Ordinary kriging in terms of the covariance function

The model:
The model assumption is:
\[ Z(s) = \mu + \delta(s) \]
where \( \delta(s) \) is a zero mean stochastic term with variogram \( 2\gamma(\cdot) \).

The Kriging System
The predictor assumption is
\[
\hat{Z}(s_0) = \sum_{i=1}^{n} w_i Z(s_i)
\]
It is a weighted average of the sample values, and \( \sum_{i=1}^{n} w_i = 1 \) to ensure unbiasedness. The \( w_i \)'s are the weights that will be estimated.

Kriging minimizes the mean squared error of prediction
\[
\min \sigma^2_e = E[(Z(s_0) - \hat{Z}(s_0))^2]
\]
or
\[
\min \sigma^2_e = E \left[ (Z(s_0) - \sum_{i=1}^{n} w_i Z(s_i))^2 \right]
\]
For second order stationary process the last equation can be written as:
\[
\sigma^2_e = C(0) - 2 \sum_{i=1}^{n} w_i C(s_i, s_i) + \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j C(s_i, s_j) \tag{1}
\]
See next page for the proof:
Let’s examine \((Z(s_0) - \sum_{i=1}^{n} w_i Z(s_i))^2:\)

\[
(z(s_0) - \sum_{i=1}^{n} w_i z(s_i) + \mu - \mu)^2 = \left\{[z(s_0) - \mu] - \sum_{i=1}^{n} w_i [z(s_i) - \mu]\right\}^2 = [z(s_0) - \mu]^2 - 2 \sum_{i=1}^{n} w_i [z(s_i) - \mu][z(s_0) - \mu] + \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j [z(s_i) - \mu][z(s_j) - \mu].
\]

If we take expectations on the last expression we have

\[
E[z(s_0) - \mu]^2 - 2 \sum_{i=1}^{n} w_i E[z(s_i) - \mu][z(s_0) - \mu] + \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j E[z(s_i) - \mu][z(s_j) - \mu]
\]

The expectations above are the covariances:

\[
C(0) - 2 \sum_{i=1}^{n} w_i C(s_0, s_i) + \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j C(s_i, s_j)
\]

Therefore kriging minimizes

\[
\sigma_e^2 = E[(Z(s_0) - \sum_{i=1}^{n} w_i Z(s_i))^2] = C(0) - 2 \sum_{i=1}^{n} w_i C(s_0, s_i) + \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j C(s_i, s_j)
\]

subject to

\[
\sum_{i=1}^{n} w_i = 1
\]

The minimization is carried out over \((w_1, w_2, ..., w_n),\) subject to the constraint \(\sum_{i=1}^{n} w_i = 1.\) Therefore the minimization problem can be written as:

\[
\min C(0) - 2 \sum_{i=1}^{n} w_i C(s_0, s_i) + \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j C(s_i, s_j) - 2\lambda(\sum_{i=1}^{n} w_i - 1) \quad (2)
\]

where \(\lambda\) is the Lagrange multiplier. After differentiating (2) with respect to \(w_1, w_2, ..., w_n,\) and \(\lambda\) and set the derivatives equal to zero we find that

\[
2 \sum_{j=1}^{n} w_j C(s_i, s_j) - 2C(s_0, s_i) - 2\lambda = 0, \quad i = 1, ..., n
\]

\[
\sum_{j=1}^{n} w_j C(s_i, s_j) - C(s_0, s_i) - \lambda = 0, \quad i = 1, ..., n
\]

and

\[
\sum_{i=1}^{n} w_i = 1
\]
Using matrix notation the previous system of equations can be written as

\[ CW = c \]

