CHAPTER 4

CIRCULAR KRIGING

4.1 Introduction

4.1.1 Objective

The objective of this chapter is to develop a circular kriging estimate of direction with optimum properties based on a sample of circular-spatial data. Kriging is a body of techniques for estimating continuous and spatially correlated data. The name of the technique is derived from Daniel G. Krige, a South African mining geologist, who originated the method for linear-spatial data. The circular kriging estimate is a linear combination of observations of direction with weights based on the spatial correlation as estimated by the cosine model fitted to the empirical cosineogram (Chapter 3), and may be imaged using arrow plots or the circular dataimage (Chapter 2). Figure 4-1 illustrates the kriging of simulated circular-spatial data. The R package CircSpatial function KrigCRF is documented in Appendix J, Section J.5.

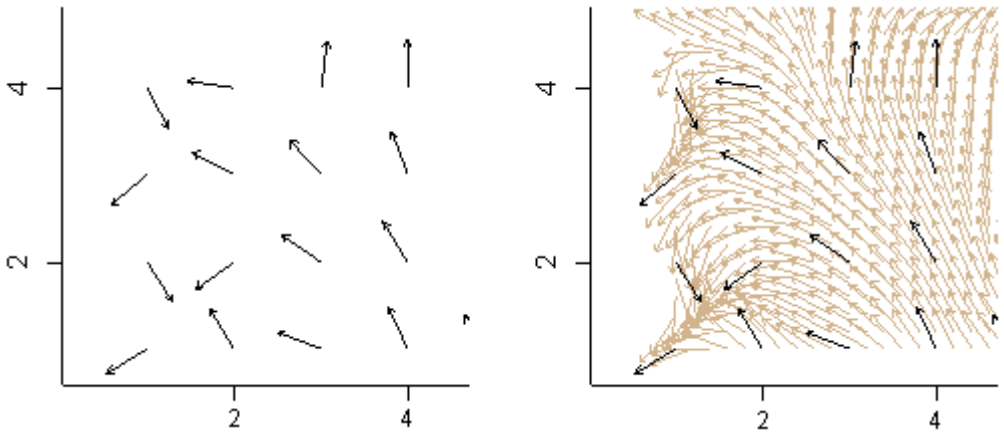

Figure 4-1. Circular Kriging, the Interpolation of Circular-Spatial Data Based on Spatial Correlation. The left plot shows simulated circular-spatial data. In the right plot, the simulated data (black) is superimposed on the kriged estimate of the simulated data (tan).

4.1.2 Chapter Organization

This chapter is organized as follows: Section 4.2 derives the circular kriging estimator, correcting the result from McNeill (1993). Section 4.3 proves that the estimator is optimal. Section 4.4 gives an alternate formula, which is computationally efficient. Section 4.5 shows the kriging behavior around a sampled location, and proves that the estimated direction at a sampled location is the observed direction. This is called “exact interpolation” in linear kriging. Section 4.6 derives the circular kriging variance of the circular kriging estimator, correcting the result from McNeill (1993). Section 4.7 shows how the circular kriging variance varies with distance and the circular-spatial correlation model. Section 4.8 concludes with the summary and description of future work.

4.2 Solution

4.2.1 A Linear Combination of Observations

The estimated spatial correlation parameters (see nugget, range, and sill in Chapter 3) are assumed to be reasonably accurate. Inaccuracy increases error and the variability of the estimate. The circular random field (CRF), as introduced in Chapter 1 and further discussed in Chapter 5, is assumed to be isotropic (spatial correlation independent of direction).

See Appendix A for a description of the notation and Appendix B for referenced Equations (B.1) to (B.12). In Chapter 3, the direction at location $\boldsymbol{x}_0$ was denoted by the scalar θ_0 in radians. The conversion from the scalar to the unit vector representation of direction is $\boldsymbol{u}_0 = \begin{bmatrix} \cos(\theta_0) \\ \sin(\theta_0) \end{bmatrix}$. Let $\hat{\boldsymbol{u}}_0$ be the unit vector estimate of the direction $\boldsymbol{u}_0$ at the

unmeasured location $\boldsymbol{x}_0$, and $\boldsymbol{u}_i, i = 1,2,\cdots,n$ be observations of direction as unit vectors at measured locations $\boldsymbol{x}_i, i = 1,2,\cdots,n,$ respectively.

Spatially correlated observations contain information about $\boldsymbol{u}_0$. Because spatial correlation increases as distance decreases, observations nearer to the estimation location carry more information about $\boldsymbol{u}_0$ than observations more distant. Hence a weighted linear combination of the observations $\boldsymbol{u}_i, i = 1,2,\cdots,n$, is required. Let $w_i, i = 1,2,\cdots,n,$ be the weights with $w_i \in R$ (the set of real numbers).

$$\begin{aligned}\hat{\boldsymbol{u}}_0 &= \boldsymbol{u}_1 w_1 + \boldsymbol{u}_2 w_2 + \cdots + \boldsymbol{u}_n w_n \\ &= \begin{bmatrix}\boldsymbol{u}_1 & \boldsymbol{u}_2 & \cdots & \boldsymbol{u}_n\end{bmatrix}\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} \\ &\equiv \boldsymbol{U}\boldsymbol{w} \Rightarrow\end{aligned}$$

$$\hat{\boldsymbol{u}}_0 = \boldsymbol{U}\boldsymbol{w} \tag{4.1}$$

$\boldsymbol{w}$ will ultimately be determined by a constrained optimization in Subsection 4.2.7 such that $\hat{\boldsymbol{u}}_0$ is a unit vector. In general, the length of the vector estimate $\hat{\boldsymbol{u}}_0$ for any $\boldsymbol{w} \in R^n$ is

$$\begin{aligned}\left\|\hat{\boldsymbol{u}}_0\right\| &\overset{(B.2)}{=} \sqrt{\hat{\boldsymbol{u}}_0^T \hat{\boldsymbol{u}}_0} \\ &\overset{(4.1)}{=} \sqrt{(\boldsymbol{U}\boldsymbol{w})^T (\boldsymbol{U}\boldsymbol{w})} \\ &= \sqrt{\boldsymbol{w}^T \boldsymbol{U}^T \boldsymbol{U}\,\boldsymbol{w}} \\ &= \sqrt{\boldsymbol{w}^T \begin{bmatrix}\boldsymbol{u}_1^T \\ \boldsymbol{u}_2^T \\ \vdots \\ \boldsymbol{u}_n^T\end{bmatrix}\begin{bmatrix}\boldsymbol{u}_1 & \boldsymbol{u}_2 & \cdots & \boldsymbol{u}_n\end{bmatrix}\boldsymbol{w}} = \sqrt{\boldsymbol{w}^T \begin{bmatrix}\boldsymbol{u}_1^T\boldsymbol{u}_1 & \boldsymbol{u}_1^T\boldsymbol{u}_2 & \cdots & \boldsymbol{u}_1^T\boldsymbol{u}_n \\ \boldsymbol{u}_2^T\boldsymbol{u}_1 & \boldsymbol{u}_2^T\boldsymbol{u}_2 & \cdots & \boldsymbol{u}_2^T\boldsymbol{u}_n \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{u}_n^T\boldsymbol{u}_1 & \boldsymbol{u}_n^T\boldsymbol{u}_2 & \cdots & \boldsymbol{u}_n^T\boldsymbol{u}_n\end{bmatrix}\boldsymbol{w}}\end{aligned}$$

$$\overset{*}{=}\sqrt{\boldsymbol{w}^T\begin{bmatrix} 1 & \boldsymbol{u}_1^T\boldsymbol{u}_2 & \cdots & \boldsymbol{u}_1^T\boldsymbol{u}_n \\ \boldsymbol{u}_2^T\boldsymbol{u}_1 & 1 & \cdots & \boldsymbol{u}_2^T\boldsymbol{u}_n \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{u}_n^T\boldsymbol{u}_1 & \boldsymbol{u}_n^T\boldsymbol{u}_2 & \cdots & 1 \end{bmatrix}\boldsymbol{w}}$$

$$\equiv\sqrt{\boldsymbol{w}^T\tilde{\boldsymbol{K}}\boldsymbol{w}}\Rightarrow$$

$$\left\|\hat{\boldsymbol{u}}_0\right\|=\sqrt{\boldsymbol{w}^T\tilde{\boldsymbol{K}}\boldsymbol{w}}. \tag{4.2}$$

* For unit vectors $\boldsymbol{u}_i$, the diagonal elements of $\boldsymbol{U}^T\boldsymbol{U}$ are
$\boldsymbol{u}_i^T\boldsymbol{u}_i \overset{(B.1)}{=} \cos(\theta_i)\cos(\theta_i)+\sin(\theta_i)\sin(\theta_i)=\cos^2(\theta_i)+\sin^2(\theta_i)=1$.

4.2.2 Optimality

Let $\boldsymbol{e}_0$ be the error vector equal to the unit vector estimate of direction $\hat{\boldsymbol{u}}_0$ minus the unobserved direction $\boldsymbol{u}_0$ (unit vector), or $\boldsymbol{e}_0=\hat{\boldsymbol{u}}_0-\boldsymbol{u}_0$. $\boldsymbol{e}_0$, $\hat{\boldsymbol{u}}_0$, and $\boldsymbol{u}_0$ are located at $\boldsymbol{x}_0$. The addition of the unit vector $\boldsymbol{u}_0$ and $\boldsymbol{e}_0$ is shown in Figure 4-2 as $\boldsymbol{u}_0+\boldsymbol{e}_0=\boldsymbol{u}_0+(\hat{\boldsymbol{u}}_0-\boldsymbol{u}_0)=\hat{\boldsymbol{u}}_0$. In words, the unobserved direction plus the error vector equals the estimate of direction. Hence, $\boldsymbol{e}_0$ is a vector from the head of $\boldsymbol{u}_0$ to the head of $\hat{\boldsymbol{u}}_0$, and the length of the error vector is the distance from the head of $\boldsymbol{u}_0$ to the head of $\hat{\boldsymbol{u}}_0$. Let θ be the angle between these vectors in $[0,\pi]$. When $\theta=0, \left\|\boldsymbol{e}_0\right\|=0,$ and when $\theta=\pi, \left\|\boldsymbol{e}_0\right\|=2.$ Hence, $0\le\left\|\boldsymbol{e}_0\right\|\le 2$.

