STANDARD PROCEDURE FOR UNIVERSAL KRIGING

Given a *universal kriging model*

$$Y(s) = x(s)'\beta + \varepsilon(s) \quad , \; s \in R$$

with (i) β unknown
(ii) $E[\varepsilon(s)] = 0$, $s \in R$
(iii) $\mathrm{cov}[\varepsilon(s), \varepsilon(v)]$ "known" for all $s, v \in R$

together with *observed data*, $y = (y_i : i = 1,..,n)'$ and $X = (x_1,...,x_n)'$, at points $\{s_i : i = 1,..,n\} \subset R$, the task is to predict the *unknown value*, $Y(s_0)$, at a *given location*, $s_0 \in R$, with *attributes*, $x_0 = x(s_0)$ [or possibly many such locations]. This is accomplished by variations of the following standard estimation procedure in software such as *Geostatistical Analyst*:

(1) Construct an *OLS estimate*, $\hat{\beta}_{OLS} = (X'X)^{-1}X'y$, of β and corresponding *residuals*, $\hat{\varepsilon}_{OLS} = y - X\hat{\beta}_{OLS}$.

(2) Use these residuals to estimate an *empirical variogram*, $\hat{\gamma}(h)$, at some set of selected distance values, $(h_i : i = 1,..,m)$.

(3) Use this data $(\hat{\gamma}_i, h_i), i = 1,..,m$ to fit (by nonlinear least squares) a *spherical variogram*, $\gamma(h\,;\hat{r},\hat{s},\hat{a})$.

(4) Then use the identity, $C(h) = \sigma^2 - \gamma(h)$, to construct the corresponding *spherical covariogram*, $\hat{C}(h) = \hat{s} - \gamma(h\,;\hat{r},\hat{s},\hat{a})$ for all distances h.

(5) If the distance between each pair of data points, s_i and s_j is denoted by h_{ij}, then the *covariance*, $\sigma_{ij} = \mathrm{cov}(\varepsilon_i, \varepsilon_j)$, between the residuals at s_i and s_j is estimated by $\hat{\sigma}_{ij} = \hat{C}(h_{ij})$ [where by definition, $\sigma_{ii} \equiv \sigma^2 \Rightarrow \hat{\sigma}_{ii} \equiv \hat{s}$], and the resulting covariance matrix between residuals at all data points $i = 1,..,n$ is given by

$$\hat{C} = \begin{pmatrix} \hat{\sigma}_{11} & \cdots & \hat{\sigma}_{1n} \\ \vdots & \ddots & \vdots \\ \hat{\sigma}_{n1} & \cdots & \hat{\sigma}_{nn} \end{pmatrix}$$

(6) Using this covariance matrix, one then obtains the *final estimate* of β by GLS as

$$\boxed{\hat{\beta} = (X'\hat{C}^{-1}X)^{-1}X'\hat{C}^{-1}y} \ .$$

(7) Note that *all* data points are used to estimate the covariogram, $\hat{C}(h)$. However, only a *subset* of these data points will be used to interpolate (krige) the value $Y(s_0)$ at s_0. In particular, if one selects an appropriate *kriging bandwidth*, h_0, about point s_0, then the simplest approach is to choose the set of all data points within distance h_0 of s_0, denoted say by $S_0 = \{s_i : i = 1,..,n_0\}$. [If the attribute vector is of dimension $k+1$ (including the intercept term) then h_0 is always chosen large enough to ensure that $n_0 \geq k+1$).

(8) The corresponding n_0-square submatrix of all covariances between residuals at points, $s_i, s_j \in S_0$, is then given by

$$\hat{C}_0 = \begin{pmatrix} \hat{\sigma}_{11} & \cdots & \hat{\sigma}_{1n_0} \\ \vdots & \ddots & \vdots \\ \hat{\sigma}_{n_0 1} & \cdots & \hat{\sigma}_{n_0 n_0} \end{pmatrix}$$

(9) In a similar manner, if the covariance between $\varepsilon(s_0)$ and each $\varepsilon(s_i)$, $i = 1,..,n_0$ is denoted by $\sigma_{0i} = \mathrm{cov}[\varepsilon(s_0), \varepsilon(s_i)]$, and if the distance from s_0 to s_i is denoted by $h_{0i} (\leq h_0)$, then the vector, $c(s_0) = (\sigma_{0i} : i = 1,..,n_0)$, of these covariances is estimated by $\hat{c}(s_0) = (\hat{\sigma}_{0i} : i = 1,..,n_0)'$, where $\hat{\sigma}_{0i} = \hat{C}(h_{0i})$.

(10) Next, if the estimate of the residual vector, $\varepsilon = (\varepsilon_i : i = 1,..,n)$, for all data points is defined (in terms of the GLS estimate of β) by $\hat{\varepsilon} = y - X\hat{\beta}$, and if the corresponding n_0-dimensional subvector of estimated residuals in S_0 is denoted by $\hat{\varepsilon}_0$, then the predicted residual, $\hat{\varepsilon}(s_0)$, is given by the (simple) kriging expression,

$$\boxed{\hat{\varepsilon}(s_0) = c(s_0)\hat{C}_0^{-1}\hat{\varepsilon}_0} \ .$$

(11) Finally, the desired *prediction*, $\hat{Y}(s_0)$, at s_0 is given by

$$\boxed{\hat{Y}(s_0) = x(s_0)'\hat{\beta} + \hat{\varepsilon}(s_0)} \ .$$