Results:

SOLUTIONS

Midterm Exam #1

1. (20 pts.) The following isotropic semivariogram model is called the *rational quadratic* model:

$$\gamma(r;\theta) = \frac{\theta_1 r^2}{1+\theta_2 r^2},$$

with parameters $\theta_1 \geq 0$ and $\theta_2 \geq 0$. Determine numerical values or expressions (simplified as much as possible) for each of the following quantities for this model, in terms of its parameters:

(a) the nugget effect

$$\lim_{r \to 0} \gamma(r) = \frac{\theta_1(0^2)}{1+\theta_2(0^2)} = 0.$$

(b) the sill

$$\lim_{r\to\infty} \gamma(r) = \lim_{r\to\infty} \frac{\theta_1}{\frac{1}{r^2}+\theta_2} = \frac{\theta_1}{\theta_2}, \quad \theta_1 \geq 0, \quad \theta_2 \geq 0.$$

(c) the effective range (i.e., the distance at which the semivariogram reaches 95% of its sill)

Solve the equation for r:

$$0.95\,\frac{\cancel{\theta_1}}{\theta_2} = \frac{\cancel{\theta_1} r^2}{1+\theta_2 r^2}$$

$$0.95(1+\theta_2 r^2) = \theta_2 r^2$$

$$0.95 = 0.05\,\theta_2 r^2$$

$$r^2 = \frac{19}{\theta_2}$$

$$r \approx \frac{4.36}{\sqrt{\theta_2}}$$

2. (15 pts.) Consider the covariance function

$$C(h_1, h_2, \ldots, h_d) = \exp\left(-\sum_{i=1}^{d} \frac{h_i^2}{\theta_i}\right) \quad -\infty < h_i < \infty \quad (i = 1, \ldots, d).$$

(a) Explain why this is a valid covariance function in d dimensions. (Hint: Show how this function may be built from other functions that you know to be valid covariance functions in one dimension.)

$$C(h_1, h_2, \ldots, h_d) = \exp\left(-\frac{h_1^2}{\theta_1}\right) \cdot \exp\left(-\frac{h_2^2}{\theta_2}\right) \cdots \exp\left(-\frac{h_d^2}{\theta_d}\right)$$

It is a product of 1-dim Gaussian covariance model, $C(r) = \theta_1^* \exp\left(-\frac{r^2}{\theta_2^*}\right)$

with $\theta_1^* = 1$, $\theta_2^* = \theta_1, \theta_2, \ldots, \theta_d$.

Thus, the product of d valid 1-dim covariance model's is a valid covariance func in d dimensions.

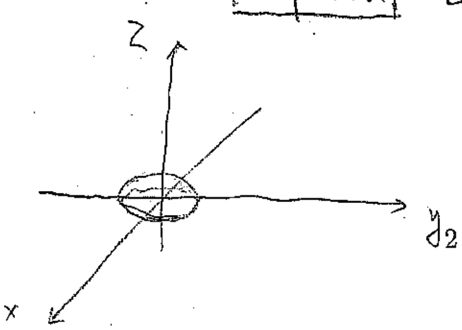
(b) Consider the special case of this covariance function for which $d = 3$, $\theta_1 = \theta_2$, and $\theta_2 \neq \theta_3$. What is the shape of the isocorrelation contours of this special case? Be as specific as possible.

$$C(h_1, h_2, h_3) = \exp\left(-\left[\frac{h_1^2}{\theta_1} + \frac{h_2^2}{\theta_1} + \frac{h_3^2}{\theta_3}\right]\right).$$

The isocorrelation contour would be for any arbitrary constant

$C > 0$, $\quad \dfrac{h_1^2 + h_2^2}{\theta_1} + \dfrac{h_3^2}{\theta_3} = C$, where $\theta_1 \neq \theta_3$.

It is a $\boxed{\text{ellipsoid}}$ centered at the origin.



horizontal cut-off surface is a circle.

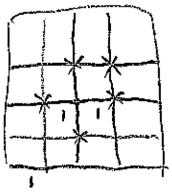
$r = \theta_1 (3r - 4r^3)$

$0 < r \le 0.5$

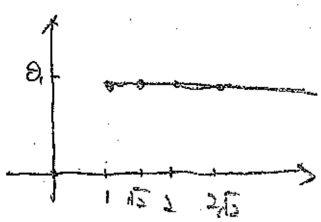
3. (20 pts.) Consider a geostatistical process (random field) that is second-order stationary and has a spherical semivariogram with range 0.5 units. Suppose further that the random field is observed at data locations which form a square grid with grid spacing 1.0 units.

Lag $= 1, \sqrt{2}, 2, 2\sqrt{2}$ ...

(a) Apart from chance fluctuations (noise), what would the sample semivariogram tend to look like? You may sketch a picture if you find it easier to answer the question this way than using words.



Since the range < grid spacing, we would only observe a relative flat trend of the sample semivariogram, i.e



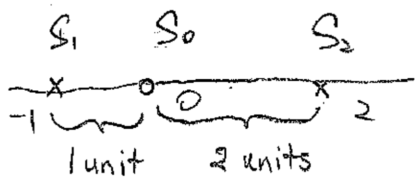
$\rightsquigarrow$ More like a pure nugget effect model

(b) Based on your answer to part (a), explain how the ordinary kriging predictor, obtained by replacing the true semivariogram with the sample semivariogram in the ordinary kriging equations, would be (adversely) affected when predicting at sites close to a data location.

