STAT 153 - Introduction to Time Series
Homework Five

Fall 2022, UC Berkeley

Due by 11:59 pm on 18 November 2022
Total Points = 114

1. Consider the dataset in "norm.nao.monthly.b5001.current.ascii04Nov2022.txt' which gives (use the third column in the dataset) monthly data on the Northern Oscillation Index (this data has been taken from `https://www.cpc.ncep.noaa.gov/products/precip/CWlink/pna/nao.shtml`; see this page for more details about the data).

   a) I want to fit the $MA(q)$ model to this dataset. Look at the sample autocorrelation function of the data and figure out an appropriate value of q. (**2 points**)

   b) Fit the $MA(q)$ (with your selected choice of q in the previous part) to the data. Use the conditional sum of squares method described in class (do not use any inbuild function in R for this part). Report point estimates and standard errors for $\mu, \theta_1, \dots, \theta_q$ (**5 points**)

   c) Compare your answers to that given by the `arima` function in `R`. (**2 points**)

   d) Use your fitted model to obtain point predictions for the next 24 months. Comment on whether the predictions appear reasonable. (**2 points**)

2. Download the FRED dataset on "Long-Term Government Bond Yields: 10-year Main (including benchmark) for the United States" from `https://fred.stlouisfed.org/series/IRLTLT01USM156N`. This is a monthly dataset (units are in percent) and it is not seasonally adjusted.

   a) Fit an $AR(p)$ model to this datset with $p = 4$. Write the model as

   $$Y_t = \hat{\phi}_0 + \hat{\phi}_1 Y_{t-1} + \cdots + \hat{\phi}_p Y_{t-p} + Z_t \tag{1}$$

   and report parameter estimates and standard errors for $\hat{\phi}_j, j = 0, 1, \dots, p$. Use the model to obtain predictions for the next 100 months. Do the predictions look reasonable? (**5 points**)

   b) Let Y_t denote the original dataset. Construct a new dataset D_t by differencing Y_t: $D_t = Y_t - Y_{t-1}$. Plot the dataset D_t with time on the x-axis. Also plot the sample autocorrelation function of $\{D_t\}$. Would the $MA(1)$ model be reasonable for $\{D_t\}$? (**3 points**)

   c) Fit the $MA(1)$ model to $\{D_t\}$ and obtain point estimates and standard errors of the parameters (you can use the `R` function `arima`). Denote this model by (**2 points**)

   $$D_t = \hat{\mu} + \epsilon_t + \hat{\theta}\epsilon_{t-1} \qquad \text{where } \epsilon_t \overset{\text{i.i.d}}{\sim} N(0, \sigma^2). \tag{2}$$

d) Rewriting (2) as

$$\epsilon_t = \left(I + \hat{\theta}B\right)^{-1}(D_t - \hat{\mu}) = \left(I - \hat{\theta}B + \hat{\theta}^2B^2 - \hat{\theta}^3B^3 + \dots\right)(D_t - \hat{\mu}),$$

approximate the model (2) by an autoregressive model of the form:

$$D_t = \hat{\psi}_0 + \hat{\psi}_1 D_{t-1} + \hat{\psi}_2 D_{t-2} + \hat{\psi}_3 D_{t-3} + \epsilon_t \tag{3}$$

and report the values of $\hat{\psi}_0, \hat{\psi}_1, \hat{\psi}_2, \hat{\psi}_3$. (**2 points**)

e) Replace $D_s = Y_s - Y_{s-1}$ on both sides of the equation in (3) to obtain an AR model for Y_t. Compare the coefficients of this AR model with those of (1). Are they similar? (**3 points**)

f) Use the AR model from the previous part to obtain predictions for the next 100 months. Compare these predictions with those obtained from part (a). Comment on the differences between these two predictions. (**5 points**).

3. Download the FRED dataset on "Retail Sales: Beer, Wine, and Liquor Stores" from `https://fred.stlouisfed.org/series/MRTSSM4453USN`. This is a monthly dataset (the units are millions of dollars) and is not seasonally adjusted.

a) Fit an $AR(p)$ model to this dataset with $p = 16$. Write the model as

$$Y_t = \hat{\phi}_0 + \hat{\phi}_1 Y_{t-1} + \cdots + \hat{\phi}_p Y_{t-p} + Z_t \tag{4}$$

and report parameter estimates and standard errors for $\hat{\phi}_j, j = 0, 1, \dots, p$. Use the model to obtain predictions for the next 36 months (3 years). Do the predictions look reasonable? (**5 points**)

b) Would any Moving Average model work directly on this dataset? Answer this question by trying out $MA(q)$ for a range of values of q. You can evaluate models by looking at their future predictions. Use the `R` function `arima` to fit models and the function `predict` to obtain future predictions (**5 points**).

c) Let Y_t denote the original dataset. Construct a new dataset D_t via:

$$D_t = Y_t - Y_{t-1} - Y_{t-12} + Y_{t-13}$$

This can be created in `R` by, for example, the command `diff(diff(Yt, lag = 12))`. Plot the dataset D_t with time on the x-axis. Also plot the sample autocorrelation function of $\{D_t\}$. Would the $MA(1)$ model be reasonable for $\{D_t\}$? (**3 points**).

