Lecture 5

Recall that, we concluded the last lecture with the following:

If our data $\{Y_1, Y_2, ..., Y_T\}$ is from a **stationary** $ARMA\ (p, q)$ given by

$$Y_t = \alpha + \sum_{i=1}^{p} \beta_i Y_{t-i} + \sum_{j=1}^{q} \theta_j \epsilon_{t-j} + \epsilon_t,$$

then the one-step-ahead forecast based on $\mathcal{F}_T$ is given by

$$\hat{Y_T}(1) = \hat{\alpha} + \sum_{i=1}^{p} \hat{\beta}_i Y_{(T+1)-i} + \sum_{j=1}^{q} \hat{\theta}_j e_{(T+1)-j}, \qquad (\mathbf{1})$$

where $\hat{\alpha}$, the $\hat{\beta}_i$'s and the $\hat{\theta}_j$'s are the MLEs of the model parameters and e_t's are the residuals which represent the unobservable errors for $t = 1, 2, ..., T$ (all of these explained in the last lecture notes).

As we are currently at time T, the forecast error associated one-step-ahead forecast of Y_{T+1} is the random variable given by

$$\begin{aligned} \text{One-step-ahead Forecast Error} &= Y_{T+1} - \hat{Y_T}(1) \\ &= \{\hat{\alpha} + \sum_{i=1}^{p} \hat{\beta}_i Y_{(T+1)-i} + \sum_{j=1}^{q} \hat{\theta}_j e_{(T+1)-j} + \epsilon_{T+1}\} - \hat{Y_T}(1) \\ &= \epsilon_{T+1} \qquad (\mathbf{2}) \end{aligned}$$

where $\hat{Y_T}(1)$ is given in **(1)** above. Hence, from **(2)** above, the corresponding one-step-ahead forecast error variance is

$$V[Y_{T+1} - \hat{Y_T}(1)] = V[\epsilon_{T+1}] = \sigma^2.$$

Now, the two-step-ahead forecast of Y_{T+2} given $\mathcal{F}_T$, the information up to the current time T, is given by

$$\begin{aligned}
\hat{Y_T}(2) &= E[Y_{T+2} \mid \mathcal{F}_T] \\
&= E[(\hat{\alpha} + \hat{\beta}_1 Y_{T+1} + \sum_{i=2}^{p} \hat{\beta}_i Y_{(T+2)-i} + \sum_{j=2}^{q} \hat{\theta}_j \epsilon_{(T+2)-j} + \hat{\theta}_1 \epsilon_{T+1} + \epsilon_{T+2}) \mid \mathcal{F}_T] \\
&= \hat{\alpha} + \hat{\beta}_1 \hat{Y_T}(1) + \sum_{i=2}^{p} \hat{\beta}_i Y_{(T+2)-i} + \sum_{j=2}^{q} \hat{\theta}_j e_{(T+2)-j} + E[\hat{\theta}_1 \epsilon_{T+1} + \epsilon_{T+2}) \mid \mathcal{F}_T]
\end{aligned}$$

Noting that the conditional expectation on the right side of the last equality above is zero because both ϵ_{T+1} and ϵ_{T+2} are independent of $\mathcal{F}_T$, we get the two-step-ahead forecast of Y_{T+2} as

$$\hat{Y_T}(2) = \hat{\alpha} + \hat{\beta}_1 \hat{Y_T}(1) + \sum_{i=2}^{p} \hat{\beta}_i Y_{(T+2)-i} + \sum_{j=2}^{q} \hat{\theta}_j e_{(T+2)-j}. \qquad \textbf{(3)}$$

As before, the associated forecast error is the random variable given by

Two-step-ahead Forecast Error

$$\begin{aligned}
&= Y_{T+2} - \hat{Y_T}(2) \\
&= \{\hat{\alpha} + \sum_{i=1}^{p} \hat{\beta}_i Y_{(T+2)-i} + \sum_{j=1}^{q} \hat{\theta}_j \epsilon_{(T+2)-j} + \epsilon_{T+2}\} - \hat{Y_T}(2) \\
&= \{\hat{\alpha} + \hat{\beta}_1 Y_{T+1} + \sum_{i=2}^{p} \hat{\beta}_i Y_{(T+2)-i} + \sum_{j=2}^{q} \hat{\theta}_j e_{(T+2)-j} + (\hat{\theta}_1 \epsilon_{T+1} + \epsilon_{T+2})\} - \hat{Y_T}(2) \\
&= \hat{\beta}_1 [Y_{T+1} - \hat{Y_T}(1)] + \hat{\theta}_1 \epsilon_{T+1} + \epsilon_{T+2}. \qquad \textbf{(4)}
\end{aligned}$$