→ If the prediction sites are not necessarily on the grid, then based on (a), the sample semivariogram looks more like a pure nugget effect model. Then the predictor will have $\boxed{\text{equal weights}}$ regardless of the data location, since there's no spatial correlation. It's less accurate and $\hat{\sigma}^2_{ok}$ will be greater.

→ However, if the prediction sites are on the grid, then it won't change too much whether we use the true semivariogram or the sample semivariogram. The spatial correlation decays at lag = 1 unit any way.

4. (20 pts.) Consider the spatial set-up in the following diagram. Here, an $\times$ marks the locations of each of two observed values of an attribute variable, and $\circ$ marks the location of an unobserved value of the same variable which we would like to predict.



The attribute variable is intrinsically stationary, with semivariogram $\gamma(r) = r$, for $r \geq 0$ (i.e., a linear semivariogram model with nugget 0 and slope 1).

(a) Write out the ordinary kriging equations for this prediction scenario; that is, write out the equations in terms of matrices and vectors, and give the numerical elements of those matrices/vectors where possible.

$$\Gamma_0 = \begin{pmatrix} 0 & 3 & 1 \\ 3 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \qquad \gamma_0 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$$

$\Gamma_0 \lambda_0 = \gamma_0$, solve for λ_0

$$\begin{pmatrix} 0 & 3 & 1 \\ 3 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} -\frac{1}{6} & \frac{1}{6} & \frac{1}{2} \\ \frac{1}{6} & -\frac{1}{6} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{3}{2} \end{pmatrix} = \begin{pmatrix} 1 & & \\ 0 & 1 & \\ 0 & & 1 \end{pmatrix}$$

(b) The inverse of the matrix on the left-hand side of the ordinary kriging equations is

$$\begin{pmatrix} -\frac{1}{6} & \frac{1}{6} & \frac{1}{2} \\ \frac{1}{6} & -\frac{1}{6} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{3}{2} \end{pmatrix}.$$

$\frac{1}{6} + \frac{3}{6} = \frac{2}{3}$

Using this information, obtain the ordinary kriging weights for the two data locations and the ordinary kriging variance corresponding to the ordinary kriging predictor.

$$\lambda_0 = \Gamma_0^{-1} \gamma_0 = \begin{pmatrix} -\frac{1}{6} & \frac{1}{6} & \frac{1}{2} \\ \frac{1}{6} & -\frac{1}{6} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{3}{2} \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} -\frac{1}{6} + \frac{2}{6} + \frac{1}{2} \\ \frac{1}{6} - \frac{2}{6} + \frac{1}{2} \\ \frac{1}{2} + \frac{2}{2} - \frac{3}{2} \end{pmatrix} = \begin{pmatrix} \frac{2}{3} \\ \frac{1}{3} \\ 0 \end{pmatrix}$$

$\lambda_1 = \frac{2}{3}, \quad \lambda_2 = \frac{1}{3}.$

$$\sigma^2_{ok} = \lambda_0' \gamma_0 = \left(\frac{2}{3} \quad \frac{1}{3} \quad 0\right) \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = \frac{2}{3} + \frac{2}{3} = \boxed{\frac{4}{3}}$$

5. (25 pts.) Consider the following set of geostatistical data, taken on a 5×5 square grid of points:

5	7	9	11	13	(4,4)
4	6	8	10	12	
3	5	7	$\boxed{9}$	11	
2	4	6	8	10	
1	3	5	7	9	
(0,0)	(1,0)				

$y = \beta_0 + \beta_1 u + \beta_2 v + \varepsilon(u, v).$

(a) Give the equation of a first-order polynomial function that would fit these data perfectly. You may assume that the origin of a two-dimensional Cartesian coordinate system coincides with the lower left-hand corner of the grid.

$$y = 1 + 2x + y + \varepsilon.$$

(b) Assuming that the data are a realization from a stationary process, obtain the sample semivariogram at lags (1,0), (0,1), (1,1), and (1,-1). (Hint: Don't waste time doing a lot of messy algebra — exploit the patterns in the data!)

h_u	$N(h_u)$	$\gamma(h_u)$
(1,0)		$\frac{2^2}{2} = 2$
(0,1)		$\frac{1^2}{2} = \frac{1}{2}$
(1,1)		$\frac{3^2}{2} = \frac{9}{2}$
(1,-1)		$\frac{1^2}{2} = \frac{1}{2}$

(c) Based on your answer to part (b), would it be unreasonable to assume that the underlying process is (i) isotropic, (ii) geometrically anisotropic, (iii) zonally anisotropic? Explain.

We observe $\hat{\gamma}(h)$ is different between (0,1) and (1,0); (1,1) and (1,-1). Thus we assume anisotropic.

But it's unusal to use zonally anisotropic.

thus, it's more reasonable to assume $\boxed{\text{geometrically anisotropic}}$

the spatial correlation in N-S direction is twice as strong as in the E-W direction.