Lecture Note 6

- Introduction to Forecasting
- Forecasting of Autoregressive models
- Forecasting of Moving average models

Introduction to Forecasting

Let $y^f_{t+\tau|t}$, $\tau = 1, 2, \cdots$ be a forecast of $y_{t+\tau}$ at time t and $e_{t+\tau|t} = y_{t+\tau} - y^f_{t+\tau|t}$ be the corresponding forecast error. In order to determine our optimal forecast of $y_{t+\tau}$, $y^f_{t+\tau|t}$, we first need to specify our loss function. During this course, we will consider quadratic loss function, i.e.

$$L\left(y_{t+\tau}, y^f_{t+\tau|t}\right) = Ae^2_{t+\tau|t} = A\left(y_{t+\tau} - y^f_{t+\tau|t}\right)^2.$$

The goal is to minimize the expected value of our loss conditional on information that we have at time t, i.e.

$$y^{f*}_{t+\tau|t} = \arg\min_{y^f_{t+\tau|t}} \mathbb{E}\left[L\left(y_{t+\tau}, y^f_{t+\tau|t}\right)|\mathcal{I}_t\right] = \arg\min_{y^f_{t+\tau|t}} \mathbb{E}\left[A\left(y_{t+\tau} - y^f_{t+\tau|t}\right)^2|\mathcal{I}_t\right],$$

where $\mathcal{I}_t$ is the set of all information available at time t. The first order condition is

$$2A\mathbb{E}\left[\left(y_{t+\tau} - y^{f*}_{t+\tau|t}\right)|\mathcal{I}_t\right] = 0 \Rightarrow \mathbb{E}\left[\left(y_{t+\tau} - y^{f*}_{t+\tau|t}\right)|\mathcal{I}_t\right] = 0.$$

Hence

$$y^{f*}_{t+\tau|t} = \mathbb{E}\left(y_{t+\tau}|\mathcal{I}_t\right), \tag{1}$$

and its corresponding forecasting error is

$$e^*_{t+\tau|t} = y_{t+\tau} - y^{f*}_{t+\tau|t} = y_{t+\tau} - \mathbb{E}\left(y_{t+\tau}|\mathcal{I}_t\right).$$

• Note that the optimal forecast based on the quadratic loss function, $y^{f*}_{t+\tau|t}$, depends only on expected value of $y_{t+\tau}$ conditional on $\mathcal{I}_t$ and the conditional variance of $y_{t+\tau}$ does not play any role here. We can interpret this as if an agent is risk neutral.

Forecasting of Autoregressive models

Recall that if y_t follows an Autoregressive of order p, AR(p), we can write

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t, \tag{2}$$

where $\varepsilon_t \sim \text{iid}(0, \sigma^2)$.

τ-step ahead forecast and forecast error

By using the optimal forecast based on the quadratic loss function given by (1), and subtituting for $y_{t+\tau}$ from (2), we have

$$y^{f*}_{t+\tau|t} = \phi_0 + \phi_1 \mathbb{E}\left(y_{t+\tau-1}|\mathcal{I}_t\right) + \phi_2 \mathbb{E}\left(y_{t+\tau-2}|\mathcal{I}_t\right) + \cdots + \phi_p \mathbb{E}\left(y_{t+\tau-p}|\mathcal{I}_t\right) + \mathbb{E}\left(\varepsilon_{t+\tau}|\mathcal{I}_t\right).$$

Note that the information set at time t includes only current and past values of y_t, i.e. $\mathcal{I}_t = \{y_t, y_{t-1}, y_{t-2}, \cdots\}$ and $\varepsilon_{t+\tau}$ is independent of current and past information at time t. Hence $\mathbb{E}\left(\varepsilon_{t+\tau}|\mathcal{I}_t\right) = 0$. Moreover, $\mathbb{E}\left(y_{t+\tau-i}|\mathcal{I}_t\right)$ for $i = 1, 2, \cdots, p$ are the optimal forecasts of $y_{t+\tau-i}$, i.e. $y^{f*}_{t+\tau-i|t}$. Therefore,

$$y^{f*}_{t+\tau|t} = \phi_0 + \phi_1 y^{f*}_{t+\tau-1|t} + \phi_2 y^{f*}_{t+\tau-2|t} + \cdots + \phi_p y^{f*}_{t+\tau-p|t}. \tag{3}$$

