ASSIGNMENT 5

(1) When variables, like weight and blood pressure, are measured on a single individual over time, we expect the observations to be correlated over time. Consider the model

\[ Y_i = \beta_0 + \epsilon_i, \quad i = 1, \ldots, n \]

for your weight over \( n \) consecutive days. Clearly, we do not expect \( \text{Var}(\epsilon) \) to be \( I\sigma^2 \). Rather, we anticipate the weight measurements to be correlated over time and, furthermore, we expect measurements on two consecutive days to be more highly correlated than measurements that are further apart in time. One of the model that is used to model such behavior is a first-order autoregressive model:

\[ \epsilon_i = \rho \epsilon_{i-1} + \eta_i, \quad i = 1, \ldots, n, \]

where \( \eta_i \) are uncorrelated \((0, \sigma^2)\) random variables. Assume that \( \epsilon_1 \) has mean 0 and variance \( \sigma^2/(1-\rho^2) \), is independent of \( \eta_i \) and \( \rho < 1 \). Find the generalized least squares estimate of \( \beta_0 \) (with its variance) and compare its variance with the ordinary least squares estimate (when you ignore the correlation structure).

(2) It is desired to estimate the daily amount of evaporation from the soil as a function of air temperature, relative humidity, and wind. Since these factors vary considerably throughout the day it is not clear what function or aspect of these variables is most important. For this reason the following ten variables relating to these factors are recorded.” The carriers include maximum, minimum, and average soil temperatures, maximum, minimum, and average air temperature, maximum, minimum, and average relative humidity, and total wind (miles per day). The averages are actually the integrated area under the curve for the particular variable. The response \( Y \) is the daily amount of evaporation from the soil. These data were recorded from June 16 to July 21 for a location in west central Texas. The variables are \( x_1 = \text{‘max soil temp’}, \ x_2 = \text{‘min soil temp’}, \ x_3 = \text{‘average soil temp’}, \ x_4 = \text{‘max air temp’}, \ x_5 = \text{‘min air temp’}, \ x_6 = \text{‘ave air temp’}, \ x_7 = \text{‘max rel humidity’}, \ x_8 = \text{‘min rel humidity’}, \ x_9 = \text{‘average rel humidity’}, \ x_{10} = \text{‘total wind’}, \ Y = \text{‘daily evaporation’}. \) The data is in the file evap.sas. Use stepwise technique to find out the important variables.