M1 APE, Econometrics 2
Exercises: session 4
Stationary processes and forecasts

2019-2020

Exercise 1

Remember that in problem set 3 exercise 2, we saw that the equation defining (y_t) can be rewritten as $\left(1-\frac{1}{3}L\right)\left(1-\frac{1}{2}L\right)y_t = 0.5 + v_t$ and that (v_t) is the innovation process of (y_t).

1) For a forecast horizon $h = 1$, we have:

$$
\begin{aligned}
y^*_{t+1|t} &= \text{BLF}(y_{t+1}|\underline{y_t}) \\
&= \text{BLF}\left(0.5 + \frac{5}{6}y_t - \frac{1}{6}y_{t-1} + v_{t+1}|\underline{y_t}\right) \\
&= 0.5 + \frac{5}{6}y_t - \frac{1}{6}y_{t-1} + \text{BLF}(v_{t+1}|\underline{y_t}) \\
y^*_{t+1|t} &= 0.5 + \frac{5}{6}y_t - \frac{1}{6}y_{t-1} \qquad \text{since } (v_t) \text{ is the innovation process of } (y_t)
\end{aligned}
$$

$$e^*_{t+1|t} = y_{t+1} - y^*_{t+1|t} = v_{t+1}$$

$$V\left(e^*_{t+1|t}\right) = V(v_{t+1}) = \sigma^2_v$$

2) For a forecast horizon $h = 2$, we have:

$$
\begin{aligned}
y^*_{t+2|t} &= \text{BLF}(y_{t+2}|\underline{y_t}) \\
&= \text{BLF}\left(0.5 + \frac{5}{6}y_{t+1} - \frac{1}{6}y_t + v_{t+2}|\underline{y_t}\right) \\
&= 0.5 + \frac{5}{6}\text{BLF}(y_{t+1}|\underline{y_t}) - \frac{1}{6}y_t + \text{BLF}(v_{t+2}|\underline{y_t}) \\
y^*_{t+2|t} &= 0.5 + \frac{5}{6}y^*_{t+1|t} - \frac{1}{6}y_t \qquad \text{since } (v_t) \text{ is the innovation process of } (y_t)
\end{aligned}
$$

$$e^*_{t+2|t} = y_{t+2} - y^*_{t+2|t} = \frac{5}{6}(y_{t+1} - y^*_{t+1|t}) + v_{t+2} = \frac{5}{6}v_{t+1} + v_{t+2}$$

$$
\begin{aligned}
V\left(e^*_{t+2|t}\right) &= \left(\frac{25}{36} + 1\right)\sigma^2_v \qquad \text{since } v_t \sim \text{WN} \\
V\left(e^*_{t+2|t}\right) &= \frac{61}{36}\sigma^2_v
\end{aligned}
$$

3) For a forecast horizon $h = 3$, we have:

$$\begin{aligned}
y^*_{t+3|t} &= \mathrm{BLF}(y_{t+3}|\underline{y_t}) \\
&= \mathrm{BLF}\left(0.5 + \frac{5}{6}y_{t+2} - \frac{1}{6}y_{t+1} + v_{t+3}|\underline{y_t}\right) \\
&= 0.5 + \frac{5}{6}\mathrm{BLF}(y_{t+2}|\underline{y_t}) - \frac{1}{6}\mathrm{BLF}(y_{t+1}|\underline{y_t}) + \mathrm{BLF}(v_{t+3}|\underline{y_t}) \\
y^*_{t+3|t} &= 0.5 + \frac{5}{6}y^*_{t+2|t} - \frac{1}{6}y^*_{t+1|t} \qquad \text{since } (v_t) \text{ is the innovation process of } (y_t)
\end{aligned}$$

$$\begin{aligned}
e^*_{t+3|t} &= y_{t+3} - y^*_{t+3|t} \\
&= \frac{5}{6}\left(y_{t+2} - y^*_{t+2|t}\right) - \frac{1}{6}\left(y_{t+1} - y^*_{t+1|t}\right) + v_{t+3} \\
&= \frac{5}{6}\left(\frac{5}{6}v_{t+1} + v_{t+2}\right) - \frac{1}{6}v_{t+1} + v_{t+3} \\
&= v_{t+3} + \frac{5}{6}v_{t+2} + \left(\frac{25}{36} - \frac{1}{6}\right)v_{t+1} \\
e^*_{t+3|t} &= v_{t+3} + \frac{5}{6}v_{t+2} + \frac{19}{36}v_{t+1}
\end{aligned}$$

$$V\left(e^*_{t+3|t}\right) = \left(1 + \frac{25}{36} + \left(\frac{19}{36}\right)^2\right)\sigma_v^2$$