d) Fit the $MA(1)$ model to $\{D_t\}$ and obtain point estimates and standard errors of the parameters (you can use the `R` function `arima`). Denote this model by (**2 points**)

$$D_t = \hat{\mu} + \epsilon_t + \hat{\theta}\epsilon_{t-1} \qquad \text{where } \epsilon_t \overset{\text{i.i.d}}{\sim} N(0, \sigma^2). \tag{5}$$

e) Rewriting (5) as

$$\epsilon_t = \left(I + \hat{\theta}B\right)^{-1}(D_t - \hat{\mu}) = \left(I - \hat{\theta}B + \hat{\theta}^2B^2 - \hat{\theta}^3B^3 + \dots\right)(D_t - \hat{\mu}),$$

approximate the model (5) by an autoregressive model of the form:

$$D_t = \hat{\psi}_0 + \hat{\psi}_1 D_{t-1} + \hat{\psi}_2 D_{t-2} + \hat{\psi}_3 D_{t-3} + \epsilon_t \tag{6}$$

and report the values of $\hat{\psi}_0, \hat{\psi}_1, \hat{\psi}_2, \hat{\psi}_3$. (**2 points**)

f) Replace $D_s = Y_s - Y_{s-1} - Y_{s-12} + Y_{s-13}$ on both sides of the equation in (6) to obtain an AR model for Y_t. Compare the coefficients of this AR model with those of (4). Are they similar? (**3 points**)

g) Use the AR model from the previous part to obtain predictions for the next 36 months. Compare these predictions with those obtained from part (a). Comment on the differences between these two predictions. (**5 points**).

4. Consider the sunspots data that we looked at in class.

a) Plot the sample acf and pacf for this dataset. Based on these plots, argue that $AR(9)$ is an appropriate model for this dataset. (**4 points**)

b) Split this dataset by removing the last 40 datapoints and keeping them aside as a test dataset. The remaining observations will form the training dataset. Fit the $AR(p)$ model for $p = 1, 2, \dots, 15$ as well as the $MA(q)$ model for $q = 1, 2, \dots, 15$ to the training dataset. You can inbuilt R functions for fitting these models. Obtain predictions for each of these models for the future 40 datapoints and compare them to the actual observations in the test dataset. Which model performs best in terms of mean squared error of prediction? Compare the performance of the best model with the $AR(9)$ model (if they are different) obtained in the previous part. (**8 points**)

5. Let Y_1 and Y_2 be two uncorrelated random variables having mean zero and the same variance. Let $Y_3 := Y_1 + Y_2$.

a) What is the Best Linear Predictor (BLP) of Y_1 in terms of Y_3? (**3 points**)

b) What is the Best Linear Predictor (BLP) of Y_2 in terms of Y_3? (**3 points**)

c) What is the partial correlation $\rho_{Y_1,Y_2|Y_3}$? (**3 points**)

6. Let Y_1, Y_2 and ϵ be three uncorrelated random variables having mean zero and the same variance. Let $Y_3 := Y_1 + Y_2 + \epsilon$.

a) What is the Best Linear Predictor (BLP) of Y_1 in terms of Y_3? (**3 points**)

b) What is the Best Linear Predictor (BLP) of Y_2 in terms of Y_3? (**3 points**)

c) What is the partial correlation $\rho_{Y_1,Y_2|Y_3}$? (**3 points**)

7. Let Y be a 4×1 random vector with components Y_1, Y_2, Y_3 and Y_4. Suppose that each Y_i has mean zero. Suppose that the covariance matrix, Σ, of Y is given by

$$\Sigma = Cov(Y) = \begin{pmatrix} 1 & 0.5 & 0 & 1 \\ 0.5 & 1.25 & 2 & -1.5 \\ 0 & 2 & 5 & -5 \\ 1 & -1.5 & -5 & 7 \end{pmatrix}$$

The inverse of Σ is given by

$$\Sigma^{-1} = \begin{pmatrix} 3.25 & -2.5 & 0 & -1 \\ -2.5 & 5 & -2 & 0 \\ 0 & -2 & 2 & 1 \\ -1 & 0 & 1 & 1 \end{pmatrix}$$

a) List all pairs (i, j) with $1 \le i < j \le 4$ such that the correlation between Y_i and Y_j is strictly positive. Give reasons for your answer. (**2 points**).

b) List all pairs (i, j) with $1 \leq i < j \leq 4$ such that the partial correlation between Y_i and Y_j given all the other Y_k's equals zero. Give reasons for your answer. (**2 points**).

c) List all pairs (i, j) with $1 \leq i < j \leq 4$ such that the partial correlation between Y_i and Y_j given all the other Y_k's is strictly positive. Give reasons for your answer. (**3 points**).

d) Let $\beta_0^* + \beta_1^* Y_1 + \beta_2^* Y_2 + \beta_3^* Y_3$ be the Best Linear Predictor of Y_4 in terms of Y_1, Y_2 and Y_3. For what values of $i \in \{0, 1, 2, 3\}$ is the coefficient β_i^* exactly zero? For what values of $i \in \{0, 1, 2, 3\}$ is the coefficient β_i^* strictly positive? Give reasons for your answers. (**2 + 2 = 4 points**).

e) What is the variance of the residual $r_{Y_4|Y_1,Y_2,Y_3}$? (**2 points**).

8. Suppose X_1, Z_2, Z_3, Z_4 are uncorrelated random variables having mean zero. Also suppose that X_1 has variance 1 while each of Z_2, Z_3, Z_4 has variance $3/4$. Using these, we define new random variables X_2, X_3, X_4 via

$$X_2 = (-0.5)X_1 + Z_2, \quad X_3 = (-0.5)X_2 + Z_3 \quad \text{and} \quad X_4 = (-0.5)X_3 + Z_4.$$

a) What is the 3×3 covariance matrix of the 3×1 random vector with components X_1, X_2, X_3? (**3 points**)

b) What is the 3×3 covariance matrix of the 3×1 random vector with components X_2, X_3, X_4? (**3 points**)

c) What is the partial correlation between X_2 and X_4 given X_3? (**3 points**)

d) What is the partial correlation between X_1 and X_4 given X_2, X_3? (**3 points**)

e) What is the best linear predictor of X_4 in terms of X_1, X_2, X_3? (**3 points**)