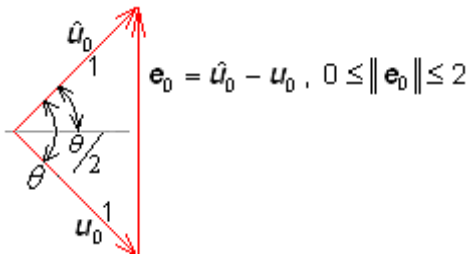

Figure 4-2. Directions Represented by the Unobserved $\boldsymbol{u}_0$, Estimate $\hat{\boldsymbol{u}}_0$, and Error $\boldsymbol{e}_0=\hat{\boldsymbol{u}}_0-\boldsymbol{u}_0$ Vectors. θ is the angle between $\boldsymbol{u}_0$ and $\hat{\boldsymbol{u}}_0$.

The unit vectors $\boldsymbol{u}_0$ and $\hat{\boldsymbol{u}}_0$ can be visualized as the hands of a clock when the tails of the vectors are positioned at the center of a clock. For any $\boldsymbol{u}_0$ and $\hat{\boldsymbol{u}}_0$, while holding θ and $\|\boldsymbol{e}_0\|$ constant, the minute and hour hands (vectors) can be transposed so $\hat{\boldsymbol{u}}_0$ is counterclockwise of $\boldsymbol{u}_0$, and the hands can be rotated so that the line joining the hands is vertical with $\boldsymbol{e}_0$ pointing upward. Since we are interested in $\|\boldsymbol{e}_0\|$ for minimization, Figure 4-2 illustrates the case of any $\boldsymbol{u}_0$ and $\hat{\boldsymbol{u}}_0$.

$\hat{\boldsymbol{u}}_0 = \boldsymbol{U}\boldsymbol{w}$, the estimate of the unobserved direction, will be considered optimal when the choice of $\boldsymbol{w}$ results a unit vector estimate $\hat{\boldsymbol{u}}_0$ with an error vector $\boldsymbol{e}_0$ of minimum squared length over all such estimates. With θ the angle between $\hat{\boldsymbol{u}}_0$ and $\boldsymbol{u}_0$, the squared length of the error vector is

$$\begin{aligned}\|\hat{\boldsymbol{u}}_0 - \boldsymbol{u}_0\|^2 &= [2\sin(\theta/2)]^2 \\ &= 4[\sin(\theta/2)]^2 \Rightarrow\end{aligned}$$

$$\|\hat{\boldsymbol{u}}_0 - \boldsymbol{u}_0\|^2 = 4[\sin(\theta/2)]^2. \tag{4.3}$$

The result (4.3) can be further transformed.

$$\begin{aligned}\cos(\theta) &= \cos(\theta/2 + \theta/2) \\ &= \cos(\theta/2)\cos(\theta/2) - \sin(\theta/2)\sin(\theta/2) \\ &= [\cos(\theta/2)]^2 - [\sin(\theta/2)]^2 \\ &= \{1 - [\sin(\theta/2)]^2\} - [\sin(\theta/2)]^2 \\ &= 1 - 2[\sin(\theta/2)]^2 \Rightarrow \\ -\cos(\theta) &= -1 + 2[\sin(\theta/2)]^2 \Rightarrow \\ 1 - \cos(\theta) &= 2[\sin(\theta/2)]^2 \Rightarrow \\ 2(1 - \cos(\theta)) &= 4[\sin(\theta/2)]^2 \\ &\overset{(4.3)}{=} \|\hat{\boldsymbol{u}}_0 - \boldsymbol{u}_0\|^2 \Rightarrow\end{aligned}$$

$$\|\hat{\boldsymbol{u}}_0 - \boldsymbol{u}_0\|^2 = 2(1 - \cos(\theta)) \tag{4.4}$$

Hence, the squared length of the error vector (4.4) is minimized by maximizing $\cos\theta$, or minimizing the angle between unit vectors $\hat{\boldsymbol{u}}_0$ and $\boldsymbol{u}_0$. With $\cos(\theta_{i0})$ denoting the cosine of the angle between observation $\boldsymbol{u}_i$ and unobserved $\boldsymbol{u}_0$, and referencing Appendix B, Equation (B.1), it follows that

$$\begin{aligned}
\cos(\theta) &\overset{(B.1)}{=} \hat{\boldsymbol{u}}_0{}^T \boldsymbol{u}_0 \\
&\overset{(4.1)}{=} (\boldsymbol{U}\boldsymbol{w})^T \boldsymbol{u}_0 \\
&= \boldsymbol{w}^T \begin{bmatrix} \boldsymbol{u}_1^T \\ \boldsymbol{u}_2^T \\ \vdots \\ \boldsymbol{u}_n^T \end{bmatrix} \boldsymbol{u}_0 \\
&= \boldsymbol{w}^T \begin{bmatrix} \boldsymbol{u}_1^T\boldsymbol{u}_0 \\ \boldsymbol{u}_2^T\boldsymbol{u}_0 \\ \vdots \\ \boldsymbol{u}_n^T\boldsymbol{u}_0 \end{bmatrix} \\
&\overset{(B.1)}{=} \boldsymbol{w}^T \begin{bmatrix} \cos(\theta_{10}) \\ \cos(\theta_{20}) \\ \vdots \\ \cos(\theta_{n0}) \end{bmatrix} \\
&\equiv \boldsymbol{w}^T \tilde{\boldsymbol{c}} \Rightarrow
\end{aligned}$$

$$\cos\theta = \boldsymbol{w}^T \tilde{\boldsymbol{c}}. \tag{4.5}$$

4.2.3 Cosineogram

$\tilde{\boldsymbol{c}}$ must be estimated since $\boldsymbol{u}_0$ is unknown. Also, the eigenvalues of $\tilde{\boldsymbol{K}} = \boldsymbol{U}^T\boldsymbol{U}$ have been observed to be not all positive indicating that $\tilde{\boldsymbol{K}}$ is generally not positive definite according to Appendix B, Equation (B.3). However, positive definiteness is required for maximum fit (Section 4.3). This is accomplished by replacing $\tilde{\boldsymbol{c}}$ and $\tilde{\boldsymbol{K}}$ with estimates $\boldsymbol{c}$ and $\boldsymbol{K}$, respectively, which are computed on smooth and positive definite functions as is done in the kriging of linear-spatial data (Bailey and Gatrell 1995).

For circular-spatial data, the empirical cosineogram, a rough plot of circular-spatial correlation (Chapter 3, Section 3.2), is modeled by the positive definite cosine model (Chapter 3, Section 3.6 and Appendix M) with best fit. The elements of $\boldsymbol{c}$ and $\boldsymbol{K}$ are computed using the selected cosine model and the distances between measurement locations. $\boldsymbol{c}$, which depends on the estimation location, must be updated for each location to be estimated. $\boldsymbol{K}$, which depends only on the distances between observations, is computed once.

4.2.4 First Iteration of the Circular Kriging Solution $\boldsymbol{w}$

Recall that $\boldsymbol{w}^T\boldsymbol{c}$ is a linear combination of the expected cosines of the angles between the unobserved direction as a unit vector $\boldsymbol{u}_0$ and the sampled directions $\boldsymbol{u}_i, i = 1,2,\cdots,n$, with $\tilde{\boldsymbol{c}}$ replaced by $\boldsymbol{c}$ in (4.5). It expresses the fit of $\hat{\boldsymbol{u}}_0$ to $\boldsymbol{u}_0$. $\boldsymbol{w}^T\boldsymbol{K}\boldsymbol{w}$ is the squared length of $\hat{\boldsymbol{u}}_0$, with $\tilde{\boldsymbol{K}}$ replaced by $\boldsymbol{K}$ in (4.2). $\boldsymbol{w}^T\boldsymbol{K}\boldsymbol{w} - 1 = 0$ expresses that the squared length of $\hat{\boldsymbol{u}}_0$ is constrained to 1. The vector of weights $\boldsymbol{w}$ will be solved by maximizing $\boldsymbol{w}^T\boldsymbol{c}$ relative to $\boldsymbol{w}$ with the maximization constrained such that $\boldsymbol{w}^T\boldsymbol{K}\boldsymbol{w}$ is equal to one. The method of Lagrange multipliers (Grossman 1988), for finding the extrema of a function of several variables subject to a constraint, introduces a new unknown scalar variable, which is called the Lagrange multiplier, and defines a new function, which is called the Lagrangian, in terms of the original function, the constraint, and the Lagrange multiplier. At the stationary point, the Lagrange multiplier is the proportionality of the gradient of the function to be maximized and the gradient of the constraint. Let ν be the Lagrange multiplier and q be the Lagrangian.

$$q = \boldsymbol{w}^T\boldsymbol{c} - \frac{\nu}{2}\left(\boldsymbol{w}^T\boldsymbol{K}\boldsymbol{w} - 1\right) \tag{4.6}$$

ν is divided by two to simplify a subsequent result.

