EXERCISES

November, 2017

---

1. Stationarity and Autocovariance.

   (a) Is the following MA(2) covariance stationary?

$$y_t = \epsilon_t + 2.4\epsilon_{t-1} + 0.8\epsilon_{t-2}$$
$$\epsilon_t \overset{iid}{\sim} \mathsf{N}(0,1).$$

If so, calculate the autocovariance function and compute the forecasts $y_{t+k|t} = E_t y_{t+k}$.
ANSWER: an MA of finite order is always covariance stationary provided the variance exists (here it does since $\mathsf{Var}(\epsilon_t) < \infty$).
To compute the ACF, we let $\gamma_k = \mathsf{Cov}(y_t, y_{t-k})$

$$\begin{aligned}
\gamma_0 &= \mathsf{Var}(\epsilon_t + 2.4\epsilon_{t-1} + 0.8\epsilon_{t-2}) = \left(1 + 2.4^2 + 0.8^2\right)\mathsf{Var}(\epsilon_t) \\
&= 1 + 2.4^2 + 0.8^2 = 7.4 \\
\gamma_1 &= \mathsf{Cov}(\epsilon_t + 2.4\epsilon_{t-1} + 0.8\epsilon_{t-2}, \epsilon_{t-1} + 2.4\epsilon_{t-2} + 0.8\epsilon_{t-3}) \\
&= \mathsf{Cov}(2.4\epsilon_{t-1} + 0.8\epsilon_{t-2}, \epsilon_{t-1} + 2.4\epsilon_{t-2}) \\
&= (2.4 + 0.8 \times 2.4)\,\mathsf{Var}(\epsilon_t) = 4.32 \\
\gamma_2 &= \mathsf{Cov}(\epsilon_t + 2.4\epsilon_{t-1} + 0.8\epsilon_{t-2}, \epsilon_{t-2} + 2.4\epsilon_{t-3} + 0.8\epsilon_{t-4}) \\
&= \mathsf{Cov}(0.8\epsilon_{t-2}, \epsilon_{t-2}) = 0.8 \\
\gamma_k &= 0, \quad \text{for } k \geq 3
\end{aligned}$$

and the ACF is $\rho_k = \gamma_k / \gamma_0$.
Now the forecasts are

$$y_{t+k|t} = E_t y_{t+k} = E\left[y_{t+k} | y_t, y_{t-1}, ...\right]$$

and since $y_t = \left(1 + 2.4L + 0.8L^2\right)\epsilon_t$, provided the MA lag polynomial is invertible (check the roots!)

$$\left(1 + 2.4L + 0.8L^2\right)^{-1} y_t = \epsilon_t$$

so the information set at t, $\mathcal{I}_t$, is equivalently generated by the sequence of observables $\{y_{t-k}\}_{k\geq 0}$ or by $\{\epsilon_{t-k}\}_{k\geq 0}$. Hence

$$y_{t+k|t} = E_t y_{t+k} = E\left[y_{t+k} | \epsilon_t, \epsilon_{t-1}, ...\right]$$

so

$$\begin{aligned}
y_{t+1|t} &= E\left[\epsilon_{t+1} + 2.4\epsilon_t + 0.8\epsilon_{t-1} | \epsilon_t, \epsilon_{t-1}, ...\right] \\
&= 2.4\epsilon_t + 0.8\epsilon_{t-1} \\
y_{t+2|t} &= E\left[\epsilon_{t+2} + 2.4\epsilon_{t+1} + 0.8\epsilon_t | \epsilon_t, \epsilon_{t-1}, ...\right] \\
&= 0.8\epsilon_t \\
y_{t+k|t} &= 0, \quad \text{for } k \geq 3.
\end{aligned}$$

   (b) Is the following AR(2) covariance stationary?

$$y_t = 1.1y_{t-1} - 0.18y_{t-2} + \epsilon_t$$
$$\epsilon_t \overset{iid}{\sim} \mathsf{N}(0,1).$$