4) *There is a typo in the text: read* $h \geq 1$ *(and* $t \geq 2$*)*

<u>For $h = 1$:</u>

$$\begin{aligned}
\mathrm{BLF}(y_{t+1}|y_t,\ldots,y_1) &= \mathrm{BLF}\left(0.5 + \frac{5}{6}y_t - \frac{1}{6}y_{t-1} + v_{t+1}|y_t,\ldots,y_1\right) \\
&= 0.5 + \frac{5}{6}y_t - \frac{1}{6}y_{t-1} + 0 \qquad \text{since } \mathrm{Cov}(v_{t+1}, y_{t-k}) = 0 \quad \forall k \geq 0 \\
&= \mathrm{BLF}(y_{t+1}|\underline{y_t})
\end{aligned}$$

<u>For $h = 2$:</u>

$$\begin{aligned}
\mathrm{BLF}(y_{t+2}|y_t,\ldots,y_1) &= \mathrm{BLF}\left(0.5 + \frac{5}{6}y_{t+1} - \frac{1}{6}y_t + v_{t+2}|y_t,\ldots,y_1\right) \\
&= 0.5 + \frac{5}{6}\mathrm{BLF}(y_{t+1}|y_t,\ldots,y_1) - \frac{1}{6}y_t + 0 \qquad \text{since } \mathrm{Cov}(v_{t+2}, y_{t-k}) = 0 \quad \forall k \geq 0 \\
&= 0.5 + \frac{5}{6}y^*_{t+1|t} - \frac{1}{6}y_t \\
&= \mathrm{BLF}(y_{t+2}|\underline{y_t})
\end{aligned}$$

<u>For $h = 3$:</u>

$$\begin{aligned}
\mathrm{BLF}(y_{t+3}|y_t,\ldots,y_1) &= \mathrm{BLF}\left(0.5 + \frac{5}{6}y_{t+2} - \frac{1}{6}y_{t+1} + v_{t+3}|y_t,\ldots,y_1\right) \\
&= 0.5 + \frac{5}{6}\mathrm{BLF}(y_{t+2}|y_t,\ldots,y_1) - \frac{1}{6}\mathrm{BLF}(y_{t+1}|y_t,\ldots,y_1) + 0 \qquad \text{since } \mathrm{Cov}(v_{t+3}, y_{t-k}) = 0 \quad \forall k \geq 0 \\
&= 0.5 + \frac{5}{6}y^*_{t+2|t} - \frac{1}{6}y^*_{t+1|t} \\
&= \mathrm{BLF}(y_{t+3}|\underline{y_t})
\end{aligned}$$

It is easy to iterate and to see that $\mathrm{BLF}(y_{t+h}|y_t,\ldots,y_1) = \mathrm{BLF}(y_{t+h}|\underline{y_t})$ for any $h \geq 1$.

Remark: This property is true for any AR process, but <u>not</u> for an MA(q) or ARMA(p,q) process!

Exercise 2

1)i) We can compute a and b using the two conditions defining this linear predictor (forecast error of expectation 0 and of minimum variance).

The first condition yields:

$$\begin{aligned} E\big(X_{t+h} - (aX_t + b)\big) = 0 \quad &\Leftrightarrow m - am - b = 0 \quad \text{as } (X_t) \text{ is stationary} \\ &\Leftrightarrow \big[b = m(1-a)\big] \end{aligned}$$

Secondly, the variance of the forecast error can be rewritten as:

$$\begin{aligned} V\big(X_{t+h} - (aX_t + b)\big) &= V(X_{t+h} - aX_t) \\ &= V(X_{t+h}) + (-a)^2 V(X_t) + 2 \times (-a) \times \mathrm{Cov}(X_{t+h}, X_t) \\ &= \big(1 + a^2\big)\gamma(0) - 2a\gamma(h) \qquad \text{as } (X_t) \text{ is stationary} \end{aligned}$$