where $\hat{Y_T}(2)$ is given in **(3)**. Hence, from **(4)** above, the corresponding two-step-ahead forecast error variance is

$$\begin{aligned}
&V[Y_{T+2} - \hat{Y_T}(2)] \\
&= V[\hat{\beta}_1 [Y_{T+1} - \hat{Y_T}(1)]] + V[\hat{\theta}_1 \epsilon_{T+1} + \epsilon_{T+2}] + 2Cov[\hat{\beta}_1 [Y_{T+1} - \hat{Y_T}(1)], [\hat{\theta}_1 \epsilon_{T+1} + \epsilon_{T+2}]] \\
&= \hat{\beta_1}^2 \sigma^2 + (\hat{\theta_1}^2 + 1)\sigma^2 + 2Cov[\hat{\beta}_1 Y_{T+1}, \hat{\theta}_1 \epsilon_{T+1}] \\
&= \hat{\beta_1}^2 \sigma^2 + (\hat{\theta_1}^2 + 1)\sigma^2 + 2\hat{\beta}_1 \hat{\theta}_1 \sigma^2 \\
&= [1 + (\hat{\beta}_1 + \hat{\theta}_1)^2]\sigma^2 \\
&\geq \sigma^2.
\end{aligned}$$

Note that

(i) on the right side of the first equality above, when you expand the last covariance term, the only possible non-zero covariance exists between Y_{T+1} and ϵ_{T+1} and the rest are zero;

(ii) the error variance for the two-step-ahead forecast is bigger than or equal to that of the one-step-ahead forecast - which is natural, as the two-step-ahead forecast is itself *based* on the one-step-ahead forecast too.

The k-step-ahead forecast $\hat{Y_T}(k)$ is computed in tha same manner as described above for $k = 3, 4, ...$

Forecasting for general $ARIMA\ (p, d, q)$ **Models**

Now, let us consider the forecasting problem for a **non stationary** time series $\{Y_t\}$. So, assume that $\{Y_t\}$ is an $ARIMA\ (p, d, q)$ model where $d \geq 1$. We will consider the case when $d = 1$, and describe the procedure of forecasting for the original non stationary $\{Y_t\}$, and the procedure can be easily generalised for $d \geq 2$.

Let $\{Y_1, Y_2, ..., Y_T\}$ be the data from $\{Y_t\}$ which is an $ARIMA\ (p, d, q)$ time series. Suppose further that the first differenced series $\{Y_t^{(1)}\}$ of $\{Y_t\}$ be a stationary time series, where

$$Y_t^{(1)} = Y_t - Y_{t-1}, \quad t = 2, 3, 4, ...$$

Then, the new data set from the stationary $\{Y_t^{(1)}\}$ is given by $\{Y_2^{(1)}, Y_3^{(1)}, ..., Y_T^{(1)}\}$ - a set of $T - 1$ observations. Now,

(i) Fit the appropriate $ARMA$ model to this data set from the fist differenced series.

(ii) Proceed to to find the forecasts $\hat{Y_T}^{(1)}(1)$, $\hat{Y_T}^{(1)}(2)$, and so on, using the forecasting method described above stationary $ARMA$ models.

We know that the one-step-ahead forecast $\hat{Y_T}^{(1)}(1)$, computed for the stationary $ARMA$ model of $\{Y_t^{(1)}\}$, is given by

$$\begin{aligned}
\hat{Y_T}^{(1)}(1) &= E[Y_{T+1}^{(1)} \mid \mathcal{F}_T] \\
&= E[Y_{T+1} - Y_T \mid \mathcal{F}_T] \quad \text{(by the definition of } Y_{T+1}^{(1)}) \\
&= E[Y_{T+1} \mid \mathcal{F}_T] - Y_T \quad \text{(since } Y_T \in \mathcal{F}_T) \\
&= \hat{Y_T}(1) - Y_T.
\end{aligned}$$

Hence, we see that

$$\hat{Y_T}(1) = Y_T + \hat{Y_T}^{(1)}(1),$$

which is the one-step-ahead forecast for the original non stationary data $\{Y_t\}$.

