Homework problems
(more problems will be added as we go along)

Stat 591, A17

1. For a stationary random field $Z(s); \ s \in D \subseteq \mathbb{R}^2$, observed at sites $s_1, \ldots, s_n$, derive the unbiased linear estimator with the smallest variance.

   *Hint:* Use a Lagrange multiplier to enforce the unbiasedness conditions.

2. For a stationary random field $Z(s); \ s \in D \subseteq \mathbb{R}^2$, observed at sites $s_1, \ldots, s_n$, show that the universal kriging estimator for $A(s) = a^T \begin{pmatrix} 1 \\ s \end{pmatrix}$ is unbiased.

3. Compare the variability of simple and ordinary kriging (this can either be done theoretically or by designing an appropriate simulation study).

4. Write R-code to put contours of a kriged surface on a grey-scale background of kriging standard errors.

5. Design a study to compare the plug-in estimate of kriging variance to the real variance of the predictor at a single point. (You do not need to implement the study, just execute a thoughtful design—see also problem 6).

6. (For those who did problem 5). Implement your study from problem 5.

7. Show that a $d$-dimensional isotropic correlation functions satisfies $\rho(v) \geq -\frac{1}{d}$.

8. Consider the correlation function $\rho(v) = 1 - v / \phi; v \leq \phi$. This is a valid correlation function in one dimension. Show that it is not valid in two dimensions.

   *Hint:* Consider points $s_{ij}$ at a 6x8 grid of size $\phi / \sqrt{2}$. Look at $\text{Var} \sum a_{ij} Z(s_{ij})$ where $a_{ij} = 1$ if $i+j$ even, -1 otherwise.

9. Compare several spatial covariance models graphically, by choosing parameters so that the range/effective range, sill and the nugget are the same for all models.

10. Consider a 2-dimensional Gaussian process in the plane with known mean

    $\mathbf{\mu} = (\mu_1, \mu_2)^T$ and covariance structure $C(h) = \begin{pmatrix} C_{11}(h) & C_{12}(h) \\ C_{21}(h) & C_{22}(h) \end{pmatrix}$.

    (a) Find the kriging estimate of the process at a point $s_0$.

    (b) If $s_0$ is one of the points of observation, under what circumstances is the kriging estimate at that point equal to the observation?
11. Develop R code that links the variogram cloud points to the geographic map, so that clicking on a point in the cloud scatter highlights the two corresponding sites, and clicking on a site highlights all the scatter points including that site.

12. For an isotropic Higdon model with kernel $(2\pi \phi)^{-\frac{1}{2}} \exp(-t^2/(2\phi))$, determine the covariance between two locations.

13. Consider iid Gaussian random variables $\varepsilon_i$ and define a spatial process $Z_i$ on a finite lattice by

$$Z_i - \theta_1(Z_{N(i)} + Z_{S(i)}) - \theta_2(Z_{E(i)} + Z_{W(i)}) = \varepsilon_i,$$

where $N(i)$ is the northern neighbor of $i$, etc. This is called a Gaussian simultaneously specified autoregression (SAR, Whittle (1954)).

(a) Show that for suitable $b_{ij}$

$$Z_i = \mu_i + \sum b_{ij}(Z_j - \mu_j) + \varepsilon_i.$$

(b) Show that $Z \sim N(\mu, \sigma^2((I - B)^{-1}((I - B^T)^{-1}))$ where $B = (b_{ij})$.

(c) Show that this model is equivalent to a CAR process.

14. Using pseudolikelihood, determine how to estimate the parameters of an Ising model on a finite rectangle lattice.

15. Apply the method in problem 14 to the data in http://www.stat.washington.edu/peter/book.data/set5 indicating the location of spotted wilt (a disease of a tomato plant) on a 24 by 60 field of plants (for example, the second plant from the right in row 1 is diseased).

16. (Borrowed from Brian Reich, NCSU) The data set http://www4.stat.ncsu.edu/~reich/st733/NARCCAP.RData contains, for each of 802 spatial locations in eastern US, elevation, lat/long, and median annual maximum precipitation from 1968-1999 ($Y_{\text{past}}$) and from 2038-2070 ($Y_{\text{future}}$), obtained from a regional climate model.

(a) Fit a trend in terms of $Y_{\text{future}}$ and elevation.

(b) Add appropriate functions of latitude and longitude to the model in (a).

Which model do you prefer, and why?

17. Fit a Matérn covariance to the data set in problem 16. Do you think an exponential covariance would be sufficient? Is there evidence of a nugget?

You may also submit solutions to the problems in any practicum that you have not used to satisfy the practicum report requirement. The set of problems in a practicum counts as ONE homework problem.