Differentiating (4.6) with respect to $\boldsymbol{w}$ according Appendix B, Equations (B.8) and (B.9),

$$\begin{aligned}\frac{\partial}{\partial \boldsymbol{w}}\left(\boldsymbol{w}^T\boldsymbol{c}-\frac{\nu}{2}\left(\boldsymbol{w}^T\boldsymbol{K}\boldsymbol{w}-1\right)\right) &= \boldsymbol{c}-\frac{\nu}{2}\left(2\boldsymbol{K}\boldsymbol{w}-0\right)\\ &= \boldsymbol{c}-\nu\boldsymbol{K}\boldsymbol{w}\Rightarrow\end{aligned}$$

$$\frac{\partial}{\partial \boldsymbol{w}}\left(\boldsymbol{w}^T\boldsymbol{c}-\frac{\nu}{2}\left(\boldsymbol{w}^T\boldsymbol{K}\boldsymbol{w}-1\right)\right)=\boldsymbol{c}-\nu\boldsymbol{K}\boldsymbol{w}. \tag{4.7}$$

Setting the derivative (4.7) equal to the zero vector $\boldsymbol{0}=\begin{bmatrix}0 & 0 & \cdots & 0\end{bmatrix}^T$, $\boldsymbol{c}-\nu\boldsymbol{K}\boldsymbol{w}=\boldsymbol{0} \Rightarrow \nu\boldsymbol{K}\boldsymbol{w}=\boldsymbol{c}$. From the invertibility of $\boldsymbol{K}$ (Appendix B, Equation (B.4)) it follows that

$$\boldsymbol{w}=\nu^{-1}\boldsymbol{K}^{-1}\boldsymbol{c}. \tag{4.8}$$

ν can be determined from the unit length constraint $\boldsymbol{w}^T\boldsymbol{K}\boldsymbol{w}=1$.

$$\begin{aligned}1&=\boldsymbol{w}^T\boldsymbol{K}\boldsymbol{w}\\ &\overset{(4.8)}{=}\left(\nu^{-1}\boldsymbol{K}^{-1}\boldsymbol{c}\right)^T\boldsymbol{K}\left(\nu^{-1}\boldsymbol{K}^{-1}\boldsymbol{c}\right)\\ &=\boldsymbol{c}^T\left(\boldsymbol{K}^{-1}\right)^T\nu^{-1}\boldsymbol{K}\nu^{-1}\boldsymbol{K}^{-1}\boldsymbol{c}\\ &\overset{(B.5)}{=}\nu^{-1}\nu^{-1}\boldsymbol{c}^T\boldsymbol{K}^{-1}\boldsymbol{K}\boldsymbol{K}^{-1}\boldsymbol{c}\\ &=\nu^{-2}\boldsymbol{c}^T\boldsymbol{K}^{-1}\boldsymbol{c}\Rightarrow\\ 1&=\nu^{-2}\boldsymbol{c}^T\boldsymbol{K}^{-1}\boldsymbol{c}\Rightarrow\\ \nu^{+2}&=\boldsymbol{c}^T\boldsymbol{K}^{-1}\boldsymbol{c}\Rightarrow\\ \nu&=\pm\sqrt{\boldsymbol{c}^T\boldsymbol{K}^{-1}\boldsymbol{c}}\end{aligned}$$

The Lagrange multiplier ν is a scalar representing the proportionality of parallel gradient vectors of the function being optimized and the constraint. The positive sign of ν is selected to maximize fit as will be explained in detail in Section 4.3**.** It follows that

$$\nu=+\sqrt{\boldsymbol{c}^T\boldsymbol{K}^{-1}\boldsymbol{c}}, \tag{4.9}$$

and $$\boldsymbol{w} \overset{(4.8)}{=} \nu^{-1}\boldsymbol{K}^{-1}\boldsymbol{c}$$

$$\overset{(4.9)}{=} \frac{1}{\sqrt{\boldsymbol{c}^T\boldsymbol{K}^{-1}\boldsymbol{c}}}\boldsymbol{K}^{-1}\boldsymbol{c} \Rightarrow$$

$$\boldsymbol{w} = \frac{1}{\sqrt{\boldsymbol{c}^T\boldsymbol{K}^{-1}\boldsymbol{c}}}\boldsymbol{K}^{-1}\boldsymbol{c}. \tag{4.10}$$

(4.10) is McNeill's result (1993, p. 40, eq. 6).

4.2.5 Length of the Estimator $\hat{\boldsymbol{u}}_0$

Substituting the solution into the estimator, the squared length is

$$\begin{aligned}
\left\|\hat{\boldsymbol{u}}_0\right\|^2 &\overset{(B.2)}{=} \left(\sqrt{\hat{\boldsymbol{u}}_0^T\hat{\boldsymbol{u}}_0}\right)^2 \\
&\overset{(4.1)}{=} \boldsymbol{w}^T\boldsymbol{U}^T\boldsymbol{U}\boldsymbol{w} \\
&\overset{(4.10)}{=} \left(\frac{1}{\sqrt{\boldsymbol{c}^T\boldsymbol{K}^{-1}\boldsymbol{c}}}\boldsymbol{K}^{-1}\boldsymbol{c}\right)^T\boldsymbol{U}^T\boldsymbol{U}\left(\frac{1}{\sqrt{\boldsymbol{c}^T\boldsymbol{K}^{-1}\boldsymbol{c}}}\boldsymbol{K}^{-1}\boldsymbol{c}\right) \\
&= \left(\boldsymbol{c}^T\left(\boldsymbol{K}^{-1}\right)^T\frac{1}{\sqrt{\boldsymbol{c}^T\boldsymbol{K}^{-1}\boldsymbol{c}}}\right)\boldsymbol{U}^T\boldsymbol{U}\left(\frac{1}{\sqrt{\boldsymbol{c}^T\boldsymbol{K}^{-1}\boldsymbol{c}}}\boldsymbol{K}^{-1}\boldsymbol{c}\right) \\
&= \frac{\boldsymbol{c}^T\boldsymbol{K}^{-1}\boldsymbol{U}^T\boldsymbol{U}\boldsymbol{K}^{-1}\boldsymbol{c}}{\boldsymbol{c}^T\boldsymbol{K}^{-1}\boldsymbol{c}} \\
&\equiv s \Rightarrow
\end{aligned}$$

$$\left\|\hat{\boldsymbol{u}}_0\right\|^2 = \frac{\boldsymbol{c}^T\boldsymbol{K}^{-1}\boldsymbol{U}^T\boldsymbol{U}\boldsymbol{K}^{-1}\boldsymbol{c}}{\boldsymbol{c}^T\boldsymbol{K}^{-1}\boldsymbol{c}}. \tag{4.11}$$

Is $\hat{\boldsymbol{u}}_0$ a unit vector?

4.2.6 $\hat{\boldsymbol{u}}_0$ Is Likely Not a Unit Vector

If $\boldsymbol{U}^T\boldsymbol{U} = \boldsymbol{K} \overset{(4.11)}{\Rightarrow} \left\|\hat{\boldsymbol{u}}_0\right\|^2 = 1$. However, with $\boldsymbol{U}^T\boldsymbol{U}$ a realization of the continuous random matrix $\boldsymbol{V}^T\boldsymbol{V}$,

$\textit{probability}\left(\boldsymbol{V}^T\boldsymbol{V}=\boldsymbol{K}\right)=0 \overset{(4.11)}{\Rightarrow} \textit{probability}\left(\left\|\hat{\boldsymbol{u}}_0\right\|=1\right)=0$. Thus, it is likely that $\hat{\boldsymbol{u}}_0$ is not a unit vector (except for a set of Borel-measure 0).

4.2.7 Corrected Circular Kriging Solution $\boldsymbol{w}$

Equation (4.11) suggests that the matrix of cosines $\boldsymbol{K}$ be scaled by s. It will be shown in Section 4.3 that scaling $\boldsymbol{K}$ by s leads to a unit vector solution. The revised function to maximize is

$$q_s \equiv \boldsymbol{w}^T\boldsymbol{c}-\frac{\nu}{2}\left(\boldsymbol{w}^T s\boldsymbol{K}\boldsymbol{w}-1\right). \tag{4.12}$$

$$\begin{aligned}\frac{\partial}{\partial \boldsymbol{w}}(q_s) &\overset{(4.12)}{=} \boldsymbol{c}-\frac{\nu}{2}(2s\boldsymbol{K}\boldsymbol{w}-0)\\ &= \boldsymbol{c}-\nu\, s\boldsymbol{K}\boldsymbol{w}\Rightarrow\end{aligned}$$

$$\frac{\partial}{\partial \boldsymbol{w}}(q_s)=\boldsymbol{c}-\nu\, s\boldsymbol{K}\boldsymbol{w} \tag{4.13}$$

Setting the derivative (4.13) equal to the zero vector $\boldsymbol{0}$,

$$\begin{aligned}\boldsymbol{c}-\nu\, s\boldsymbol{K}\boldsymbol{w}&=\boldsymbol{0}\Rightarrow\\ \nu\, s\boldsymbol{K}\boldsymbol{w}&=\boldsymbol{c}\Rightarrow\end{aligned}$$

$$\boldsymbol{w}=\nu^{-1}s^{-1}\boldsymbol{K}^{-1}\boldsymbol{c}. \tag{4.14}$$

From the constraint $\boldsymbol{w}^T s\boldsymbol{K}\boldsymbol{w}=1$ in (4.12), and keeping the sign of ν from (4.9),

$$\begin{aligned}1&=\boldsymbol{w}^T s\boldsymbol{K}\boldsymbol{w}\\ &\overset{(4.14)}{=}\left(\nu^{-1}s^{-1}\boldsymbol{K}^{-1}\boldsymbol{c}\right)^T s\boldsymbol{K}\left(\nu^{-1}s^{-1}\boldsymbol{K}^{-1}\boldsymbol{c}\right)\\ &=\boldsymbol{c}^T\left(\boldsymbol{K}^{-1}\right)^T s^{-1}\nu^{-1}s\boldsymbol{K}\nu^{-1}s^{-1}\boldsymbol{K}^{-1}\boldsymbol{c}\\ &\overset{(B.5)}{=}\nu^{-1}\nu^{-1}s^{-1}ss^{-1}\boldsymbol{c}^T\boldsymbol{K}^{-1}\boldsymbol{K}\boldsymbol{K}^{-1}\boldsymbol{c}\\ &=\nu^{-2}s^{-1}\boldsymbol{c}^T\boldsymbol{K}^{-1}\boldsymbol{c}\Rightarrow\\ 1&=\nu^{-2}s^{-1}\boldsymbol{c}^T\boldsymbol{K}^{-1}\boldsymbol{c}\Rightarrow\\ \nu^2&=s^{-1}\boldsymbol{c}^T\boldsymbol{K}^{-1}\boldsymbol{c}\Rightarrow\end{aligned}$$

$$\nu=+\sqrt{s^{-1}\boldsymbol{c}^T\boldsymbol{K}^{-1}\boldsymbol{c}}. \tag{4.15}$$

Substituting s (4.11) and ν (4.15) into $\boldsymbol{w}$ (4.14), we arrive at the principal result of this chapter.

$$\begin{aligned}
\boldsymbol{w} &\overset{(4.14)}{=} \nu^{-1} s^{-1} \boldsymbol{K}^{-1} \boldsymbol{c} \\
&\overset{(4.15)}{=} \frac{1}{\sqrt{s^{-1} \boldsymbol{c}^T \boldsymbol{K}^{-1} \boldsymbol{c}}} s^{-1} \boldsymbol{K}^{-1} \boldsymbol{c} \\
&= \frac{\boldsymbol{K}^{-1} \boldsymbol{c}}{\sqrt{s\, \boldsymbol{c}^T \boldsymbol{K}^{-1} \boldsymbol{c}}} \\
&\overset{(4.11)}{=} \frac{\boldsymbol{K}^{-1} \boldsymbol{c}}{\sqrt{\left(\frac{\boldsymbol{c}^T \boldsymbol{K}^{-1} \boldsymbol{U}^T \boldsymbol{U} \boldsymbol{K}^{-1} \boldsymbol{c}}{\boldsymbol{c}^T \boldsymbol{K}^{-1} \boldsymbol{c}} \right) \boldsymbol{c}^T \boldsymbol{K}^{-1} \boldsymbol{c}}} \\
&= \frac{\boldsymbol{K}^{-1} \boldsymbol{c}}{\sqrt{\boldsymbol{c}^T \boldsymbol{K}^{-1} \boldsymbol{U}^T \boldsymbol{U} \boldsymbol{K}^{-1} \boldsymbol{c}}} \Rightarrow
\end{aligned}$$

$$\boldsymbol{w} = \frac{\boldsymbol{K}^{-1} \boldsymbol{c}}{\sqrt{\boldsymbol{c}^T \boldsymbol{K}^{-1} \boldsymbol{U}^T \boldsymbol{U} \boldsymbol{K}^{-1} \boldsymbol{c}}} \tag{4.16}$$

This result differs from McNeill (1993, p. 40, eq. 6), who obtained $\boldsymbol{w} = \frac{\boldsymbol{K}^{-1} \boldsymbol{c}}{\sqrt{\boldsymbol{c}^T \boldsymbol{K}^{-1} \boldsymbol{c}}}$.