Now, the two-step-ahead forecast $\hat{Y_T}^{(1)}(2)$ is given by

$$\begin{aligned}
\hat{Y_T}^{(1)}(2) &= E[Y_{T+2}^{(1)} \mid \mathcal{F}_T] \\
&= E[Y_{T+2} - Y_{T+1} \mid \mathcal{F}_T] \quad \text{(by the definition of } Y_{T+2}^{(1)}) \\
&= E[Y_{T+2} \mid \mathcal{F}_T] - E[Y_{T+1} \mid \mathcal{F}_T] \\
&= \hat{Y_T}(2) - \hat{Y_T}(1).
\end{aligned}$$

Hence, we see that

$$\hat{Y_T}(2) = \hat{Y_T}(1) + \hat{Y_T}^{(1)}(2),$$

which is the two-step-ahead forecast for the original non stationary $\{Y_t\}$.

Proceeding the same way, we compute the other forecasts for the non stationary $\{Y_t\}$ from those of the stationary $\{Y_t^{(1)}\}$, the first differenced series of $\{Y_t\}$.

Now, suppose that $\{Y_t\}$ is non stationary such that its second differenced series $\{Y_t^{(2)}\}$ is stationary, where

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

Here, it is to be noted that the original series $\{Y_t\}$, and its first differenced series $\{Y_t^{(1)}\}$ are **both** non stationary.

So, with the help of the forecasts for $\{Y_t^{(2)}\}$, we get the forecasts for $\{Y_t^{(1)}\}$ in the same way as described above. Again, now using the forecasts for $\{Y_t^{(1)}\}$, we get the forecasts for the original non stationary $\{Y_t\}$.

Note: Suppose that the data is the stock price, as has been empirically proved, then this time series is *not* stationary. Also, as well documented in Finance, the log return series of the stock prices is stationary. So, if we choose to work on the forecasting of stock price, letting $\{Y_t\}$ denote the stock price series and confirming that it is not stationary, we construct the following transformed series $\{r_t\}$ where

$$r_t = \log\left(\frac{Y_t}{Y_{t-1}}\right) = \log(Y_t) - \log(Y_{t-1}), \quad t = 2, 3, ...$$

Here, r_t is known as the log return of the stock at time t.

After confirming the stationarity of $\{r_t\}$, (a) get forecasts for this log return, and (b) transform this log return forecast back to the forecast of the stock price.

Exercise 1 - (Answers due by 18:00 hours on November 04, 2021 (Thursday))

Reference: My e-mail on the links to data sets

With reference to my above e-mail, download a data set of your choice (with a minimum of 100 observations and a maximum of your choice, say in thousands). Perform the following tasks on this data (steps from **Lecture 4**) using either R or $Python$:

1. Prepare the data: (a) Chronological ordering - the first column containing the chronology, and the second containing the data; (b) Any clean-up procedure applied, like capping the outliers (extreme values), replacing the missing data if any. Then, divide this data set into (a) data for model building, and (b) data for testing. In practice, it is usually 95% to 98% of the total data for model building and the rest for testing. If it is financial data, keep almost the entire data for model building and the last 5 to 10 observations for testing, as long term forecasting in Finance performs very badly in reality.

2. Plot of the total data obtained from the above step.

3. Descriptive Statistics of this data (mean, median, minimum, maximum, standard deviation, skewness, kurtosis and the percentage of observation in (-3*SD, +3*SD) and any other which you feel as important)

4. Check for the stationarity of the prepared data using the three tests - ADF, PP and $KPSS$. If the data is not stationary, come up with the required transformation(s) to get stationarity.

5. Fit the appropriate $ARMA$ model to the stationary data (given in Step 4 above) - (a) using AIC; (b) using BIC.

6. Get the forecasts corresponding to the size of the testing data; that is, forecast the testing data from the model data. Compute the mean square error (MSE) of the forecast given by

$$MSE = \frac{\sum(\text{original test data value - forecast value})^2}{\text{number of forecasts}}.$$

Also, get the graph of the test data and the forecast values - for visual comparison.