Appendix S1

Kriging Models for Linear Networks and non-Euclidean Distances:

Cautions and Solutions

Jay M. Ver Hoef

Marine Mammal Laboratory, NOAA Fisheries Alaska Fisheries Science Center
7600 Sand Point Way NE, Seattle, WA 98115
tel: (907) 347-5552 E-mail: jay.verhoef@noaa.gov

January 11, 2018

SUPPLEMENTAL MATERIAL

Estimation Methods

I use two methods to fit theoretical semivariograms eqn 3 to empirical semivariograms eqn 5. The first is simple weighted least squares. To show the dependence of the theoretical semivariogram on parameters, write any of the models, eqn 3, in semivariogram form with a nugget effect, $\gamma(h_k|\boldsymbol{\theta}) = \sigma_0^2 + \sigma_p^2(1 - \rho_m(h_k|\alpha))$, where $\boldsymbol{\theta} = (\sigma_p^2, \sigma_0^2, \alpha)$. Then the weighted least squares estimator of $\boldsymbol{\theta}$ is,

$$\hat{\boldsymbol{\theta}}_{WLS} = \underset{\boldsymbol{\theta}}{\operatorname{argmin}} \sum_{k=1}^{K} [N(\mathcal{D}_k)](\hat{\gamma}(h_k) - \gamma(h_k|\boldsymbol{\theta}))^2.$$

Cressie's weighted least squares estimate of $\boldsymbol{\theta}$ is,

$$\hat{\boldsymbol{\theta}}_{CWLS} = \underset{\boldsymbol{\theta}}{\operatorname{argmin}} \sum_{k=1}^{K} [N(\mathcal{D}_k)] \left(\frac{\hat{\gamma}(h_k)}{\gamma(h_k|\boldsymbol{\theta})} - 1 \right)^2.$$

REML does not use an empirical semivariogram. Rather, let $\mathbf{y}$ be a vector of observed data of length n, $\mathbf{X}$ a fixed effects design matrix with n rows and p linearly independent columns, $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ an $n \times n$ covariance matrix, in the same order as the data, that depends on distances between observations, and a set of parameters, as given in eqn 2. Note that I show the dependence of $\boldsymbol{\Sigma}$ on $\boldsymbol{\theta}$ with a subscript. Then REML estimates are given by

$$\begin{aligned} \hat{\boldsymbol{\theta}}_{REML} = \underset{\boldsymbol{\theta}}{\operatorname{argmin}}\ & [(\boldsymbol{Y} - \mathbf{X}\boldsymbol{\beta}_g)'\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{-1}(\boldsymbol{Y} - \mathbf{X}\boldsymbol{\beta}_g) + \log(|\boldsymbol{\Sigma}_{\boldsymbol{\theta}}|) + \\ & \log(|\mathbf{X}'\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{-1}\mathbf{X}|) + (n-p)\log(2\pi)], \end{aligned} \qquad \text{eqn S.1}$$

where

$$\boldsymbol{\beta}_g = (\mathbf{X}'\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{-1}\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{-1}\boldsymbol{Y} \qquad \text{eqn S.2}$$

is the generalized least squares estimator of $\boldsymbol{\beta}$. Note that for our case, $\mathbf{X} = \mathbf{1}$, where $\mathbf{1}$ is a vector of all 1s, and $\boldsymbol{\beta} = \mu$, a scalar.

Simple Example on Negative Variances from Improper Covariance Matrices

For a very simple, worked example in `R` on how a covariance matrix that is not positive definite can lead to negative variances, consider the 4 locations in a linear network shown in Figure S1.

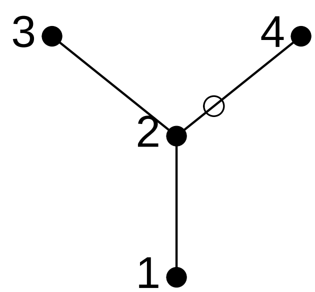

Figure S1: A simple 4-location network, where each location is given by a solid circle numbered from 1 to 4, along with a prediction location, shown by the open circle.