4.3 Verification of Optimality

In this section, it will be proven that the estimated direction, which is a linear combination of the observations of direction as unit vectors, $\boldsymbol{Uw}$, is also a unit vector, that the expression of constrained optimization of the cosine of the angle between the direction to be estimated and the estimator has derivatives of zero at the estimated direction, and that the estimated direction has a maximum fit as opposed to a minimum fit to the direction being estimated.

First, it will be shown that $\hat{\boldsymbol{u}}_0$ is a unit vector.

$$\begin{aligned}
\left\| \hat{\boldsymbol{u}}_0 \right\|^2 &\overset{(4.2)}{=} \boldsymbol{w}^T \boldsymbol{U}^T \boldsymbol{U} \boldsymbol{w} \\
&\overset{(4.16)}{=} \left(\frac{\boldsymbol{K}^{-1}\boldsymbol{c}}{\sqrt{\boldsymbol{c}^T \boldsymbol{K}^{-1} \boldsymbol{U}^T \boldsymbol{U} \boldsymbol{K}^{-1} \boldsymbol{c}}} \right)^T \boldsymbol{U}^T \boldsymbol{U} \frac{\boldsymbol{K}^{-1}\boldsymbol{c}}{\sqrt{\boldsymbol{c}^T \boldsymbol{K}^{-1} \boldsymbol{U}^T \boldsymbol{U} \boldsymbol{K}^{-1} \boldsymbol{c}}} \\
&= \frac{\boldsymbol{c}^T \left(\boldsymbol{K}^{-1} \right)^T}{\sqrt{\boldsymbol{c}^T \boldsymbol{K}^{-1} \boldsymbol{U}^T \boldsymbol{U} \boldsymbol{K}^{-1} \boldsymbol{c}}} \boldsymbol{U}^T \boldsymbol{U} \frac{\boldsymbol{K}^{-1}\boldsymbol{c}}{\sqrt{\boldsymbol{c}^T \boldsymbol{K}^{-1} \boldsymbol{U}^T \boldsymbol{U} \boldsymbol{K}^{-1} \boldsymbol{c}}} \\
&\overset{(B.5)}{=} \frac{\boldsymbol{c}^T \boldsymbol{K}^{-1} \boldsymbol{U}^T \boldsymbol{U} \boldsymbol{K}^{-1} \boldsymbol{c}}{\boldsymbol{c}^T \boldsymbol{K}^{-1} \boldsymbol{U}^T \boldsymbol{U} \boldsymbol{K}^{-1} \boldsymbol{c}} \\
&= 1 \Rightarrow
\end{aligned}$$

$$\left\| \hat{\boldsymbol{u}}_0 \right\|^2 = 1 \tag{4.17}$$

Thus, the squared length of the estimate of direction as a vector is one. Hence, the direction estimate is a unit vector such as the observations.

Next, it will be shown that the vector of derivatives of the expression of constrained optimization (4.12) at the solution $\boldsymbol{w}$ is the zero vector $\boldsymbol{0}$.

$$\begin{aligned}
\frac{\partial}{\partial \boldsymbol{w}} (q_s) &\overset{(4.13)}{=} \boldsymbol{c} - \nu\, s \boldsymbol{K} \boldsymbol{w} \\
&\overset{(4.15),(4.16)}{=} \boldsymbol{c} - \sqrt{s^{-1} \boldsymbol{c}^T \boldsymbol{K}^{-1} \boldsymbol{c}}\; s \boldsymbol{K} \left(\frac{\boldsymbol{K}^{-1}\boldsymbol{c}}{\sqrt{\boldsymbol{c}^T \boldsymbol{K}^{-1} \boldsymbol{U}^T \boldsymbol{U} \boldsymbol{K}^{-1} \boldsymbol{c}}} \right) \\
&= \boldsymbol{c} - s^{-1/2}\, s \boldsymbol{K} \left(\boldsymbol{K}^{-1}\boldsymbol{c} \Big/ \frac{\sqrt{\boldsymbol{c}^T \boldsymbol{K}^{-1} \boldsymbol{U}^T \boldsymbol{U} \boldsymbol{K}^{-1} \boldsymbol{c}}}{\sqrt{\boldsymbol{c}^T \boldsymbol{K}^{-1} \boldsymbol{c}}} \right) \\
&= \boldsymbol{c} - s^{-1/2}\, s \boldsymbol{K} \left(\boldsymbol{K}^{-1}\boldsymbol{c} \Big/ \sqrt{\frac{\boldsymbol{c}^T \boldsymbol{K}^{-1} \boldsymbol{U}^T \boldsymbol{U} \boldsymbol{K}^{-1} \boldsymbol{c}}{\boldsymbol{c}^T \boldsymbol{K}^{-1} \boldsymbol{c}}} \right) \\
&\overset{(4.11)}{=} \boldsymbol{c} - s^{-1/2}\, s \left(\frac{\boldsymbol{K}\boldsymbol{K}^{-1}\boldsymbol{c}}{\sqrt{s}} \right) \\
&= \boldsymbol{c} - s^{-1/2} s^{+1/2}\, \boldsymbol{c} \\
&= \boldsymbol{c} - \boldsymbol{c} \\
&= \boldsymbol{0} \Rightarrow
\end{aligned}$$

$$\frac{\partial}{\partial \boldsymbol{w}} (q_s) = \boldsymbol{0} \tag{4.18}$$

Thus, the vector of derivatives at the solution $\boldsymbol{w}$ is the zero vector, $\boldsymbol{0}$. Hence, the direction estimate has either a minimum or a maximum fit to the direction being estimated.

Last, it will be shown that the solution has a maximum fit. The quadratic part of (4.12) is $-0.5\nu\left(\boldsymbol{w}^T s\boldsymbol{K}\boldsymbol{w}\right) = \boldsymbol{w}^T\left(-0.5\nu\, s\boldsymbol{K}\right)\boldsymbol{w}$. $\boldsymbol{K}$ is positive definite (Chapter 3, Subsection 3.6.3) and symmetric (Appendix B, Section B.2), hence, $\boldsymbol{K}$ is orthonormally diagonalizable. Let $\boldsymbol{Q}$ be the diagonalizing matrix of eigenvectors of $\boldsymbol{K}$, and $\boldsymbol{\Lambda}$ be the diagonal matrix of the eigenvalues of $\boldsymbol{K}$. Hence, the diagonalization of the symmetric matrix $-0.5\nu\, s\boldsymbol{K}$ is $\boldsymbol{Q}^T\left(-0.5\nu\, s\boldsymbol{K}\right)\boldsymbol{Q} = -0.5\nu\, s\,\boldsymbol{Q}^T\boldsymbol{K}\,\boldsymbol{Q} = -0.5\nu\, s\,\boldsymbol{\Lambda}$. By Appendix B, Equation (B.3), the eigenvalues of $\boldsymbol{K}$, which are the elements λ_i of $\boldsymbol{\Lambda}$, are all positive. Therefore, the eigenvalues of $-0.5\nu\, s\,\boldsymbol{K}$, which are the diagonal elements of the matrix $-0.5\nu\, s\,\boldsymbol{\Lambda}$, are all negative. Since all the eigenvalues are negative, $-0.5\,\nu\, s\,\boldsymbol{K}$ is negative definite by equivalence (B.7). The Hessian of (4.12) is $-\nu\, s\,\boldsymbol{K}$ by Equation (B.11). Hence, the Hessian is also negative definite. The point of zero derivatives (4.18) has a negative definite Hessian. According to Appendix B, Subsection B.8.1, the point of zero derivatives is a maximum. Hence, the direction estimate has maximum fit to the direction being estimated.

4.4 Computationally Efficient Formula

From (4.1) and (4.16), $\hat{\boldsymbol{u}}_0 = \boldsymbol{U}\boldsymbol{K}^{-1}\boldsymbol{c} / \sqrt{\boldsymbol{c}^T\boldsymbol{K}^{-1}\boldsymbol{U}^T\boldsymbol{U}\boldsymbol{K}^{-1}\boldsymbol{c}}$. The denominator $\sqrt{\boldsymbol{c}^T\boldsymbol{K}^{-1}\boldsymbol{U}^T\boldsymbol{U}\boldsymbol{K}^{-1}\boldsymbol{c}}$ scales the vector $\boldsymbol{U}\boldsymbol{K}^{-1}\boldsymbol{c}$ to a unit vector, but does not affect the signs and the ratio of the magnitudes of the components of the vector $\boldsymbol{U}\boldsymbol{K}^{-1}\boldsymbol{c}$. Hence, computational efficiency may be obtained by eliminating the computation of

$\sqrt{\boldsymbol{c}^T\boldsymbol{K}^{-1}\boldsymbol{U}^T\boldsymbol{U}\boldsymbol{K}^{-1}\boldsymbol{c}}$ and computing direction directly from the components of $\boldsymbol{U}\boldsymbol{K}^{-1}\boldsymbol{c}$.

