STP598sta: Spatiotemporal Analysis

Homework 1

Name: Your name; NetID: Your ID

Due 11:59pm Sunday Sept 15 2024

Question 1

(a) Given latitude (θ) and longitude (λ) of two locations P_1 and P_2. Obtain the formula to calculate the geodesic distance D.

(b) Use the earth radius $R = 6371$ km. Can you compute the geodesic distances: (1) between Phoenix (33.28N, 112.03W) and Chicago (41.52N, 87.39W), and (2) between New York (40.43N, 73.56W) and Los Angeles (34.03N, 118.15W)?

Question 2

Consider the time series, $Y_t = X = \sin(\omega t + \theta)$ (so X is the amplitude, ω is the frequency and θ is the phase) where X is distributed with mean 0 and variance 1 independent of $\theta \sim \text{unif}(-\pi, \pi)$). Show that Y_t is weakly stationary.

Question 3

Show that a valid variogram $\gamma(\cdot)$ satisfies the negative definiteness condition, i.e. for any set of locations $\mathbf{s}_1, \cdots, \mathbf{s}_n$ and any set of $a_1, \cdots, a_n$ such that $\sum_{i=1}^{n} a_i = 0$, we have

$$\sum_i \sum_j a_i a_j \gamma(\mathbf{s}_i - \mathbf{s}_j) \leq 0$$

Question 4

Consider the `coalash` data frame in the `gstat` package in `R` and available from here. This data comes from the Pittsburgh coal seam on the Robena Mine Property in Greene County, PA (Cressie, 1993, p. 32). This data frame contains 208 coal ash core samples (the variable coal in the data frame) collected on a grid given by x and y planar coordinates (*not* latitude and longitude).

(a) Plot the sampled sites embedded on a map of the region. Add contour lines to the plot.

(b) Provide a descriptive summary (histograms, stems, quantiles, means, range, etc.) of the variable coal in the data frame.

(c) Plot variograms and correlograms of the response and comment on the need for spatial analysis here.

(d) If you think that there is need for spatial analysis, arrive at your best estimates of the range, nugget, and sill.