

# Mathematical Development from Kriging to Semivariogram Modeling

Instructor: Jonnathan Frame

January 27, 2025

## 1 Kriging as Unbiased Linear Estimator

Consider the estimation of an unknown spatial variable  $z_0$  as a linear combination of observed spatial data  $z_i$  with weights  $\lambda_i$ :

$$Z^* = \lambda_0 + \sum_{i=1}^n \lambda_i z_i$$

where  $\lambda_0$  is a shift parameter (often zero if the estimator is unbiased), and  $\lambda_i$  are weights derived from the semivariogram.

## 2 Spatial Interpolation Using Kriging

For spatial interpolation, such as estimating porosity at an unsampled location  $x_0$  using observed data from surrounding locations  $x_i$ :

$$Z^*(x_0) = \sum_{i=1}^n \lambda_i (Z(x_i) - m) + m$$

where  $m$  is the mean of the observed values  $Z(x_i)$ , assumed known or estimated as the average of the  $Z(x_i)$ .

## 3 Modeling the Semivariogram

To determine the weights  $\lambda_i$  in the kriging formula, we model the semivariogram, which describes how spatial correlation changes with distance.

### 3.1 Linear Semivariogram

A simple model to begin with is the linear semivariogram:

$$\gamma(h) = \text{nugget} + \text{slope} \times h$$

where  $h$  is the lag or distance between observations, and *nugget* and *slope* are parameters that describe the variance at zero distance and the rate of increase in variance with distance, respectively.

### 3.2 Spherical Semivariogram Model

A more realistic model for many geostatistical applications is the spherical model, defined as:

$$\gamma(h) = \begin{cases} c_0 + c \left( \frac{3h}{2a} - \frac{h^3}{2a^3} \right) & \text{if } h \leq a \\ c_0 + c & \text{if } h > a \end{cases}$$

where:

- $c_0$  is the nugget effect,
- $c$  is the sill minus the nugget effect,
- $a$  is the range beyond which there is no correlation.

### 3.3 Exponential Semivariogram Model

The exponential semivariogram model is another common choice, particularly suitable for processes where correlation decreases exponentially with distance:

$$\gamma(h) = c_0 + c \left( 1 - e^{-\frac{h}{a}} \right)$$

where  $c_0$ ,  $c$ , and  $a$  have similar meanings as in the spherical model, with  $a$  representing the practical range at which the correlation is close to 0.37 of its original value.

### 3.4 Gaussian Semivariogram Model

For a smoother correlation decay, the Gaussian semivariogram model can be applied, which is characterized by a squared exponential decrease:

$$\gamma(h) = c_0 + c \left( 1 - e^{-\left(\frac{h}{a}\right)^2} \right)$$

This model assumes a faster decline in correlation near the origin compared to the exponential model, making it appropriate for highly continuous data.

## 4 Using the Semivariogram in Kriging

Once the semivariogram model is chosen and its parameters are estimated from the data, the model is used to calculate the weights  $\lambda_i$  for the kriging estimator. These weights minimize the estimation variance based on the modeled spatial correlations provided by the semivariogram.

When we transition from using kriging as an unbiased linear estimator to employing a semivariogram model, we are essentially shifting from a theoretical framework to a practical application. The kriging method, at its core, promises an unbiased linear estimation by assuming a known mean and spatial dependencies described by a semivariogram. However, in real-world applications, the semivariogram isn't directly observable and must be modeled from data.

The reason we model the semivariogram, rather than using empirical semivariogram values calculated from observed lags, is to ensure the estimator remains efficient and stable across different spatial scales and conditions. The semivariogram model provides a smoothed, consistent estimation of spatial dependencies, which helps in calculating the weights for the kriging equations more accurately and predictably. This model-based approach allows us to handle irregularities and variations in data, which might not be well-represented by a direct empirical calculation.

Moreover, using a modeled semivariogram still maintains the unbiasedness of the kriging estimator. The estimator is designed to minimize the mean squared error of the predictions, ensuring that it is the best linear unbiased predictor (BLUP). The semivariogram model, by providing a stable and smooth description of spatial variation (covariance structure), ensures that the calculated weights are optimal under the kriging framework. This approach avoids the pitfalls of overfitting to noisy or sparse data points, which can occur if the raw semivariogram values are used directly.