Let *h* and *v* be the horizontal and vertical components of the vector $\boldsymbol{U}\boldsymbol{K}^{-1}\boldsymbol{c}$, respectively, i.e., $\begin{bmatrix} h \\ v \end{bmatrix} = \boldsymbol{U}\boldsymbol{K}^{-1}\boldsymbol{c}$. Then, the estimated direction in $[0, 2\pi)$ radians at location $\boldsymbol{x}_0$ is

$$\hat{\theta}_0 = \begin{cases} \tan^{-1}(v/h), & h > 0, v \geq 0 \\ \pi/2, & h = 0, v > 0 \\ \tan^{-1}(v/h) + \pi, & h < 0 \\ \frac{3}{2}\pi, & h = 0, v < 0 \\ \tan^{-1}(v/h) + 2\pi, & h > 0, v < 0 \\ \text{undefined}, & h = v = 0. \end{cases} \tag{4.19}$$

This is called the quadrant specific inverse tangent as in Chapter 3, Subsection 3.3.1, (3.1).

4.5 Kriging Behavior Around a Sampled Location

The kriging behavior around an observation location depends on which cosine model is used. Figure 4-3, which was constructed using the R code in Appendix L, Section L.14, shows the kriging estimate in degrees around a direction of 90º observed at location 0 with nearest observations of 0º at a distance of 10 units away, which is the range. The curve of estimated direction from the spherical cosine model is dashed and red, the curve from the gaussian cosine model is tan and thick, and the curve from the exponential model is solid and black. The exponential curve of direction vs. location has a discontinuity in the derivative of direction with respect to location at the observation location. With nugget=0.0, the Gaussian and spherical curves appear smooth. With nugget = 0.1, the spherical curve is smooth, but the exponential and gaussian curves spike with a discontinuity at the observation location. With or without a nugget, the kriging solution produces “exact interpolation” at a sampled location (estimate=observation). Exact interpolation will now be proven.

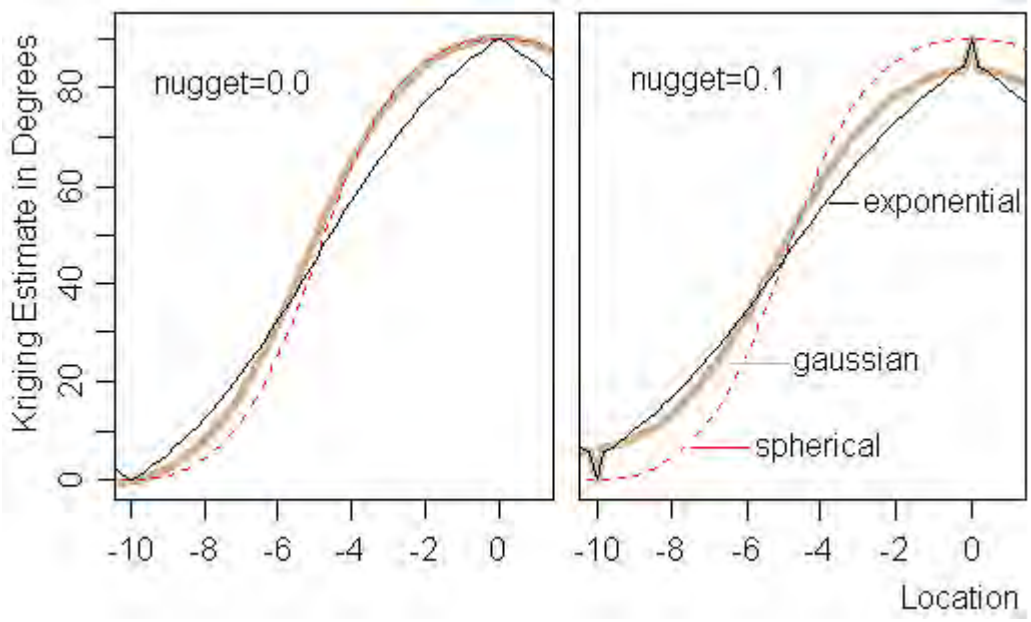

Figure 4-3. Effect of Cosine Model on the Kriging Estimate Around the Measurement Location. The curve from the spherical cosine model is dashed and red, the curve from the gaussian cosine model is tan and thick, and the curve from the exponential model is solid and black. Estimation at a sampled location produces exact interpolation.

With $\varsigma(d_{ij})$ being the mean cosine computed from the cosine model (Chapter 3, Subsection 3.6.2) on the distance from the i[th] to the j[th] observation d_{ij}, $i, j = 1,2,\cdots,n$, the unit vector estimate $\hat{\boldsymbol{u}}_j$ at sample location $\boldsymbol{x}_j$ requires the vector of cosines

$\boldsymbol{c}_j^T = \left[\varsigma(d_{1,j}) \;\; \varsigma(d_{2,j}) \;\; \varsigma(d_{3,j}) \;\; \cdots \;\; \varsigma(d_{j,j}) \;\; \cdots \;\; \varsigma(d_{n-1,j}) \;\; \varsigma(d_{n,j})\right]$ and the symmetric matrix of cosines

$$\boldsymbol{K} = \begin{bmatrix} \varsigma(d_{1,1})=1 & \varsigma(d_{1,2}) & \varsigma(d_{1,3}) & \cdots & \varsigma(d_{1,j}) & \cdots & \varsigma(d_{1,n-1}) & \varsigma(d_{1,n}) \\ \varsigma(d_{2,1}) & \varsigma(d_{2,2})=1 & \varsigma(d_{2,3}) & \cdots & \varsigma(d_{2,j}) & \cdots & \varsigma(d_{2,n-1}) & \varsigma(d_{2,n}) \\ \varsigma(d_{3,1}) & \varsigma(d_{3,2}) & \varsigma(d_{3,3})=1 & \cdots & \varsigma(d_{3,j}) & \cdots & \varsigma(d_{3,n-1}) & \varsigma(d_{3,n}) \\ \vdots & \vdots & \vdots & \cdots & \vdots & \cdots & \vdots & \vdots \\ \varsigma(d_{j,1}) & \varsigma(d_{j,2}) & \varsigma(d_{j,3}) & \cdots & \varsigma(d_{j,j})=1 & \cdots & \varsigma(d_{j,n-1}) & \varsigma(d_{j,n}) \\ \vdots & \vdots & \vdots & \cdots & \vdots & \cdots & \vdots & \vdots \\ \varsigma(d_{n-1,1}) & \varsigma(d_{n-1,2}) & \varsigma(d_{n-1,3}) & \cdots & \varsigma(d_{n-1,j}) & \cdots & \varsigma(d_{n-1,n-1})=1 & \varsigma(d_{n-1,n}) \\ \varsigma(d_{n,1}) & \varsigma(d_{n,2}) & \varsigma(d_{n,3}) & \cdots & \varsigma(d_{n,j}) & \cdots & \varsigma(d_{n,n-1}) & \varsigma(d_{n,n})=1 \end{bmatrix}.$$

The key to this proof is to observe that $\boldsymbol{c}_j$ is the j^{th} row or column of $\boldsymbol{K}$.

With $\boldsymbol{P}$ the diagonalizing orthonormal matrix of eigenvectors (Appendix B, Section B.2, point 4) of the positive definite matrix $\boldsymbol{K}$ (Chapter 3, Subsection 3.6.3), $\boldsymbol{\Lambda}$ the diagonal matrix of eigenvalues of $\boldsymbol{K}$ (Appendix B, Section B.2, point 4), $\boldsymbol{I}$ the diagonal matrix of 1s (the identity matrix), and $(\boldsymbol{M})_{\text{Col } j}$ the j^{th} column of matrix $\boldsymbol{M}$, and the fact (1) that the j^{th} column of the product of matrix $\boldsymbol{P}$ post multiplied by a matrix $\left(\boldsymbol{\Lambda}\boldsymbol{P}^T\right)$ equals the product of the matrix $(\boldsymbol{P}\boldsymbol{\Lambda})$ post multiplied by the j^{th} column of $\boldsymbol{P}^T$, the numerator of the solution vector $\boldsymbol{w}_j$ (4.16) is

$$
\begin{aligned}
\boldsymbol{K}^{-1}\boldsymbol{c}_j = & \\
&= \left(\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^T\right)^{-1}\boldsymbol{c}_j \\
&= \boldsymbol{P}\boldsymbol{\Lambda}^{-1}\boldsymbol{P}^T(\boldsymbol{K})_{\text{Col } j} \\
&= \boldsymbol{P}\boldsymbol{\Lambda}^{-1}\boldsymbol{P}^T\left(\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^T\right)_{\text{Col } j} \\
&\overset{(1)}{=} \boldsymbol{P}\boldsymbol{\Lambda}^{-1}\boldsymbol{P}^T\boldsymbol{P}\boldsymbol{\Lambda}\left(\boldsymbol{P}^T\right)_{\text{Col } j} \\
&= \boldsymbol{P}\boldsymbol{\Lambda}^{-1}\underbrace{\underbrace{\boldsymbol{P}^T\boldsymbol{P}}_{\boldsymbol{P}\text{ Orthonomal}}}_{=\boldsymbol{I}}\boldsymbol{\Lambda}\left(\boldsymbol{P}^T\right)_{\text{Col } j} \\
&= \boldsymbol{P}\underbrace{\boldsymbol{\Lambda}^{-1}\boldsymbol{I}\boldsymbol{\Lambda}}_{=\boldsymbol{I}}\left(\boldsymbol{P}^T\right)_{\text{Col } j} \\
&= \boldsymbol{P}\boldsymbol{I}\left(\boldsymbol{P}^T\right)_{\text{Col } j} \\
&\overset{(1)}{=} \left(\boldsymbol{P}\boldsymbol{P}^T\right)_{\text{Col } j} \\
&= (\boldsymbol{I})_{\text{Col } j} \\
&= [0 \quad 0 \quad 0 \quad \cdots \quad 1 \quad \cdots \quad 0 \quad 0]^T.
\end{aligned}
$$

So $\boldsymbol{K}^{-1}\boldsymbol{c}_j$ is a column vector of 0's with 1 in the j^{th} position. Using this result we will see that the estimated direction is the observed direction.

