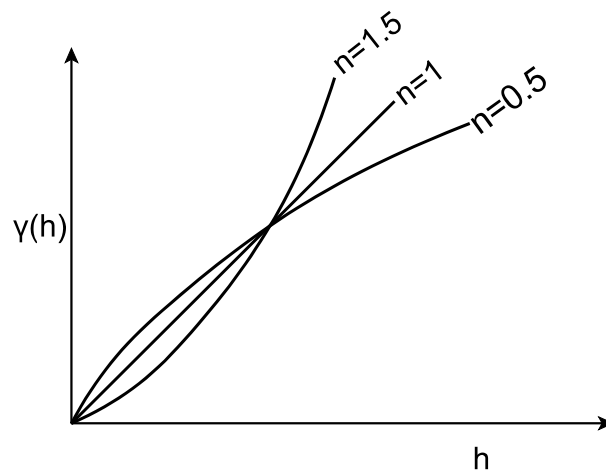


§ Theoretical Variogram

(positive definite function)

1. linear variogram model

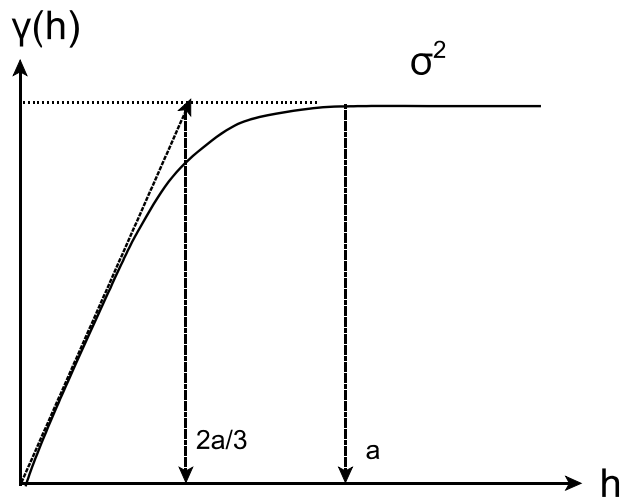
$$\gamma(h) = \sigma^2 |h|^n$$



2. Spherical variogram model:

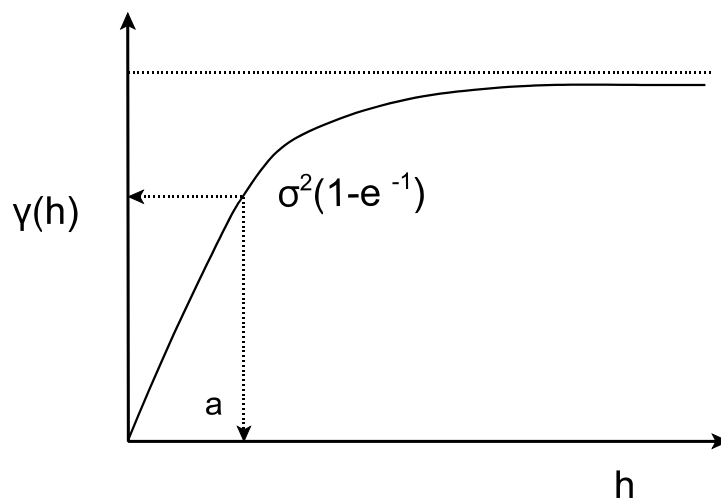
$$\gamma(h) = \sigma^2 \left[\frac{3}{2} \left(\frac{h}{a} \right) - \frac{1}{2} \left(\frac{h}{a} \right)^3 \right] \quad h < a$$