Let $f(a) = \big(1 + a^2\big)\gamma(0) - 2a\gamma(h)$. We have:

$$\begin{aligned} &f'(a) = 0 \Leftrightarrow 2a\gamma(0) - 2\gamma(h) = 0 \Leftrightarrow a = \frac{\gamma(h)}{\gamma(0)} \Leftrightarrow \big[a = \rho(h)\big] \\ &f''(a) = 2\gamma(0) > 0 \ \forall a \quad \text{so } a = \rho(h) \text{ is indeed a minimum} \end{aligned}$$

This predictor is thus obtained by choosing $a = \rho(h)$ and $b = m(1-a) = m(1-\rho(h))$.

ii) $BLF(X_{t+h}|X_t) = \alpha X_t + \beta$ has been defined in the course by the two following properties:

$$\begin{cases} E\big(X_{t+h} - (\alpha X_t + \beta)\big) = 0 \\ \mathrm{Cov}\big(X_{t+h} - (\alpha X_t + \beta), X_t\big) = 0 \end{cases} \Leftrightarrow \begin{cases} m - \alpha m - \beta = 0 \\ \gamma(h) - \alpha\gamma_y(0) = 0 \end{cases} \Leftrightarrow \begin{cases} \beta = m(1-\alpha) = m(1-\rho(h)) = b \\ \alpha = \dfrac{\gamma(h)}{\gamma(0)} = \rho(h) = a \end{cases}$$

The previous linear predictor does coincide with $\mathrm{BLF}(X_{t+h}|X_t)$.

This explains why $\mathrm{BLF}(X_{t+h}|X_t)$ is named as *best* linear forecast: it is the best one since it is associated to a forecast error which has minimum variance.

Remark: The coefficients a, b, α and β obviously depend on the forecast horizon h and we should rather denote them a_h, b_h, α_h and β_h (but we don't, in order to have simpler notations).

2)i) We can compute a_0, a_1 and b using the two conditions defining this linear predictor (forecast error of expectation 0 and of minimum variance).

The first condition yields:

$$\begin{aligned} E\big(X_{t+h} - (a_0X_t + a_1X_{t-1} + b)\big) = 0 \quad &\Leftrightarrow m - a_0m - a_1m - b = 0 \quad \text{as } (X_t) \text{ is stationary} \\ &\Leftrightarrow \big[b = m(1 - a_0 - a_1)\big] \qquad (1) \end{aligned}$$

Secondly, the variance of the forecast error can be rewritten as:

$$\begin{aligned} V\big(X_{t+h} - (a_0X_t + a_1X_{t-1} + b)\big) &= V(X_{t+h} - a_0X_t - a_1X_{t-1}) \\ &= V(X_{t+h}) + a_0^2V(X_t) + a_1^2V(X_{t-1}) \\ &\quad -2a_0\,\mathrm{Cov}(X_{t+h}, X_t) - 2a_1\,\mathrm{Cov}(X_{t+h}, X_{t-1}) + 2a_0a_1\,\mathrm{Cov}(X_t, X_{t-1}) \\ &= \big(1 + a_0^2 + a_1^2\big)\gamma(0) - 2a_0\gamma(h) - 2a_1\gamma(h+1) + 2a_0a_1\gamma(1) \qquad \text{as } (X_t) \text{ is stationary} \end{aligned}$$

Let $g(a_0, a_1) = \big(1 + a_0^2 + a_1^2\big)\gamma(0) - 2a_0\gamma(h) - 2a_1\gamma(h+1) + 2a_0a_1\gamma(1)$.

The first derivatives are:

$$\begin{aligned}
\frac{\partial g}{\partial a_0}(a_0, a_1) &= 2a_0\gamma(0) - 2\gamma(h) + 2a_1\gamma(1) \\
\frac{\partial g}{\partial a_1}(a_0, a_1) &= 2a_1\gamma(0) - 2\gamma(h+1) + 2a_0\gamma(1)
\end{aligned}$$