If so, compute the forecasts y_{t+}
ANSWER: To check whether y_t is stationary, we compute the roots of the AR lag polynomial:

$$\left(1 - 1.1L + 0.18L^2\right) y_t = \epsilon_t.$$

the roots are

$$\frac{1.1 \pm \sqrt{1.1^2 - 4 \times 1 \times 0.18}}{2 \times 0.18} = \{5.0, 1.11\}$$

i.e. they are both greater than unity in absolute value, so the AR(2) is stationary.
Now the forecasts are

$$\begin{aligned}
y_{t+k|t} &= E_t y_{t+k} = E\left[y_{t+k}|y_t, y_{t-1}, ...\right] \\
y_{t+1|t} &= 1.1y_t - 0.18y_{t-1} \\
y_{t+2|t} &= 1.1y_{t+1|t} - 0.18y_t = 1.1\left(1.1y_t - 0.18y_{t-1}\right) - 0.18y_t \\
&= \left(1.1^2 - 0.18\right) y_t - 1.1 \times 0.18y_{t-1} \\
y_{t+k|t} &= 1.1y_{t+k-1|t} - 0.18y_{t+k-2|t}
\end{aligned}$$

which can be computed forward.

2. Consider the VAR(1)

$$\begin{aligned}
x_t &= x_{t-1} + \varepsilon_t \\
y_t &= \beta x_t + \eta_t
\end{aligned}$$

where $(\varepsilon_t, \eta_t)' \sim \text{iid}\,(\mathbf{0},\Omega)$.

(a) Are x_t and y_t cointegrated?
ANSWER: We see that x_t is non-stationary and integrated of order 1 (a random walk). Also η_t is stationary so y_t is non-stationary since it is the sum of a stationary and of a nonstationary process.

$$\begin{aligned}
y_t &= \beta x_t + \eta_t = \beta x_{t-1} + \beta\varepsilon_t + \eta_t \\
&= \left[y_{t-1} - \eta_{t-1}\right] + \beta\varepsilon_t + \eta_t \\
y_t &= y_{t-1} + \beta\varepsilon_t + \eta_t - \eta_{t-1}
\end{aligned}$$

where $\beta\varepsilon_t + \eta_t - \eta_{t-1}$ is stationary, so Δy_t is stationary and y_t is integrated of order 1.
We see that $y_t - \beta x_t = \eta_t$ is stationary, so there exists a linear combination of y_t and x_t that is stationary: the two processes are cointegrated.

(b) Write the model in Vector Error (or Equilibrium) form (VEC).
ANSWER: We see from the expressions above that

$$\begin{aligned}
x_t &= x_{t-1} + \varepsilon_t \\
y_t &= \beta x_{t-1} + \beta\varepsilon_t + \eta_t
\end{aligned}$$

where $u_t = \beta\varepsilon_t + \eta_t$ denotes an *iid* process. Hence

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \beta & 0 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_t \\ u_t \end{bmatrix}$$

where

$$\begin{bmatrix} \varepsilon_t \\ u_t \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \beta & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_t \\ \eta_t \end{bmatrix} \sim \text{iid}\left(\mathbf{0}, \begin{bmatrix} 1 & 0 \\ \beta & 1 \end{bmatrix} \Omega \begin{bmatrix} 1 & 0 \\ \beta & 1 \end{bmatrix}'\right)?$$

We notice in passing that y_t does not Granger-Cause x_t.

(c) Suggest a method for computing $\partial x_{t+j}/\partial\eta_t$ for $j = 0, 1, ...$ (THIS IS A BIT MORE COMPLICATED)

ANSWER: The question is how to forecast future x_{t+j} when there is a shock to η_t. If ε_t and η_t do not correlate, then it's easy as the VAR rewrites

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \beta & 0 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ \beta & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_t \\ \eta_t \end{bmatrix}$$

