ST 733 Practice Problems for Final Exam

Spring 2005

Final Exam on March 3, 2005

1. Consider a small data set:

   $Z[1,1] = 1, Z[2,2] = 2, Z[1,4] = 4, Z[3,1] = 3$ and $Z[3,4] = 20$

   (a) Compute the classical semivariogram estimate as lag distances $\sqrt{2}, 2, \sqrt{5}, 3$ and $\sqrt{13}$.

   (b) What is the impact of the extreme observation $Z[3,4]$?

   (c) Remove the observation $Z[3,4]$ and compute the semivariogram estimate at lag distances $2, \sqrt{5}, 3$ and $\sqrt{13}$.

   (d) Based on the results in (a) and (c) above would you delete the extreme observation? Justify your answer.

   (e) Compute the robust semivariogram estimate based on all observations. Compare this robust semivariogram estimate with that obtained in (a).

   Hint: Read pages 158-161 of the optional textbook

2. Consider the following data:

   $Z[0,4] = 5, Z[4,0] = 10, Z[0,8] = 15$ and $Z[4,8] = 6$.

   Suppose we want to predict at four sites $(2,4), (3,2), (4,0)$ and $(5,-2)$.

   (a) Plot the observed sites along with the sites to be predicted on the same graph.

   (b) Consider two isotropic exponential semivariogram models with:

   A: practical range 8.5, sill 1.0 and no nugget

   B: practical range 8.5, partial sill 0.5 and nugget 0.5.

   Let the model be represented as: $Z[s_i] = \mu + \epsilon_i$.

   i. Obtain GLS estimate of μ under models A and B.

   ii. Obtain the kriging weights under models A and B for each of the four target prediction sites.

iii. Obtain the prediction values under models A and B at each of the four target prediction sites.

iv. Obtain the prediction variance under models A and B at each of the four target prediction sites.

(c) Compare the prediction values at site $(4, 0)$ under models A and B. Compare prediction values to the observed value at $(4, 0)$?

(d) Compare the prediction variances at site $(4, 0)$ under models A and B.

(e) Is kriging a perfect inperpolator?

(f) Consider model A. Compare the kriging weights at site $(2, 4)$ to kriging weigths at other sites in terms of their standard deviations. In other words, compute the standard deviation of the (four) kriging weights at four sites.

(g) Consider model B. Compare the kriging weights at site $(2, 4)$ to kriging weigths at other sites in terms of their standard deviations.

(h) Is it true that if a prediction site is equally close to (or far from) all observed sites, then the kriging weights are all equal (and hence the standard deviation of the kriging weights is zero)?

(i) Is it true that if a site that is closest to a given prediction site gets the highest kriging weights?

Hint: Read section 5.4.3 of the optional textbook

3. Prove the prediction result:

   For any random vector U and any random variable Y, we have either of the following for every function g:

   (a) $E[(Y - g(U))^2] = \infty$

   (b) $E[(Y - E[Y|U])^2] \leq E[(Y - g(U))^2]$ *with equality only if* $g(U) = E[Y|U]$.

4. Consider prediction under squared error loss. Let $p_0(Z, s_0) = E[Z(s_0)|Z]$. Establish that,

$$E[(Z(s_0) - p_0(Z, s_0))^2] = Var[Z(s_0)] - Var[p_0(Z, s_0)]$$

5. Consider the following two models:

   A: $Z_i = \mu + \epsilon_{Ai},\ E[\epsilon_{Ai}] = 0, Cov[\epsilon_{Ai}, \epsilon_{Aj}] = \sigma^2 I(i = j)$

   B: $Z_i = \mu + \epsilon_{Bi},\ E[\epsilon_{Bi}] = 0, Cov[\epsilon_{Bi}, \epsilon_{Bj}] = \sigma^2 \rho^{I(i \neq j)}$

   where $I(C) = 1$ if the statement C is correct and $I(C) = 0$ otherwise.

   (a) Obtain the OLS estimator of μ under model A.

(b) Compute the variance of the OLS estimator obtained in (a).

(c) Obtain the GLS estimator of μ under model B.

(d) Compute the variance of the GLS estimator obtained in (b). Compare this variance with that in (b) when $\rho = 0$. What happens to this variance when $\rho = 1$? If you use the OLS estimator obtained in (a) under model B, what is the variance of this estimator under model B?

(e) Suppose you want to predict a new observation Y_0.

i. Obtain the best linear unbiased predictor (BLUP) of Y_0 under model A.

ii. Compute the variance of BLUP obtained in i.

iii. Obtain the best linear unbiased predictor (BLUP) of Y_0 under model B.

iv. Compute the variance of BLUP obtained in iii.

v. Compare the variances in ii. and iv.

Hint: Read section 1.5 of the optional textbook

6. Show that a valid semivariogram of an intrinsically stationary process, is conditionally negative definite, i.e.,

$$\sum_{i=1}^{m}\sum_{j=1}^{m} a_i a_j \gamma(s_i - s_j) \leq 0,$$

for any real numbers $a_1, \ldots, a_m$ such that $\sum_{i=1}^{m} a_i = 0$.
Recall that $\gamma(s_i - s_j) = \frac{1}{2}E[(Z(s_i) - Z(s_j))^2]$.

7. Suppose that $\{Z(s),\ s \in D\}$ is a weakly stationary Gaussian process with semivariance function, $\gamma(h) = \frac{1}{2}E[(Z(s+h) - Z(s))^2]$.

(a) Show that $\frac{(Z(s+h)-Z(s))^2}{2\gamma(h)} \sim \chi^2_1$.

(b) Use (a) to deduce that, $Var[(Z(s+h) - Z(s))^2] = 8\gamma(h)^2$.

(c) Use (b) to justify the approximate formula $Var[\hat{\gamma}(h)] \approx 2\frac{\gamma(h)^2}{N(h)}$, where $N(h)$ denotes the number of sites at a lag h.

Hint: Read section 4.4 of the optional textbook

8. Read section 5.2 of the optional text. Recreate Table 5.1 using `geoR` based on the data set given in example 5.5 (p.229).

9. Review all HW problems especially HW7 and HW8 that involves block kriging and cokriging, respectively.