$$= \sigma^2 \quad h > a$$



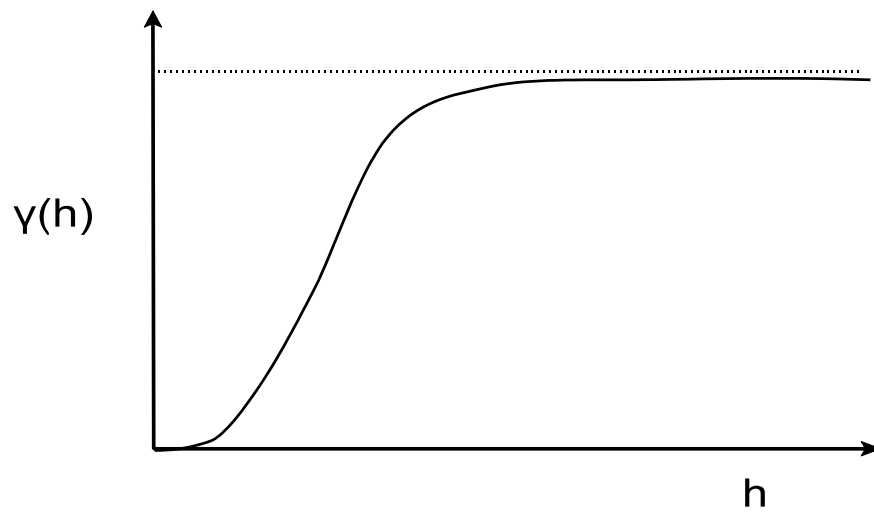
3. Exponential Model

$$\sigma^2 \left[1 - \exp\left(\frac{-|h|}{a}\right) \right]$$



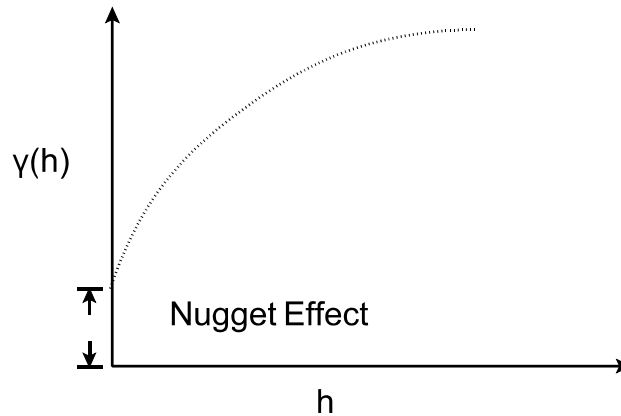
4. Gaussian Model

$$\sigma^2 \left(1 - \exp \left[- \left(\frac{h}{a} \right)^2 \right] \right)$$



There are many other models that will satisfy the positive definite requirement (de Marsily, 1986 and Isaaks and Srivastava, 1989).

Nugget Effect: The variogram is discontinuous near the origin.



- 1) measurement errors
- 2) Sampling interval is too large

$$\gamma(h) = c[1 - \delta(h)] + \gamma'(h)$$

where

$$\delta(h) = 1 \quad \text{if } h = 0 \quad \text{and} \quad = 0 \quad \text{if } h \neq 0$$

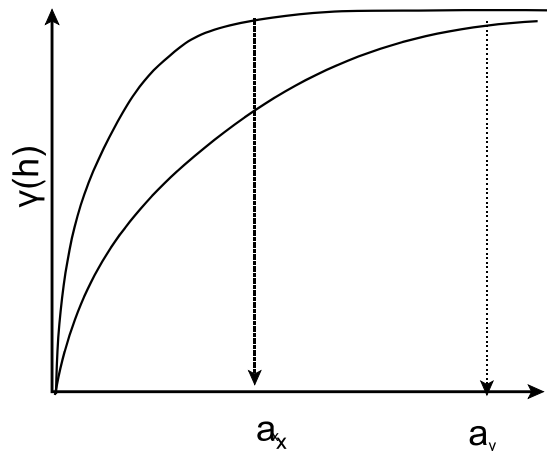
Anisotropic Variogram:

Similar to correlation scales, the variogram may have different ranges in different directions due to geological deposition processes. Again, it has been observed that most hydraulic conductivity data tend to show a longer range in the horizontal plane than in the vertical direction. This type of the variogram is called anisotropic variogram. Theoretically, the variograms in different directions should have the same sill but different ranges. For example,

Head \rightarrow longer range in the direction perpendicular to mean flow direction

→ small in the direction of mean flow.

Log Transformed hydraulic conductivity: The range in the horizontal direction is much larger than the range in the vertical direction—implications of layering structure of geological formation.



If the variograms in different directions have the same sill values but different ranges. This anisotropy is called the geometric anisotropy. Some field data, however, show that the variograms may have different sill values and the ranges values in different directions. The anisotropy is called the zonal anisotropy. This type anisotropy is mainly attributed to insufficient samples. That is, the sample area is small compared to the anisotropic structure of the geological formation. For example, if one intensively samples only over a 50 m x 50 m x 50m volume of a geological formation which has layers that extend horizontally over 500 m, the analysis of the samples likely yields a zonal anisotropy.

In general, T and K are assumed to have lognormal distributions. That is, $\ln T$ and $\ln K$ are normally distributed. As a result, the linear predictor (e.g., kriging) yields exact conditional expectation (mean).

Questions.

There are many theoretical variogram models. How does the accuracy of the model affect kriging?