Time Series Analysis

Prof. Lea Petrella

MEMOTEF Department
Sapienza, University of Rome, Italy

Class 11

Forecasting with ARMA

- After the model is chosen and diagnostic is done, the $ARMA(p, q)$ can be used to forecast future values of the time series.
- In probabilistic terms, the forecast is considered on the basis of the collected information on the process at time t, denoted $\mathcal{F}_t$, where

$$\mathcal{F}_t = \{X_t, X_{t-1}, X_{t-2}, \cdots, X_0, X_{-1}, \cdots\}.$$

- Goal: To forecast the value of the process at time $t+k$, X_{t+k}, conditioning on $\mathcal{F}_t$.
- Define $\hat{X}_{t+k}$ the predictor of X_{t+k} and $\hat{x}_{t+k}$ its realization.
- Choose $\hat{X}_{t+k}$ such that it minimizes the forecast men squared error

$$\mathbb{E}(e^2_{t+k}) = \mathbb{E}(X_{t+k} - \hat{X}_{t+k})^2,$$

where $e_{t+k} = X_{t+k} - \hat{X}_{t+k}$ is the forecast error and is a measure of the model predictive ability.

- It can be shown that the best predictor $\hat{X}_{t+k}$ is given by its conditional expected value:

$$\hat{X}_{t+k} = \mathbb{E}(X_{t+k}|\mathcal{F}_t) = \mathbb{E}_t(X_{t+k}).$$

- Consider that for conditional expected values we have

$$\begin{aligned} \mathbb{E}(X_{t+j}|\mathcal{F}_t) &= x_{t+j} \quad \textit{if } j \leq 0 \\ \mathbb{E}(X_{t+j}|\mathcal{F}_t) &= \hat{x}_{t+j} \quad \textit{if } j > 0 \end{aligned}$$

- Setting $e_t = X_t - \hat{X}_t$,

$$\begin{aligned} \mathbb{E}(\epsilon_{t+j}|\mathcal{F}_t) &= e_{t+j} \quad \textit{if } j \leq 0 \\ \mathbb{E}(\epsilon_{t+j}|\mathcal{F}_t) &= 0 \qquad \textit{if } j > 0 \end{aligned}$$

- Recall that e_{t+k} is a random variable thus, it is worth to evaluate its probabilistic structure.

Parameter estimation for ARMA processes

- For the *ARMA* process it can be shown that as the forecasting horizon increases ($k \to \infty$):
- the forecast tends to the unconditional mean of the process

$$\lim_{k \to \infty} \hat{X}_{t+j} = \mathbb{E}(X_{t+j}),$$

- the variance of the error tends to the unconditional variance of the process

$$\lim_{k \to \infty} Var(e_{t+j}) = Var(X_t).$$

Least squares

- Consider a stationary $AR(1)$ process

$$X_t = \delta + \varphi X_{t-1} + \epsilon_t.$$

- If the parameters were known, the one step ahead $(k = 1)$ forecast would be

$$\hat{x}_{t+1} = \mathbb{E}_t(X_{t+1}) = \mathbb{E}_t(\delta + \varphi X_t + \epsilon_{t+1}) = \delta + \varphi x_t.$$

- The forecasting error

$$e_{t+1} = X_{t+1} - \hat{X}_{t+1} = \delta + \varphi x_t + \epsilon_{t+1} - (\delta + \varphi x_t) = \epsilon_{t+1}$$

$$\Rightarrow Var(e_{t+1}) = Var(\epsilon_{t+1}) = \sigma^2.$$

- In order to forecast $k = 2$ steps ahead:

$$\hat{x}_{t+2} = \mathbb{E}_t(X_{t+2}) = \mathbb{E}_t(\delta + \varphi X_{t+1} + \epsilon_{t+2}) = \delta + \mathbb{E}_t(\varphi X_{t+1})$$

$$= \delta + \varphi \hat{x}_{t+1} = \delta + \varphi(\delta + \varphi x_t) = \delta + \delta\varphi + \varphi^2 x_t = \delta(1 + \varphi) + \varphi^2 x_t.$$

- The forecasting error becomes

$$e_{t+2} = X_{t+2} - \hat{X}_{t+2} = \delta + \varphi X_{t+1} + \epsilon_{t+2} - \delta - \varphi \hat{X}_{t+1}$$

$$= \epsilon_{t+2} + \varphi(X_{t+1} - \hat{X}_{t+1}) = \epsilon_{t+2} + \varphi e_{t+1} = \epsilon_{t+2} + \varphi \epsilon_{t+1}.$$

- Therefore,

$$Var(e_{t+2}) = Var(\epsilon_{t+2} + \varphi \epsilon_{t+1}) = \sigma^2(1 + \varphi^2).$$

- By iterating the procedure, the general k-th steps ahead forecst will be

$$\hat{x}_{t+k} = \delta(1 + \varphi + \varphi^2 + \cdots + \varphi^{k-1}) + \varphi^k x_t$$

$$Var(e_{t+k}) = \sigma^2(1 + \varphi^2 + \varphi^4 + \cdots + \varphi^{2(k-1)}).$$

- As the horizon increases:

$$\lim_{k\to\infty} \varphi^k = 0$$

$$\lim_{k\to\infty}(1 + \varphi + \varphi^2 + \cdots + \varphi^{k-1}) = \frac{1}{1-\varphi}$$

$$\lim_{k\to\infty}(1 + \varphi^2 + \cdots + \varphi^{2(k-1)}) = \frac{1}{1-\varphi^2}$$

$$\lim_{k\to\infty} \hat{x}_{t+k} = \frac{\delta}{(1-\varphi)}.$$

- That is, for the $AR(1)$, the point forecast converges to the unconditional mean. If $\delta = 0$, then the forecast converges to zero.
- The variance tends to the unconditional variance of the process

$$\lim_{k\to\infty} Var(e_{t+k}) = \frac{\sigma^2}{1-\varphi^2}.$$

- The parameters in the model are replaced by their estimates $\hat{\delta}, \quad \hat{\varphi}, \quad \hat{\sigma^2}$.
- Same conditions hold for the $AR(p)$. (As k increases, the estimate becomes the unconditional mean of the process and the variance becomes the unconditional variance of the process.)