$$
\begin{aligned}
\hat{\boldsymbol{u}}_j &\overset{(4.1)}{=} \boldsymbol{U}\boldsymbol{w}_j \\
&\overset{(4.16)}{=} \boldsymbol{U}\frac{\boldsymbol{K}^{-1}\boldsymbol{c}_j}{\sqrt{\boldsymbol{c}_j^T\boldsymbol{K}^{-1}\boldsymbol{U}^T\boldsymbol{U}\boldsymbol{K}^{-1}\boldsymbol{c}_j}}
\end{aligned}
$$

$$= \frac{\begin{bmatrix}\boldsymbol{u}_1 & \boldsymbol{u}_2 & \boldsymbol{u}_3 & \cdots & \boldsymbol{u}_j & \cdots & \boldsymbol{u}_{n-1} & \boldsymbol{u}_n\end{bmatrix}\begin{bmatrix}0 & 0 & 0 & \cdots & 1 & \cdots & 0 & 0\end{bmatrix}^T}{\sqrt{\begin{bmatrix}0 & 0 & 0 & \cdots & 1 & \cdots & 0 & 0\end{bmatrix}\boldsymbol{U}^T\boldsymbol{U}\begin{bmatrix}0 & 0 & 0 & \cdots & 1 & \cdots & 0 & 0\end{bmatrix}^T}}$$

$$= \frac{\begin{bmatrix}\boldsymbol{u}_1 & \boldsymbol{u}_2 & \boldsymbol{u}_3 & \cdots & \boldsymbol{u}_j & \cdots & \boldsymbol{u}_{n-1} & \boldsymbol{u}_n\end{bmatrix}\begin{bmatrix}0 & 0 & 0 & \cdots & 1 & \cdots & 0 & 0\end{bmatrix}^T}{\sqrt{\left\{\begin{bmatrix}0 & 0 & 0 & \cdots & 1 & \cdots & 0 & 0\end{bmatrix}\boldsymbol{U}^T\right\}\left\{\boldsymbol{U}\begin{bmatrix}0 & 0 & 0 & \cdots & 1 & \cdots & 0 & 0\end{bmatrix}^T\right\}}}$$

$$= \frac{\boldsymbol{u}_j}{\sqrt{\boldsymbol{u}_j{}^T\boldsymbol{u}_j}}$$

$$= \frac{\boldsymbol{u}_j}{\sqrt{1}}$$

$$= \boldsymbol{u}_j$$

Thus, at a sampled location, the estimated direction equals the observed direction, which is a unit vector. This is called “exact interpolation” as in the kriging of linear RVs.

4.6 Circular Kriging Variance, σ^2_{CK}

In this section, let $\boldsymbol{U}$ be the random sample of directions. Let $\boldsymbol{u}_i, i = 1,2,\cdots n$ be the random direction at location $\boldsymbol{x}_i, i = 1,2,\cdots n$, $\hat{\boldsymbol{u}}_i$ be the estimator of $\boldsymbol{u}_i$, and $\boldsymbol{e}_i$ be the error vector of $\hat{\boldsymbol{u}}_i$ (Subsection 4.2.2, Figure 4-2). Let $\boldsymbol{u}_0$ be the random direction at unobserved (unsampled) location $\boldsymbol{x}_0$, $\hat{\boldsymbol{u}}_0$ be the estimator of $\boldsymbol{u}_0$, $\boldsymbol{e}_0$ be the random error vector equal to $\hat{\boldsymbol{u}}_0 - \boldsymbol{u}_0$ (Subsection 4.2.2, Figure 4-2), and Θ be the random angle between $\hat{\boldsymbol{u}}_0$ and $\boldsymbol{u}_0$. Let σ^2_{CK} be called the circular kriging variance and be defined as the mean squared length of the error vector, which is a measure of the variability of the circular kriging estimator.

At a sampled location, $\hat{\boldsymbol{u}}_i = \boldsymbol{u}_i$ (Section 4.5), $\boldsymbol{e}_i = \boldsymbol{0}$, and $\sigma^2_{CK} = \mathrm{E}\left\{\left\|\boldsymbol{e}_i\right\|^2\right\} = \mathrm{E}\left\{\left\|\boldsymbol{0}\right\|^2\right\} = 0$. If Θ were always the maximum of π, per Figure 4-2 $\left\|\boldsymbol{e}_0\right\|$ would always be 2 and σ^2_{CK} would always be 4. However, in a circular random field, Θ is random and cannot always be π. Hence, $0 \le \sigma^2_{CK} < 4$.

The estimate of the circular kriging variance will now be derived. With vector $\boldsymbol{c}$ and matrix $\boldsymbol{K}$ containing real-valued constants computed from the cosine model ς of circular-spatial correlation (Chapter 3, Subsection 3.6.2), let

$$\boldsymbol{X} = \boldsymbol{U}^T \boldsymbol{u}_0 \tag{4.20}$$

$$Y = \boldsymbol{c}^T \boldsymbol{K}^{-1} \boldsymbol{U}^T \boldsymbol{U} \boldsymbol{K}^{-1} \boldsymbol{c} \tag{4.21}$$

$$g(\boldsymbol{X}, Y) = \frac{\boldsymbol{X}}{\sqrt{Y}}. \tag{4.22}$$

Note that $\boldsymbol{X}$ is a random vector and Y is a random scalar. By the above definition of σ^2_{CK},

$$\begin{aligned}
\sigma^2_{CK} &= \mathrm{E}\left\{ \left\| \hat{\boldsymbol{u}}_0 - \boldsymbol{u}_0 \right\|^2 \right\} \\
&\overset{(4.4)}{=} \mathrm{E}\{ 2(1 - \cos(\Theta)) \} \\
&= 2\left(1 - \mathrm{E}\{\cos(\Theta)\}\right) \\
&\overset{(B.1)}{=} 2\left(1 - \mathrm{E}\left\{ \hat{\boldsymbol{u}}_0^T \boldsymbol{u}_0 \right\}\right) \\
&\overset{(4.1)}{=} 2\left(1 - \mathrm{E}\left\{ (\boldsymbol{U}\boldsymbol{w})^T \boldsymbol{u}_0 \right\}\right) \\
&= 2\left(1 - \mathrm{E}\left\{ \boldsymbol{w}^T \boldsymbol{U}^T \boldsymbol{u}_0 \right\}\right) \\
&\overset{(4.16)}{=} 2\left(1 - \mathrm{E}\left\{ \left(\frac{\boldsymbol{K}^{-1}\boldsymbol{c}}{\sqrt{\boldsymbol{c}^T \boldsymbol{K}^{-1} \boldsymbol{U}^T \boldsymbol{U} \boldsymbol{K}^{-1} \boldsymbol{c}}} \right)^T \boldsymbol{U}^T \boldsymbol{u}_0 \right\} \right) \\
&\overset{(B.5)}{=} 2 - 2\mathrm{E}\left\{ \frac{\boldsymbol{c}^T \boldsymbol{K}^{-1}}{\sqrt{\boldsymbol{c}^T \boldsymbol{K}^{-1} \boldsymbol{U}^T \boldsymbol{U} \boldsymbol{K}^{-1} \boldsymbol{c}}} \boldsymbol{U}^T \boldsymbol{u}_0 \right\} \\
&= 2 - 2\boldsymbol{c}^T \boldsymbol{K}^{-1} \mathrm{E}\left\{ \frac{\boldsymbol{U}^T \boldsymbol{u}_0}{\sqrt{\boldsymbol{c}^T \boldsymbol{K}^{-1} \boldsymbol{U}^T \boldsymbol{U} \boldsymbol{K}^{-1} \boldsymbol{c}}} \right\} \\
&\overset{(4.20),(4.21)}{=} 2 - 2\boldsymbol{c}^T \boldsymbol{K}^{-1} \mathrm{E}\left\{ \frac{\boldsymbol{X}}{\sqrt{Y}} \right\} \\
&\overset{(4.22)}{=} 2 - 2\boldsymbol{c}^T \boldsymbol{K}^{-1} \mathrm{E}\{ g(\boldsymbol{X}, Y) \} \Rightarrow
\end{aligned}$$

$$\sigma^2_{CK} = 2 - 2\boldsymbol{c}^T \boldsymbol{K}^{-1} \mathrm{E}\{ g(\boldsymbol{X}, Y) \}. \tag{4.23}$$

$g(\boldsymbol{X}, Y)$ is a nonlinear function of $\boldsymbol{X}$ and Y. When confronted with a nonlinear function, we can approximate using a method in probability and statistics called the

"delta method" or "propagation of errors" (Rice 1995, p. 149). With $(\boldsymbol{\mu}_{\boldsymbol{X}}, \mu_Y)$ being the expectation of $(\boldsymbol{X}, Y)$, $g(\boldsymbol{X}, Y)$ is approximated by a Taylor series about the fixed point $(\boldsymbol{\mu}_{\boldsymbol{X}}, \mu_Y)$ with expansion to the random point $(\boldsymbol{X}, Y)$. The expansion consists of one term which is nonrandom and random terms containing powers of the deltas $(\boldsymbol{X} - \boldsymbol{\mu}_{\boldsymbol{X}})$ and $(Y - \mu_Y)$. Hence, the method is called the "delta method."

The Taylor series expansion of $g(\boldsymbol{X}, Y)$ to a first order or linear approximation is

$$g(\boldsymbol{X}, Y) \approx g(\boldsymbol{\mu}_{\boldsymbol{X}}, \mu_Y) + (\boldsymbol{X} - \boldsymbol{\mu}_{\boldsymbol{X}}) \left\{ \left. \frac{\partial g(\boldsymbol{X}, Y)}{\partial \boldsymbol{X}} \right|_{(\boldsymbol{\mu}_{\boldsymbol{X}}, \mu_Y)} \right\} + (Y - \mu_Y) \left\{ \left. \frac{\partial g(\boldsymbol{X}, Y)}{\partial Y} \right|_{(\boldsymbol{\mu}_{\boldsymbol{X}}, \mu_Y)} \right\}.$$

This approximation improves as the joint probability of $\boldsymbol{X}$ and Y increases in an area of the domain of $g(\boldsymbol{X}, Y)$ where $g(\boldsymbol{X}, Y)$ is approximately linear.

