STAT 153 - Introduction to Time Series Homework Four

Fall 2022, UC Berkeley

Due by 11:59 pm on 01 November 2022
Total Points = 63

1. Consider the dataset `lynx` that is available in base R. This gives the annual numbers of lynx trappings for 1821-1934 in Canada. Type `help(lynx)` to learn more about the dataset.

    a) Fit the $AR(2)$ model to the first 90 observations of this dataset. Report the estimates of ϕ_1, ϕ_2 and σ along with uncertainty quantification. (**3 points**)

    b) Write down an explicit formula for the predictions generated by your fitted $AR(2)$ model for Y_t for $t \geq 91$. (**4 points**)

    c) Use your $AR(2)$ to predict the data from time points $t = 91, \ldots, 114$. Also compute the standard deviations corresponding to the accuracy of prediction. (**4 points**).

    d) Compare your predictions with the actual values from the dataset. Comment on the accuracy of the predictions. (**2 points**)

2. Consider the US population dataset from `https://fred.stlouisfed.org/series/POPTHM` that we have worked with in class.

    a) Fit an $AR(2)$ model to this dataset. Report the estimates of ϕ_1, ϕ_2 and σ along with uncertainty quantification (**3 points**).

    b) Write down an explicit formula for the predictions generated by your fitted $AR(2)$ model for Y_t for the future months. (**4 points**)

    c) Use your $AR(2)$ model to predict the data for 36 months immediately succeeding the last month in the dataset. Plot these predictions and uncertainty indicators along with the original data. Do these predictions make intuitive sense? (**6 points**)

    d) Suppose that we want to predict the US population for the months preceding January 1959. For this purpose, fit the model:

    $$Y_t = \alpha_0 + \alpha_1 Y_{t+1} + \alpha_2 Y_{t+2} + Z_t$$

    Compare your fitted model with the forward model fitted earlier. Are there any similarities between the two models? (**4 points**)

    e) Using your model from the previous part, predict the US population for the 36 months immediately preceding January 1959. Plot these predictions and uncertainty indicators along with the original data. Do these predictions make intuitive sense? (**6 points**).

3. Consider again the US population dataset from `https://fred.stlouisfed.org/series/POPTHM` that we have worked with in class.

   a) Fit the $AR(3)$ model to this dataset and obtain point estimates of ϕ_1, ϕ_2, ϕ_3 and σ. (**3 points**).

   b) For better interpretability, I want to fit a model to the twice differenced series:

   $$D_t = (Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2}) = Y_t - 2Y_{t-1} + Y_{t-2}.$$

   Fit the $AR(1)$ model to D_t. Rewrite your fitted model:

   $$D_t = \hat{\alpha}_0 + \hat{\alpha}_1 D_{t-1} + Z_t.$$

   back in terms of Y_t by substituting $D_t = Y_t - 2Y_{t-1} + Y_{t-2}$ in the above equation. Compare this model to the $AR(3)$ model fitted in the previous part. Are they similar? (**4 points**)

   c) Compare the predictions of the two models for the next 60 time points. Are the predictions similar? (**4 points**).

4. 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. Separate the last 48 observations from this dataset and keep them as a test dataset. Fit the $AR(p)$ model for each $p = 1, 2, \ldots, 24$ for the training dataset and use it to predict the observations in the test dataset. Evaluate the 24 models based on the accuracy of prediction and report the model with the best prediction accuracy. (**10 points**).

5. We have seen in class that the $AR(1)$ model is said to belong to the causal stationary regime when $|\phi_1| < 1$. A similar characterization exists for the $AR(2)$ model.

   a) Show that the for the $AR(2)$ model, the roots of the characteristic equation are given by (**2 points**)

   $$\frac{\phi_1 - \sqrt{\phi_1^2 + 4\phi_2}}{2} \quad \text{and} \quad \frac{\phi_1 + \sqrt{\phi_1^2 + 4\phi_2}}{2}.$$

   b) Show that both the roots have modulus strictly greater than 1 if and only if the pair (ϕ_1, ϕ_2) satisfy all the following three inequalities: (**4 points**)

   $$\phi_2 + \phi_1 < 1 \qquad \phi_2 - \phi_1 < 1 \qquad |\phi_2| < 1.$$