We can also show that

$$e^{*}_{t+\tau|t} = \phi_1 e^{*}_{t+\tau-1|t} + \phi_2 e^{*}_{t+\tau-2|t} + \cdots + \phi_p e^{*}_{t+\tau-p|t} + \varepsilon_{t+\tau}. \tag{4}$$

From (3), we can see that the optimal forecast of y at time $t + \tau$ depends on the optimal forecasts of y over its past p time periods, $y^{f*}_{t+\tau-i|t}$ for $i = 1, 2, \cdots, p$. As a result of it, we can see in (4) that the optimal forecast error at time $t + \tau$ also depends on the optimal forecast errors over its past p time periods in addition to $\varepsilon_{t+\tau}$.

Q: Given (3), find one step ahead optimal forecast of y, $y^{f*}_{t+1|t}$ and its corresponding forecast error.

A: We have

$$y^{f*}_{t+1|t} = \phi_0 + \phi_1\ y^{f*}_{t|t} + \phi_2 y^{f*}_{t-1|t} + \cdots + \phi_p\ y^{f*}_{t-p+1|t}.$$

But $y^{f*}_{t-i|t} = \mathbb{E}\left(y_{t-i}|\mathcal{I}_t\right) = y_{t-i}$ for $i = 0, 1, 2, \cdots$. Therefore,

$$y^{f*}_{t+1|t} = \phi_0 + \phi_1 y_t + \phi_2 y_{t-1} + \cdots + \phi_p\ y_{t-p+1}.$$

Regarding the forecast error, we have

$$e^*_{t+1|t} = y_{t+1} - y^{f*}_{t+1|t} = \varepsilon_{t+1}.$$

Q: Given (3), find two and three steps ahead optimal forecasts of y, and their corresponding forecast error.

A: Using similar lines of arguments as in previous answer, we have

$$y^{f*}_{t+2|t} = \phi_0 + \phi_1 y^{f*}_{t+1|t} + \phi_2 y_t + \cdots + \phi_p \ y_{t-p+2},$$

and

$$e^*_{t+2|t} = y_{t+2} - y^{f*}_{t+2|t} = \varepsilon_{t+2} + \phi_1 \left(y^{f*}_{t+1|t} - y_{t+1} \right) = \varepsilon_{t+2} + \phi_1 \varepsilon_{t+1}.$$

We also have

$$y^{f*}_{t+3|t} = \phi_0 + \phi_1 y^{f*}_{t+2|t} + \phi_2 y^{f*}_{t+1|t} + \phi_3 y_t + \cdots + \phi_p \ y_{t-p+3},$$

and

$$\begin{aligned} e^*_{t+3|t} =& y_{t+3} - y^{f*}_{t+3|t} = \varepsilon_{t+3} + \phi_1 \left(y^{f*}_{t+2|t} - y_{t+2} \right) + \phi_2 \left(y^{f*}_{t+1|t} - y_{t+1} \right) \\ =& \varepsilon_{t+3} + \phi_1 \left(\varepsilon_{t+2} + \phi_1 \varepsilon_{t+1} \right) + \phi_2 \varepsilon_{t+1} = \varepsilon_{t+3} + \phi_1 \varepsilon_{t+2} + \left(\phi_1^2 + \phi_2 \right) \varepsilon_{t+1} \end{aligned}$$

Q: Find variance of one, two and three steps ahead optimal forecast errors of y.

A: We have

$$\text{var}(e^*_{t+1|t}) = \text{var}(\varepsilon_{t+1}) = \sigma^2,$$

$$\text{var}(e^*_{t+2|t}) = \text{var}(\varepsilon_{t+2} + \phi_1 \varepsilon_{t+1}) = \left(1 + \phi_1^2\right) \sigma^2,$$

and

$$\text{var}\left(e^*_{t+3|t}\right) = \text{var}\left[\varepsilon_{t+3} + \phi_1 \varepsilon_{t+2} + \left(\phi_1^2 + \phi_2\right) \varepsilon_{t+1}\right] = \left[1 + \phi_1^2 + \left(\phi_1^2 + \phi_2\right)^2\right] \sigma^2,$$

