Indian Institute of Technology Kanpur

Department of Mathematics and Statistics

Time Series Analysis (MTH442)

Assignment 3, Due date: October 16, 2024, Wednesday

Answers should be provided neatly. In case the handwriting is unreadable, the instructor and the teaching assistants hold the right to give a zero score. In case of cheating, all students involved will get zero, irrespective of who copied from whom. Write on a paper using some dark ink and scan using a good scanner to obtain a clearly readable PDF file. For the coding related question, use R markdown and generate the PDF output with echo = TRUE mode (so that the codes are also visible along with the outputs). Finally, join the two PDFs and submit the single PDF file only. The final file should be named as RollNo_Lastname_Firstname.pdf. If the file is not submitted in this nomenclature format, marks will be zero. No request should be made in that case.

1. In the context of equation $\boldsymbol{\Gamma}_T \boldsymbol{\phi}_T = \boldsymbol{\gamma}_T$, show that, if $\gamma(0) > 0$ and $\gamma(h) \to 0$ as $h \to \infty$, then $\boldsymbol{\Gamma}_T$ is positive definite. (1.5 points)

2. Generate 1000 realizations of length $T = 50, 200, 500$ (3 cases) each of an ARMA(1,1) process with $\phi = 0.9, \theta = 0.5,$ and $\sigma^2 = 1$. Find the MLEs of the three parameters in each case and compare the estimators to the true values in terms of mean square error, mean absolute deviation, and coverage of the 95% confidence intervals (3 measures). The final answers should be printed as a 3-by-3 matrix. (2 points)

3. Fit an AR(2) model to the `cmort` dataset from the `astsa` package. Assuming the fitted model is the true model, find the forecasts over a eight-week horizon X_{T+m}^T, for $m = 1, \ldots, 8$ using linear regression, and the corresponding 95% prediction intervals. Further, estimate the model parameters using Yule-Walker equations and compare them with the estimates based on linear regression. (1.5+1=2.5 points)

4. Suppose X_t is stationary with zero mean. Let $\varepsilon_t = X_t - \sum_{i=1}^{h-1} a_i X_{t-i}$ and $\delta_{t-h} = X_{t-h} - \sum_{j=1}^{h-1} b_j X_{t-j}$ be the two residuals where $a_1, \ldots, a_{h-1}$ and $b_1, \ldots, b_{h-1}$ are chosen so that they minimize the mean-squared errors $E[\varepsilon_t^2]$ and $E[\delta_{t-h}^2]$. The PACF at lag h was defined as the cross-correlation between $E[\varepsilon_t]$ and $E[\delta_{t-h}]$; that is,

$$\phi_{h,h} = \frac{E[\varepsilon_t \delta_{t-h}]}{\sqrt{E[\epsilon_t^2] E[\delta_{t-h}^2]}}.$$

   Let $\boldsymbol{R}_h$ be the $h \times h$ matrix with elements $\rho(i-j)$ for $i, j = 1, \ldots, h$, and let $\boldsymbol{\rho}_h = (\rho(1), \rho(2), \ldots, \rho(h))$ be the vector of lagged autocorrelations, $\rho(h) = \mathrm{corr}(X_{t+h}, X_t)$. Let $\tilde{\boldsymbol{\rho}}_h = (\rho(h), \rho(h-1), \ldots, \rho(1))$ be the reversed vector. In addition, let X_t^h denote the BLP of X_t given $\{X_{t-1}, \ldots, X_{t-h}\}$, i.e., $X_t^h = \alpha_{h,1} X_{t-1} + \ldots + \alpha_{h,h} X_{t-h}$. Show that

$$\phi_{h,h} = \frac{\rho(h) - \tilde{\boldsymbol{\rho}}_{h-1}' \boldsymbol{R}_{h-1}^{-1} \boldsymbol{\rho}_h}{1 - \tilde{\boldsymbol{\rho}}_{h-1}' \boldsymbol{R}_{h-1}^{-1} \tilde{\boldsymbol{\rho}}_{h-1}} = \alpha_{h,h}. \quad \text{(2.5 points)}$$

5. Suppose we wish to find a prediction function $g(\cdot)$ that minimizes $MSE = E[(Y - g(X))^2]$, where X and Y are jointly distributed random variables with density function $f(x, y)$. Suppose we restrict our choices for the function $g(\cdot)$ to linear functions of the form $g(x) = a + bx$ and determine a and b to minimize MSE. Show that $a = 1$, $b = 0$, and $MSE = 3$. What do you interpret this to mean? (1.5 points)