MIT Sloan School of Management

Analytics of Finance 15.457
Hui Chen Spring 2021

Problem Set 5

(Due: 7:30 AM, Tuesday, April 27)

1. **Interview question**: When we estimate an AR(1) model using OLS, is the OLS estimator still unbiased? Is it consistent? Please explain.

2. Suppose that the daily log return of a security follows the model

$$r_t = 0.01 + 0.1 r_{t-2} + \varepsilon_t$$

where ε_t is IID normal with mean zero and variance 0.02.

   (a) What are the mean and variance of the return series r_t?

   (b) Compute the lag-1 and lag-2 autocorrelations of r_t.

   (c) Assume that $r_{100} = 0.01$, and $r_{99} = 0.02$. Compute the 1- and 2-step-ahead forecasts of the return series at the forecast origin $t = 100$. What are the associated standard deviations of the forecast errors?

3. **Modeling inflation.**

   (a) Download the file "CPI.csv", which contains the data of seasonally adjusted monthly Consumer Price Index for all urban consumers from the U.S. Bureau of Labor Statistics. Compute the monthly inflation rate as $\pi_t = \ln(P_t/P_{t-1})$, where P_t is the price index level in quarter t. Plot the series of π_t. Also plot the autocorrelation function ρ_k from 1 to 12 lags. You can either use the definition below or the function *acf* in R to produce the autocorrelation function.

$$\rho_k = \frac{Cov(x_t, x_{t-k})}{\sqrt{Var(x_t)}\sqrt{Var(x_{t-k})}}$$

   (b) Use AIC and BIC to select the optimal AR(p) model for monthly inflation. Code your own version of model selection, and compare the results with those from an auto-selection function such as auto.arima in R.

4. **Forecasting corporate earnings:** It is the beginning of 2020. You are an equity analyst in charge of building a model to forecast the quarterly earnings of Wal-Mart Stores, Inc. Download the data file "WMT.csv", which contains quarterly earnings (based on net income) for Wal-Mart.

(a) Denote the log quarterly earnings by x_t. Plot the time series of the first difference $\Delta x_t = x_t - x_{t-1}$, and the seasonal difference $\Delta_4 x_t = x_t - x_{t-4}$. Explain their economic meanings and comment on their statistical properties.

(b) As a first attempt, build an ARIMA(0,1,1) for x_t,

$$x_{t+1} - x_t = a_0 + \varepsilon_{t+1} - \theta\varepsilon_t.$$

Use data from 1971Q1 to 2015Q4 to estimate this model. (In R, you can use the function *arima* to do the estimation.)

(c) Next, build an "airline model" for x_t,

$$(x_t - x_{t-1}) - (x_{t-4} - x_{t-5}) = (\varepsilon_t - \theta_1\varepsilon_{t-1}) - \theta_4(\varepsilon_{t-4} - \theta_1\varepsilon_{t-5}).$$

Again, use data from 1971Q1 to 2015Q4 to estimate this model. Explain the economic meanings of θ_1 and θ_4.

(d) Use the period from 2016Q1 to 2019Q4 as the testing period. Generate your one-quarter ahead forecast[1] for Wal-Mart's earnings in fiscal quarter Q1 of 2016 through Q4 of 2019 based on the two estimated models. (Hint: You can use the function *predict* in R to produce the forecasts. Also, notice the constant term a_0 in the ARIMA(0,1,1), which makes its forecast different from a simple exponential moving average.)

(e) Plot the forecasting errors of the two models from 2016Q1 to 2019Q4. Summarize the forecasting errors of the two models using the mean-squared error (MSE), defined as

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(x_{t+i} - E_{t+i-1}[x_{t+i}])^2,$$

where x_{t+i} is the realized earnings in quarter $t+i$, and $E_{t+i-1}[x_{t+i}]$ is your prediction for that quarter. Which model performs better?

(f) Finally, extend the one-quarter-ahead forecasts to include the fiscal year 2020 without updating the model parameters. Comment on the reasons behind the (deteriorating) performances of the two models during the COVID-19 pandemic. What would you have done differently in real time to improve the forecast accuracy?

[1] One-quarter-ahead forecast means predicting x_{t+1} while taking all the observations up to x_t as given.