MA(1)

- Consider the $MA(1)$ process

$$X_t = \delta + \theta \epsilon_{t-1} + \epsilon_t.$$

- The one-step ahead forecast is given by

$$\hat{x}_{t+1} = \mathbb{E}_t(X_{t+1}) = \mathbb{E}_t(\delta + \theta \epsilon_t + \epsilon_{t+1}) =$$

$$= \delta + \theta \mathbb{E}_t(\epsilon_t) + \mathbb{E}_t(\epsilon_{t+1}) = \delta + \theta e_t.$$

- Assuming $\epsilon_t = e_t$, the one-step ahead forecast error is

$$e_{t+1} = X_{t+1} - \hat{X}_{t+1} = \delta + \theta \epsilon_t + \epsilon_{t+1} - (\delta + \theta e_t) = \epsilon_{t+1}.$$

- The expectation is null and the variance

$$Var(e_{t+1}) = Var(\epsilon_{t+1}) = \sigma^2.$$

- The two steps ahead forecast becomes

$$\hat{x}_{t+2} = \mathbb{E}_t(X_{t+2}) = \mathbb{E}_t(\delta + \theta\epsilon_{t+1} + \epsilon_{t+2}) =$$

$$= \delta + \theta\mathbb{E}_t(\epsilon_{t+1}) + \mathbb{E}_t(\epsilon_{t+2}) = \delta.$$

- The forecast with an $MA(1)$ remains constant and equal to the unconditional expectation of X_t after two steps ahead.
- The same holds for an $MA(q)$.
- In general: not trivial forecasts are obtained only for $k \leq q$. For horizon $k > q$ the forecast coincides with the unconditional mean of the model.

- The forecasting error is

$$e_{t+2} = X_{t+2} - \hat{X}_{t+2} = \delta + \theta\epsilon_{t+1} + \epsilon_{t+2} - \delta = \theta\epsilon_{t+1} + \epsilon_{t+2}$$

- It follows that

$$Var(e_{t+2}) = Var(\theta\epsilon_{t+1} + \epsilon_{t+2}) = \sigma^2(1 + \theta^2)$$

- for any step ahead with $k > 2$, the variance of the error becomes $\sigma^2(1 + \theta^2)$, that is the unconditional variance of the $MA(1)$ process.
- The same can be generalized for an $MA(q)$ process with variance $\sigma^2(1 + \theta_1^2 + \cdots + \theta_q^2)$.

Forecast with ARMA(1,1)

- Consider the $ARMA(1,1)$

$$X_t = \delta + \varphi X_{t-1} + \theta \epsilon_{t-1} + \epsilon_t.$$

- The uno step ahead forecast is given by

$$\hat{x}_{t+1} = \mathbb{E}_t(X_{t+1}) = \mathbb{E}_t(\delta + \varphi X_t + \theta \epsilon_t + \epsilon_{t+1}) = \delta + \varphi x_t + \theta e_t.$$

- The forecast error is

$$e_{t+1} = X_{t+1} - \hat{X}_{t+1} = \epsilon_{t+1}.$$

- Its variance is

$$Var(e_{t+1}) = \sigma^2.$$

- The forecast for $k = 2$ is

$$\hat{x}_{t+2} = \mathbb{E}_t(X_{t+2}) = \mathbb{E}_t(\delta + \varphi X_{t+1} + \theta\epsilon_{t+1} + \epsilon_{t+2}) = \delta + \varphi\hat{x}_{t+1} =$$

$$= \delta(1+\varphi) + \varphi^2 x_t + \theta\varphi e_t.$$

- The forecast error is

$$e_{t+2} = X_{t+2} - \hat{X}_{t+2} = \delta + \varphi X_{t+1} + \theta\epsilon_{t+1} + \epsilon_{t+2} - (\delta + \varphi\hat{X}_{t+1}) =$$

$$= \varphi e_{t+1} + \theta\epsilon_{t+1} + \epsilon_{t+2} = (\varphi + \delta)\epsilon_{t+1} + \epsilon_{t+2}.$$

- Notice that it has null expected value and variance equal to

$$Var(e_{t+2}) = \sigma^2((\varphi + \delta)^2 + 1).$$

- By iterating, the k steps ahead forecast is given by

$$\hat{x}_{t+k} = \delta(1 + \varphi + \varphi^2 + \cdots + \varphi^{k-1}) + \varphi^k x_t + \theta\varphi^{k-1} e_t.$$

- As k goes to ∞ the variance tends to the unconditional variance of the process.
- Notice that after the second step ahead the predictor resenbles that of an $AR(1)$. Indeed, its asymptotic behaviour is exactly that of an $AR(1)$.

$$\Longrightarrow$$

- When $k > 1$ the behaviour of the forecast is dominated by the autoregressive part.

- For a general $ARMA(p, q)$ similar results to those seen for the $ARMA(1, 1)$ can be obtained.
- When $k > q$ the autoregressive part drives the forecast that converges to the unconditional mean of the $ARMA(p, q)$ as k tends to ∞.
- Similarly, the variance of the forecast error converges to the unconditional variance.