Therefore the weights \( w_1, w_2, ..., w_n \) and the Lagrange multiplier \( \lambda \) can be obtained by

\[ W = C^{-1} c \]

where

\[ W = (w_1, w_2, ..., w_n, -\lambda) \]

\[ c = (C(s_0, s_1), C(s_0, s_2), ..., C(s_0, s_n), 1)' \]

\[ C = \begin{cases} 
C(s_i, s_j), & i = 1, 2, ..., n, \ j = 1, 2, ..., n, \\
1, & i = n + 1, \ j = 1, ..., n, \\
1, & j = n + 1, \ i = 1, ..., n, \\
0, & i = n + 1, \ j = n + 1.
\end{cases} \]

The variance of the estimator:
So far, we found the weights and therefore we can compute the estimator: \( \hat{Z}(s_0) = \sum_{i=1}^{n} w_i Z(s_i) \).
How about the variance of the estimator, namely \( \sigma^2_e \)?

We multiply

\[ \sum_{j=1}^{n} w_j C(s_i, s_j) - C(s_0, s_i) - \lambda = 0, \ i = 1, ..., n \]

by \( w_i \) and we sum over all \( i = 1, \cdots, n \) to get:

\[ \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j C(s_i, s_j) - \sum_{i=1}^{n} w_i C(s_0, s_i) - \sum_{i=1}^{n} w_i \lambda = 0 \]

Therefore,

\[ \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j C(s_i, s_j) = \sum_{i=1}^{n} w_i C(s_0, s_i) + \lambda \]

If we substitute this result into equation (1) we finally get:

\[ \sigma^2_e = C(0) - \sum_{i=1}^{n} w_i C(s_i, s_0) + \lambda \] (3)
The kriging system in terms of covariance

\[
\begin{pmatrix}
C(s_1, s_1) & C(s_1, s_2) & C(s_1, s_3) & \cdots & C(s_1, s_n) & 1 \\
C(s_2, s_1) & C(s_2, s_2) & C(s_2, s_3) & \cdots & C(s_2, s_n) & 1 \\
\vdots & \vdots & \ddots & \cdots & \vdots & \vdots \\
C(s_n, s_1) & C(s_n, s_2) & C(s_n, s_3) & \cdots & C(s_n, s_n) & 1 \\
1 & 1 & \cdots & \cdots & 1 & 0
\end{pmatrix}
\begin{pmatrix}
w_1 \\
w_2 \\
\vdots \\
w_n \\
-\lambda
\end{pmatrix}
= 
\begin{pmatrix}
C(s_0, s_1) \\
C(s_0, s_2) \\
\vdots \\
C(s_0, s_n)
\end{pmatrix}
\]

Again we observe that the matrix $C$ must be positive definite and this ensured by a choice of a model covariance function.
Short code for ordinary kriging in terms of variogram:

```r
a <- read.table("kriging_1.txt", header=TRUE)
b <- read.table("kriging_11.txt", header=TRUE)

x <- as.matrix(cbind(a$x, a$y))

x1 <- rep(rep(0,8),8)  #Initialize
dist <- matrix(x1,nrow=8,ncol=8)  #the distance matrix

for (i in 1:8){
  for (j in 1:8){
    dist[i,j]=((x[i,1]-x[j,1])^2+(x[i,2]-x[j,2])^2)^.5
  }
}

c0 <- 0
c1 <- 10
alpha <- 3.33

x1 <- rep(rep(0,8),8)  #Initialize
G <- matrix(x1,nrow=8,ncol=8)  #the GAMMA matrix

for(i in 1:8){
  for (j in 1:8){
    G[i,j]=c1*(1-exp(-dist[i,j]/alpha))
    if(i==j){G[i,j]=0}
    if(i==8){G[i,j]=1}
    if(j==8){G[i,j]=1}
    if(i==8 & j==8) {G[i,j]=0}
  }
}

g <- rep(0,8)  #Initialize
               #the g vector

for(j in 1:8){
g[j]=c1*(1-exp(-dist[8,j]/alpha))
  if(j == 8) {g[j]=1}
}

w <- solve(G) %*% g  #Obtain the weights and the Lagrange parameter

z_hat <- w[-8] %*% b$z  #Compute the estimate
var_z_hat <- t(w) %*% g  #Compute the variance of the estimate
```