

Homework 8 - Stat 534
Due Friday March 27, 2017

1. The New York Leukemia data are attached along with homework as a comma delimited text file. You are to explore how excluding 3 potential outliers (observations 110, 120, and 121) affects the residual spatial autocorrelation, regression results, and the conclusions obtained from them. Is weighting still necessary? Fit *OLS*, *WLS*, and appropriate *GLS* models following my steps in the example in class.
2. Imagine a lattice process on a 2×3 rectangle. The sites \mathbf{s}_1 , \mathbf{s}_2 , and \mathbf{s}_3 make up the first row, the remaining sites make up the second row. Assume that the spatial connectivity matrix is given by

$$\mathbf{W} = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}$$

For a simultaneous and conditional autoregressive scheme with

$$\text{Var}[\mathbf{Z}(\mathbf{s})] = \sigma^2 (\mathbf{I} - \rho \mathbf{W})^{-1} (\mathbf{I} - \rho \mathbf{W}')^{-1}$$

and

$$\text{Var}[\mathbf{Z}(\mathbf{s})] = \sigma^2 (\mathbf{I} - \rho \mathbf{W})^{-1}$$

respectively, and with $\rho = 0.25$ do the following:

- (a) Identify the neighbors of lattice cell \mathbf{s}_2 .
 - (b) Compute the variance-covariance matrices for the SAR and CAR schemes.
 - (c) Determine which of the processes is second-order stationary. Justify your answer.
 - (d) Describe the correlation patterns that result. Are observations equicorrelated that are the same distance apart? Do correlations decrease with increasing lag distance?
3. We have 5 binomial count responses at 5 locations, i.e. the number of successes out of $n(\mathbf{s}_i)$ trials. Assume a single covariate X_1 and that there is no overdispersion. Find the diagonal elements of $\mathbf{V}_\mu^{1/2}$ being sure to express these in terms of β_0 and β_1 .