As we can see in the above answer, the variance of optimal forecast error increases as we increase our foresting time horizon. In what follows we are going to formalize the finding. In order to do so, suppose that the AR(p) model is weakly stationary. In this case, we can present the model as a linear times model, that is

$$y_t = \mu + \sum_{i=0}^{\infty} \delta_i \varepsilon_{t-i}.$$

So we can write, $y_{t+\tau} = \mu + \sum_{i=0}^{\infty} \delta_i \varepsilon_{t+\tau-i}$ and

$$y_{t+\tau|t}^{f*} = \mathbb{E}(y_{t+\tau}|\mathcal{I}_t) = \mu + \sum_{i=\tau}^{\infty} \delta_i \varepsilon_{t+\tau-i}.$$

Therefore,

$$e_{t+\tau|t}^{*} = y_{t+\tau} - y_{t+\tau|t}^{f*} = \sum_{i=0}^{\tau-1} \delta_i \varepsilon_{t+\tau-i},$$

and hence

$$\text{var}\left(e_{t+\tau|t}^{*}\right) = \sigma^2 \sum_{i=0}^{\tau-1} \delta_i^2. \tag{5}$$

It is clear from (5) that as τ increases the variance of optimal forecast error increases as well. Note that

$$\lim_{\tau \to \infty} \text{var}\left(e_{t+\tau|t}^{*}\right) = \sigma^2 \sum_{i=0}^{\infty} \delta_i^2 = \text{var}(y_t).$$

It can be also shown that

$$\lim_{\tau \to \infty} y_{t+\tau|t}^{f*} = \mu = \mathbb{E}(y_t).$$

Forecasting of Moving average models

Recall that if y_t follows a Moving average of order q, MA(q), we can write

$$y_t = c_0 + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q}, \tag{6}$$

where $\varepsilon_t \sim \text{iid}(0, \sigma^2)$. So the $\tau-$step ahead optimal forecast of y_t, $y_{t+\tau|t}^{f*}$ is

$$y_{t+\tau|t}^{f*} = c_0 + \mathbb{E}\left(\varepsilon_{t+\tau}|\mathcal{I}_t\right) - \theta_1 \mathbb{E}\left(\varepsilon_{t+\tau-1}|\mathcal{I}_t\right) - \theta_2 \mathbb{E}\left(\varepsilon_{t+\tau-2}|\mathcal{I}_t\right) - \cdots - \theta_q \mathbb{E}\left(\varepsilon_{t+\tau-q}|\mathcal{I}_t\right).$$

Note that, $\mathbb{E}\left(\varepsilon_{t+i}|\mathcal{I}_t\right) = 0$ for $i = 1, 2, \cdots$. So, for $\tau = 1, 2, \cdots, q$, we have

$$y^{f*}_{t+\tau|t} = c_0 - \theta_\tau \varepsilon_t - \theta_{\tau+1}\varepsilon_{t-1} - \cdots - \theta_q \varepsilon_{t+\tau-q}.$$

and hence

$$e^{*}_{t+\tau|t} = \varepsilon_{t+\tau} - \theta_1\varepsilon_{t+\tau-1} - \theta_2\varepsilon_{t+\tau-2} - \cdots - \theta_{\tau-1}\varepsilon_{t+1}.$$

Therefore,

$$\text{var}\left(e^{*}_{t+\tau|t}\right) = \left(1 + \theta_1^2 + \theta_2^2 + \cdots + \theta_{\tau-1}^2\right)\sigma^2.$$

Also, for any $\tau = q+1, q+2, \cdots$, we have

$$y^{f*}_{t+\tau|t} = c_0 = \mathbb{E}(y_t),$$

and

$$e^{*}_{t+\tau|t} = \varepsilon_{t+\tau} - \theta_1\varepsilon_{t+\tau-1} - \theta_2\varepsilon_{t+\tau-2} - \cdots - \theta_q\varepsilon_{t+1}.$$

Therefore,

$$\text{var}\left(e^{*}_{t+\tau|t}\right) = \left(1 + \theta_1^2 + \theta_2^2 + \cdots + \theta_q^2\right)\sigma^2 = \text{var}\left(y_t\right).$$