Taking the expectation of the first order linear approximation,

$$\begin{aligned}
&\mathrm{E}\{g(\boldsymbol{X}, Y)\} \\
&\approx \mathrm{E}\left\{ g(\boldsymbol{\mu}_{\boldsymbol{X}}, \mu_Y) \;+\; (\boldsymbol{X} - \boldsymbol{\mu}_{\boldsymbol{X}}) \left\{ \left. \frac{\partial g(\boldsymbol{X}, Y)}{\partial \boldsymbol{X}} \right|_{(\boldsymbol{\mu}_{\boldsymbol{X}}, \mu_Y)} \right\} \;+\; (Y - \mu_Y) \left\{ \left. \frac{\partial g(\boldsymbol{X}, Y)}{\partial Y} \right|_{(\boldsymbol{\mu}_{\boldsymbol{X}}, \mu_Y)} \right\} \right\} \\
&= \mathrm{E}\left\{ \underbrace{g(\boldsymbol{\mu}_{\boldsymbol{X}}, \mu_Y)}_{\text{Constant}} \right\} + \mathrm{E}\left\{ \underbrace{(\boldsymbol{X} - \boldsymbol{\mu}_{\boldsymbol{X}})}_{\text{Random}} \underbrace{\left\{ \left. \frac{\partial g(\boldsymbol{X}, Y)}{\partial \boldsymbol{X}} \right|_{(\boldsymbol{\mu}_{\boldsymbol{X}}, \mu_Y)} \right\}}_{\text{Constant}} \right\} + \mathrm{E}\left\{ \underbrace{(Y - \mu_Y)}_{\text{Random}} \underbrace{\left\{ \left. \frac{\partial g(\boldsymbol{X}, Y)}{\partial Y} \right|_{(\boldsymbol{\mu}_{\boldsymbol{X}}, \mu_Y)} \right\}}_{\text{Constant}} \right\} \\
&= g(\boldsymbol{\mu}_{\boldsymbol{X}}, \mu_Y) \quad + \left\{ \left. \frac{\partial g(\boldsymbol{X}, Y)}{\partial \boldsymbol{X}} \right|_{(\boldsymbol{\mu}_{\boldsymbol{X}}, \mu_Y)} \right\} \mathrm{E}\{\boldsymbol{X} - \boldsymbol{\mu}_{\boldsymbol{X}}\} \quad + \left\{ \left. \frac{\partial g(\boldsymbol{X}, Y)}{\partial Y} \right|_{(\boldsymbol{\mu}_{\boldsymbol{X}}, \mu_Y)} \right\} \mathrm{E}\{Y - \mu_Y\} \\
&= g(\boldsymbol{\mu}_{\boldsymbol{X}}, \mu_Y) \quad + \left\{ \left. \frac{\partial g(\boldsymbol{X}, Y)}{\partial \boldsymbol{X}} \right|_{(\boldsymbol{\mu}_{\boldsymbol{X}}, \mu_Y)} \right\} \boldsymbol{0} \qquad + \left\{ \left. \frac{\partial g(\boldsymbol{X}, Y)}{\partial Y} \right|_{(\boldsymbol{\mu}_{\boldsymbol{X}}, \mu_Y)} \right\} 0 \\
&= g(\boldsymbol{\mu}_{\boldsymbol{X}}, \mu_Y) \quad + 0 \qquad + 0 \\
&= g(\mathrm{E}\{\boldsymbol{X}\}, \mathrm{E}\{Y\}) \Rightarrow
\end{aligned}$$

$$\mathrm{E}\{g(\boldsymbol{X}, Y)\} \approx g(\mathrm{E}\{\boldsymbol{X}\}, \mathrm{E}\{Y\}). \tag{4.24}$$

In general, $\mathrm{E}\{g(\boldsymbol{X}, Y)\} \neq g(\mathrm{E}\{\boldsymbol{X}\}, \mathrm{E}\{Y\})$. What (4.24) means is that $g(\mathrm{E}\{\boldsymbol{X}\}, \mathrm{E}\{Y\})$ is a first order linear approximation of $\mathrm{E}\{g(\boldsymbol{X}, Y)\}$.

Next, with $\varsigma(d_{ij})$ the expectation of the cosine of the angle between observations *i* and *j* estimated by the cosine model of circular-spatial correlation (Chapter 3, Section 3.6), $g(\mathrm{E}\{\boldsymbol{X}\},\mathrm{E}\{Y\})$ is evaluated.

$$g(\mathrm{E}\{\boldsymbol{X}\},\mathrm{E}\{Y\}) \overset{(4.22)}{=} \frac{\mathrm{E}\{\boldsymbol{X}\}}{\sqrt{\mathrm{E}\{Y\}}}$$

$$\overset{(4.20),(4.21)}{=} \frac{E\{\boldsymbol{U}^T\boldsymbol{u}_0\}}{\sqrt{E\{\boldsymbol{c}^T\boldsymbol{K}^{-1}\boldsymbol{U}^T\boldsymbol{U}\boldsymbol{K}^{-1}\boldsymbol{c}\}}} =$$

$$= E\left\{\begin{bmatrix}\boldsymbol{u}_1^T\\ \vdots\\ \boldsymbol{u}_n^T\end{bmatrix}\boldsymbol{u}_0\right\} \Big/ \sqrt{\boldsymbol{c}^T\boldsymbol{K}^{-1}E\{\boldsymbol{U}^T\boldsymbol{U}\}\boldsymbol{K}^{-1}\boldsymbol{c}}$$

$$= E\left\{\begin{bmatrix}\boldsymbol{u}_1^T\boldsymbol{u}_0\\ \vdots\\ \boldsymbol{u}_n^T\boldsymbol{u}_0\end{bmatrix}\right\} \Big/ \sqrt{\boldsymbol{c}^T\boldsymbol{K}^{-1}E\left\{\begin{bmatrix}1 & \boldsymbol{u}_1^T\boldsymbol{u}_2 & \cdots & \boldsymbol{u}_1^T\boldsymbol{u}_n\\ \boldsymbol{u}_2^T\boldsymbol{u}_1 & 1 & \cdots & \boldsymbol{u}_2^T\boldsymbol{u}_n\\ \vdots & \vdots & \vdots & \vdots\\ \boldsymbol{u}_n^T\boldsymbol{u}_1 & \boldsymbol{u}_n^T\boldsymbol{u}_2 & \cdots & 1\end{bmatrix}\right\}\boldsymbol{K}^{-1}\boldsymbol{c}}$$

$$\overset{(B.12)}{=} \begin{bmatrix}E\{\boldsymbol{u}_1^T\boldsymbol{u}_0\}\\ \vdots\\ E\{\boldsymbol{u}_n^T\boldsymbol{u}_0\}\end{bmatrix} \Big/ \sqrt{\boldsymbol{c}^T\boldsymbol{K}^{-1}\begin{bmatrix}E\{1\} & E\{\boldsymbol{u}_1^T\boldsymbol{u}_2\} & \cdots & E\{\boldsymbol{u}_1^T\boldsymbol{u}_n\}\\ E\{\boldsymbol{u}_2^T\boldsymbol{u}_1\} & E\{1\} & \cdots & E\{\boldsymbol{u}_2^T\boldsymbol{u}_n\}\\ \vdots & \vdots & \vdots & \vdots\\ E\{\boldsymbol{u}_n^T\boldsymbol{u}_1\} & E\{\boldsymbol{u}_n^T\boldsymbol{u}_2\} & \cdots & E\{1\}\end{bmatrix}\boldsymbol{K}^{-1}\boldsymbol{c}}$$

$$= \begin{bmatrix}\varsigma(d_{01})\\ \vdots\\ \varsigma(d_{0n})\end{bmatrix} \Big/ \sqrt{\boldsymbol{c}^T\boldsymbol{K}^{-1}\underbrace{\begin{bmatrix}\varsigma(d_{1,1}) & \varsigma(d_{1,2}) & \cdots & \varsigma(d_{1,n})\\ \varsigma(d_{2,1}) & \varsigma(d_{2,2}) & \cdots & \varsigma(d_{2,n})\\ \vdots & \vdots & \vdots & \vdots\\ \varsigma(d_{n,1}) & \varsigma(d_{n,2}) & \cdots & \varsigma(d_{n,n})\end{bmatrix}}_{\boldsymbol{K},\text{p. 89}}\boldsymbol{K}^{-1}\boldsymbol{c}}$$

$$= \begin{bmatrix}c_{01}\\ \vdots\\ c_{0n}\end{bmatrix} \Big/ \sqrt{\boldsymbol{c}^T\boldsymbol{K}^{-1}\underbrace{\boldsymbol{K}\boldsymbol{K}^{-1}}_{\boldsymbol{I}}\boldsymbol{c}}$$

$$= \boldsymbol{c}\Big/\sqrt{\boldsymbol{c}^T\boldsymbol{K}^{-1}\boldsymbol{c}}$$

$$\overset{(4.24)}{\approx} \mathrm{E}\{g(\boldsymbol{X},Y)\} \Rightarrow$$

$$\mathrm{E}\{g(\boldsymbol{X},Y)\} \approx \frac{\boldsymbol{c}}{\sqrt{\boldsymbol{c}^T\boldsymbol{K}^{-1}\boldsymbol{c}}} \tag{4.25}$$

Last, the first order linear approximation of the circular kriging variance will be completed.

$$\begin{aligned}
\sigma^2_{CK} &\overset{(4.23)}{=} 2 - 2\boldsymbol{c}^T\boldsymbol{K}^{-1}\mathrm{E}\{g(\boldsymbol{X},Y)\} \\
&\overset{(4.25)}{\approx} 2 - 2\boldsymbol{c}^T\boldsymbol{K}^{-1}\frac{\boldsymbol{c}}{\sqrt{\boldsymbol{c}^T\boldsymbol{K}^{-1}\boldsymbol{c}}} \\
&= 2 - 2\frac{\boldsymbol{c}^T\boldsymbol{K}^{-1}\boldsymbol{c}}{\sqrt{\boldsymbol{c}^T\boldsymbol{K}^{-1}\boldsymbol{c}}} \\
&= 2 - 2\sqrt{\boldsymbol{c}^T\boldsymbol{K}^{-1}\boldsymbol{c}} \\
&\equiv \hat{\sigma}^2_{CK} \Rightarrow
\end{aligned}$$

$$\hat{\sigma}^2_{CK} = 2 - 2\sqrt{\boldsymbol{c}^T\boldsymbol{K}^{-1}\boldsymbol{c}} \tag{4.26}$$