Let the linear distance between each connected location be 1 unit, so the distance matrix among the 4 locations, numbered sequentially for the rows and columns, is

```
linDmat = rbind(
  c(0,1,2,2),
  c(1,0,1,1),
  c(2,1,0,2),
  c(2,1,2,0))
```

$$\mathbf{D} = \begin{pmatrix} 0 & 1 & 2 & 2 \\ 1 & 0 & 1 & 1 \\ 2 & 1 & 0 & 2 \\ 2 & 1 & 2 & 0 \end{pmatrix}$$

I will use the Gaussian autocorrelation model, eqn 3, with $\sigma_p^2 = 1$, $\alpha = 3$, and a small nugget effect, $\sigma_0 = 0.01$.

```r
Sig = exp(-(linDmat/3)^2) + diag(rep(0.01, times = 4))
```

$$\mathbf{\Sigma} = \begin{pmatrix} 1.010 & 0.895 & 0.641 & 0.641 \\ 0.895 & 1.010 & 0.895 & 0.895 \\ 0.641 & 0.895 & 1.010 & 0.641 \\ 0.641 & 0.895 & 0.641 & 1.010 \end{pmatrix} \qquad \text{eqn S.3}$$

The spectral decomposition, $\mathbf{\Sigma} = \mathbf{Q\Lambda Q}'$ (eqn 10) is

```r
Lambda = diag(eigen(Sig)$values)
Q = eigen(Sig)$vectors
```

$$\mathbf{\Lambda} = \begin{pmatrix} 3.328 & 0.000 & 0.000 & 0.000 \\ 0.000 & 0.369 & 0.000 & 0.000 \\ 0.000 & 0.000 & 0.369 & 0.000 \\ 0.000 & 0.000 & 0.000 & -0.026 \end{pmatrix} \quad \mathbf{Q} = \begin{pmatrix} -0.480 & 0.000 & 0.816 & -0.321 \\ -0.556 & -0.000 & 0.000 & 0.831 \\ -0.480 & -0.707 & -0.408 & -0.321 \\ -0.480 & 0.707 & -0.408 & -0.321 \end{pmatrix} \qquad \text{eqn S.4}$$

The eigenvectors, $\mathbf{v}_i; i = 1, \ldots, 4$, in $\mathbf{Q} = [\mathbf{v}_1|\mathbf{v}_2|\mathbf{v}_3|\mathbf{v}_4]$ are orthonormal, which means that $\mathbf{v}_i'\mathbf{v}_j = 0$ if $i \neq j$, but $\mathbf{v}_i'\mathbf{v}_i = 1$.

```r
t(Q[,1]) %*% Q[,4]

##               [,1]
## [1,] 2.775558e-17

t(Q[,4]) %*% Q[,4]

##      [,1]
## [1,]    1
```

Now, consider 4 random variables, $\boldsymbol{Y} = \{Y_1, Y_2, Y_3, Y_4\}$. The linear combination $\mathbf{v}_4'\boldsymbol{Y} = -0.321Y_1 + 0.831Y_2 - 0.321Y_3 - 0.321Y_4$ is a perfectly valid construction, and must have a positive variance. However, if $\boldsymbol{Y}$ has covariance matrix $\mathbf{\Sigma}$ in eqn S.3, then $\text{var}(\mathbf{v}_4'\boldsymbol{Y}) = \mathbf{v}_4'\mathbf{\Sigma}\mathbf{v}_4 = -0.026$, which is the 4th eigenvalue,

```
v4 = Q[,4]
t(v4) %*% Sig %*% v4

##               [,1]
## [1,] -0.02611639
```

which is not a valid variance, so $\mathbf{\Sigma}$ in eqn S.3 is not a valid covariance matrix.

To show how this works for kriging, consider predicting the location shown with the open circle in Figure S1, which is 3/10 of the way from location 2 to location 4. Then the distance from the 4 locations with solid circles in Figure S1 to the prediction location is the vector $(1.3, 0.3, 1.3, 0.7)$, and the covariances between the prediction location and the 4 locations with solid circles in Figure S1 is

```
cvec = exp(-(c(1.3, 0.3, 1.3, 0.7)/3)^2)
cvec

## [1] 0.8287989 0.9900498 0.8287989 0.9470111
```

Using eqn 7, the prediction variance of the location with the open circle, using data from the locations with the solid black circles, would be computed as

```
(1 + 0.01) - t(cvec) %*% solve(Sig) %*% cvec +
  (1 - (sum(solve(Sig) %*% cvec))^2)/sum(solve(Sig))

##              [,1]
## [1,] -0.0425027
```

which is negative, so we see that the larger matrix, where $\mathbf{\Sigma}$ is appended with covariances that include the prediction location, eqn 9, is not a valid covariance matrix.