So the first order conditions yield:

$$\left\{\begin{array}{l} a_0\gamma(0) - \gamma(h) + a_1\gamma(1) = 0 \\ a_1\gamma(0) - \gamma(h+1) + a_0\gamma(1) = 0 \end{array}\right. \tag{2}$$

$$\Leftrightarrow \left\{\begin{array}{l} a_0 - \rho(h) + a_1\rho(1) = 0 \\ a_1 - \rho(h+1) + a_0\rho(1) = 0 \end{array}\right. \quad \text{dividing by } \gamma(0)\ (\neq 0)$$

$$\Leftrightarrow \left\{\begin{array}{l} a_0 = \rho(h) - a_1\rho(1) \\ a_1\left(1 - \rho(1)^2\right) = \rho(h+1) - \rho(h)\rho(1) \end{array}\right.$$

$$\Leftrightarrow \left\{\begin{array}{l} a_0 = \dfrac{\rho(h) - \rho(1)\rho(h+1)}{1 - \rho(1)^2} \\ a_1 = \dfrac{\rho(h+1) - \rho(1)\rho(h)}{1 - \rho(1)^2} \end{array}\right.$$

This extremum is a global minimum if the Hessian matrix (the second derivatives matrix) is positive definite.

$$H_g(a_0, a_1) = \begin{pmatrix} 2\gamma(0) & 2\gamma(1) \\ 2\gamma(1) & 2\gamma(0) \end{pmatrix}$$

Sylvester's criterion states that a Hermitian matrix is positive definite if and only if its leading principal minors (the determinants of the upper-left successive submatrices) are all strictly positive. This is the case here:

- the upper left term of $H_g(a_0, a_1)$ is $2\gamma(0) > 0$
- $\det\left(H_g(a_0, a_1)\right) = 4\left(\gamma(0)^2 - \gamma(1)^2\right) > 0$, since by Schwarz inequality, $(\mathrm{Cov}(X_t, X_{t-1}))^2 \leq V(X_t)V(X_{t-1})$

ii) $BLF(X_{t+h}|X_t) = \alpha_0 X_t + \alpha_1 X_{t-1} + \beta$ has been defined by:

$$\left\{\begin{array}{l} E\left(X_{t+h} - (\alpha_0 X_t + \alpha_1 X_{t-1} + \beta)\right) = 0 \\ \mathrm{Cov}\left(X_{t+h} - (\alpha_0 X_t + \alpha_1 X_{t-1} + \beta), X_t\right) = 0 \\ \mathrm{Cov}\left(X_{t+h} - (\alpha_0 X_t + \alpha_1 X_{t-1} + \beta), X_{t-1}\right) = 0 \end{array}\right. \quad \Leftrightarrow \quad \left\{\begin{array}{l} \beta = (1 - \alpha_0 - \alpha_1)m \\ \gamma(h) = \alpha_0\gamma(0) + \alpha_1\gamma(1) \\ \gamma(h+1) = \alpha_0\gamma(1) + \alpha_1\gamma(0) \end{array}\right.$$

This system corresponds to the equation (1) and the system of two equations (2), so that $\alpha_0 = a_0$, $\alpha_1 = a_1$, $\beta = b$. We have thus proved that $\mathrm{BLF}(X_{t+h}|X_t, X_{t-1})$ is the best predictor of X_{t+h} as a linear function of X_t and X_{t-1} since it minimizes the variance of the forecast error.

Remarks:

- *Again, all the coefficients obviously depend on the forecast horizon h and should actually be indexed by h, but we don't in order to have simpler notations.*
- *As it is written in the subject, the proof would be identical for* $\mathrm{BLF}(X_{t+h}|X_t, ..., X_{t-p})$ *which has been characterized in the course by*
  $$\left\{\begin{array}{l} \mathrm{BLF}(X_{t+h}|X_t, ..., X_{t-p}) = \alpha_0 X_t + ... + \alpha_p X_{t-p} + \beta \\ E\left(X_{t+h} - (\alpha_0 X_t + ... + \alpha_p X_{t-p} + \beta)\right) = 0 \\ \mathrm{Cov}\left(X_{t+h} - (\alpha_0 X_t + ... + \alpha_p X_{t-p} + \beta), X_{t-j}\right) = 0 \quad \forall j = \{0, ..., p\} \end{array}\right.$$
  and which can be proved to be the linear predictor which minimizes the forecast error variance. This property remains true for $\mathrm{BLF}(X_{t+h}|\underline{X_t})$ *but the proof is more complicated.*