4.7 How Distance and the Cosine Model Affect $\hat{\sigma}^2_{CK}$

When direction is estimated at a sample location $\boldsymbol{x}_j$, $j = 1,2,\cdots n,$ the vector of mean cosines $\boldsymbol{c}_j$ is the j^{th} row or column of the positive definite matrix of mean cosines $\boldsymbol{K}$ (Section 4.5). From p. 90, we know that $\boldsymbol{K}^{-1}\boldsymbol{c}_j = \begin{bmatrix}0 & 0 & 0 & \cdots & 1 & \cdots & 0 & 0\end{bmatrix}^T$ with 1 in the jth position, hence,

$$\begin{aligned}
\boldsymbol{c}_j^T\boldsymbol{K}^{-1}\boldsymbol{c}_j &= \boldsymbol{c}_j^T\begin{bmatrix}0 & 0 & 0 & \cdots & \underbrace{1}_{\text{Position } j} & \cdots & 0 & 0\end{bmatrix}^T \\
&= 1 \Rightarrow \\
\hat{\sigma}^2_{CK}(\boldsymbol{x}_j) &\overset{(4.26)}{=} 2 - 2\sqrt{\boldsymbol{c}_j^T\boldsymbol{K}^{-1}\boldsymbol{c}_j} \\
&= 2 - 2\sqrt{1} \\
&= 0 \Rightarrow
\end{aligned}$$

$$\hat{\sigma}^2_{CK}(\boldsymbol{x}_j) = 0. \tag{4.27}$$

Further, in Section 4.5, it was proven that the estimate of direction at a sampled location is the observed direction. Then, at a sampled location, the error vector

$\boldsymbol{e}_j, j = 1,2,\cdots,n$ is always the zero vector **0** (Figure 4-2). Hence, at a sampled location, $\sigma^2_{CK} = 0$, and $\hat{\sigma}^2_{CK} = \sigma^2_{CK} = 0$. In contrast to result (4.26), McNeill (1993, p. 46) obtained $\sqrt{\boldsymbol{c}^T \boldsymbol{K}^{-1} \boldsymbol{c}}$ for the circular kriging variance. By (4.26), $\hat{\sigma}^2_{CK}$ decreases when $\sqrt{\boldsymbol{c}^T \boldsymbol{K}^{-1} \boldsymbol{c}}$ increases. Hence, McNeill's result is a measure of concentration, which is opposite the sense of spread or variance.

The model parameters are the range, mean resultant length ρ of the circular distribution component of the CRF, and the nugget n_g from measurement error and close sampling. To see their effects, Figure 4-4 was constructed using the R code in Appendices K.18 and L.14, which simulates observation one with location at the origin of location coordinates, observation two with location at 1 unit of distance due north of observation one, and the estimation location at a variable distance due east of observation one. Hence, zero distance corresponds to estimation at the location of observation one.

The shape of the $\hat{\sigma}^2_{CK}$ curve resembles the inverted curve of the circular-spatial correlation model identified in the legend. The spherical curve attains the maximum at distances greater than and equal to the range. The exponential and gaussian curves appear to be asymptotic with the gaussian curve exceeding the exponential curve at a distance approximately equal to the range.

The upper left and lower right plots indicate that increasing the range increases the distance at which the maximum $\hat{\sigma}^2_{CK}$ occurs. The upper plots indicate that when ρ is increased (the distribution of the CRV is more concentrated), the estimator of direction, which depends on the variability of the CRV, is also more concentrated, and hence the $\hat{\sigma}^2_{CK}$ is reduced. The left plots indicate that the nugget introduces a

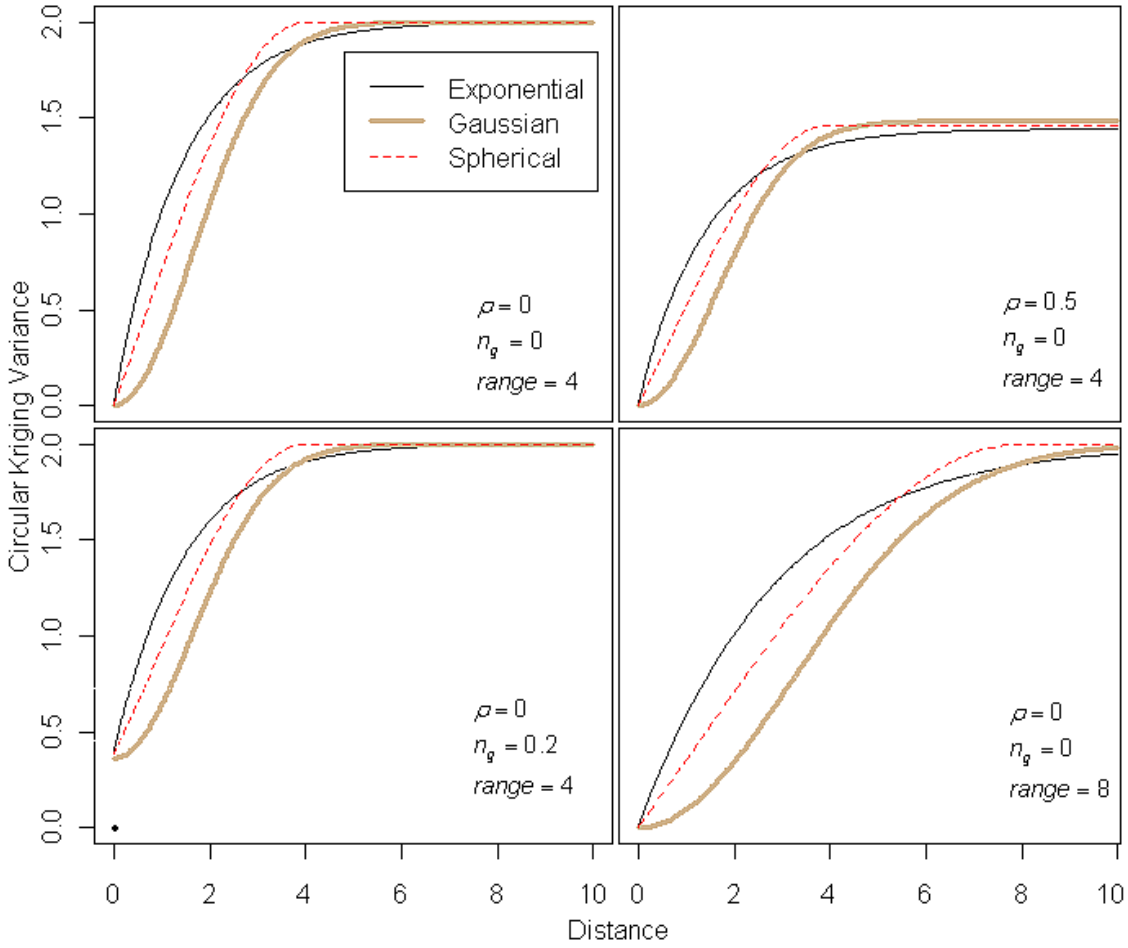

Figure 4-4. Effect of Range, Mean Resultant Length ρ, and nugget n_g on the Circular Kriging Variance $\hat{\sigma}^2_{CK}$.

discontinuity at zero distance and increases $\hat{\sigma}^2_{CK}$ at small distances relative to the range. The maximum precision of the estimator at nonzero distances is limited by the nugget.

4.8 Chapter Summary and Future Work

The principal result of this chapter is the vector of weights $\boldsymbol{w}$ of the circular kriging solution, which is

$$\boldsymbol{w} = \boldsymbol{K}^{-1}\boldsymbol{c} \Big/ \sqrt{\boldsymbol{c}^T \boldsymbol{K}^{-1}\boldsymbol{U}^T \boldsymbol{U}\boldsymbol{K}^{-1}\boldsymbol{c}}\,.$$

It was derived in full detail, and shown to be optimal producing a unit vector estimate (4.17) at a stationary point (4.18) of maximum fit (Section 4.3). The approach avoided the first order Taylor series approximation of McNeill (1993, p. 39), which results in a nonunit vector estimator of direction (4.11).

A computationally efficient form of the estimator of direction (4.19) was derived with elimination of the scalar function $\sqrt{\boldsymbol{c}^T \boldsymbol{K}^{-1}\boldsymbol{U}^T \boldsymbol{U}\boldsymbol{K}^{-1}\boldsymbol{c}}$ in the denominator of the kriging solution. The scalar function does not affect the signs and the ratio of the magnitudes of the vector components. With *h* and *v* being the horizontal and vertical components of the vector $\boldsymbol{U}\boldsymbol{K}^{-1}\boldsymbol{c}$, respectively, i.e., $\begin{bmatrix} h \\ v \end{bmatrix} = \boldsymbol{U}\boldsymbol{K}^{-1}\boldsymbol{c}$, the estimated direction in $[0, 2\pi)$ radians at location $\boldsymbol{x}_0$ is

$$\hat{\theta}_0 = \begin{cases} \tan^{-1}(v/h), & h > 0,\ v \geq 0 \\ \pi/2, & h = 0,\ v > 0 \\ \tan^{-1}(v/h) + \pi, & h < 0 \\ \frac{3}{2}\pi, & h = 0,\ v < 0 \\ \tan^{-1}(v/h) + 2\pi, & h > 0,\ v < 0 \\ \text{undefined}, & h = v = 0. \end{cases}$$

The estimated direction at a sampled location was proven to be the observed direction. This is called “exact interpolation” as in the kriging of linear RV.

The variability of the circular kriging estimator σ_{CK}^2 was defined as the mean squared length of the error vector, $0 \le \sigma_{CK}^2 < 4$. The first order linear approximation was derived (4.26) as

$$\hat{\sigma}_{CK}^2 = 2 - 2\sqrt{\boldsymbol{c}^T \boldsymbol{K}^{-1} \boldsymbol{c}}\,.$$

It was proven that the circular kriging variance at a sampled location is zero. The effects on $\hat{\sigma}_{CK}^2$ of the distance to an observation and the cosine model parameters were shown.

Future work includes derivation of $\hat{\sigma}_{CK}^2$ to a higher order approximation to increase the accuracy. A nonzero nugget n_g such as from measurement error has the effect of smoothing the estimates at locations where data does not exist and not smoothing the estimates at observation locations (exact interpolation). This suggests deriving a circular kriging solution for estimation “without measurement error” as in linear kriging where the smoothing of estimates at all locations varies with the magnitude of the nugget.