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = A \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + B \begin{bmatrix} \varepsilon_t \\ \eta_t \end{bmatrix}$$

$$\begin{bmatrix} x_{t+j} \\ y_{t+j} \end{bmatrix} = A^{j+1} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \sum_{k=0}^{j} A^k B \begin{bmatrix} \varepsilon_{t+j-k} \\ \eta_{t+j-k} \end{bmatrix}$$

where $A^2 = \begin{bmatrix} 1 & 0 \\ \beta & 0 \end{bmatrix}^2 = A$ and so $A^k = A$. Hence

$$\begin{bmatrix} x_{t+j} \\ y_{t+j} \end{bmatrix} = A \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + AB \sum_{k=0}^{j} \begin{bmatrix} \varepsilon_{t+j-k} \\ \eta_{t+j-k} \end{bmatrix}$$

with $AB = \begin{bmatrix} 1 & 0 \\ \beta & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \beta & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \beta & 0 \end{bmatrix}$ and

$$\frac{\partial x_{t+j}}{\partial \eta_t} = 0$$

Now if ε_t and η_t correlate, with correlation ρ, then when η_t experiences a change, ε_t should change by $\rho\frac{\sigma_\varepsilon}{\sigma_\eta}$ times the change experienced by η_t (this is in fact the coefficient of the regression of ε_t on η_t) and we see that the impact of

$$\frac{\partial x_{t+j}}{\partial \eta_t} = \rho\frac{\sigma_\varepsilon}{\sigma_\eta}\frac{\partial x_{t+j}}{\partial \varepsilon_t} = \rho\frac{\sigma_\varepsilon}{\sigma_\eta}$$

3. Consider the following AR(1) data generating process (DGP)

$$y_t = \tau + \delta \times 1_{\{t>t_0\}} + \rho y_{t-1} + \varepsilon_t$$

where $1_{\{\cdot\}}$ denotes the indicator function that takes value 1 if $\{\cdot\}$ is true and zero otherwise. We assume $|\rho| < 1$.

(a) Is y_t defined above stationary?

ANSWER: no, since the expectation shifts at time t_0+1 from $\tau/(1-\rho)$ to $(\tau+\delta)/(1-\rho)$.

(b) We now want to compare several forecasting techniques. We write $y_{T+1|T}$ the forecast of y_{T+1} made at time T and $e_{T+1|T}$ the forecast error $y_{T+1} - y_{T+1|T}$. For each of the models below, compute the forecast errors $e_{t_0|t_0-1}$, $e_{t_0+1|t_0}$ and $e_{t_0+2|t_0+1}$, their expectations and variances.

(i) forecasting model $y^i_{t+1|t} = \tau + \rho y_t$

(ii) forecasting model $y^{ii}_{t+1|t} = \tau + \rho y_t + e^i_{t|t-1}$

(iii) forecasting model $y^{iii}_{t+1|t} = y_t$.

What forecasting model seems best to you?
ANSWER: Model (i).

$$\begin{aligned} y^{i}_{t_0|t_0-1} &= \tau + \rho y_{t_0-1}, \quad e^{i}_{t_0|t_0-1} = [\tau + \rho y_{t_0-1} + \varepsilon_{t_0}] - [\tau + \rho y_{t_0-1}] = \varepsilon_{t_0} \\ y^{i}_{t_0+1|t_0} &= \tau + \rho y_{t_0}, \quad e^{i}_{t_0+1|t_0} = [\tau + \delta + \rho y_{t_0} + \varepsilon_{t_0+1}] - [\tau + \rho y_{t_0}] = \delta + \varepsilon_{t_0+1} \\ y^{i}_{t_0+2|t_0+1} &= \tau + \rho y_{t_0+1}, \quad e^{i}_{t_0+2|t_0+1} = [\tau + \delta + \rho y_{t_0} + \varepsilon_{t_0+2}] - [\tau + \rho y_{t_0+1}] = \delta + \varepsilon_{t_0+2} \end{aligned}$$

Model (ii)

$$\begin{aligned} y^{ii}_{t_0|t_0-1} &= \tau + \rho y_{t_0-1} + [\varepsilon_{t_0-1}], \quad e^{ii}_{t_0|t_0-1} = [\tau + \rho y_{t_0-1} + \varepsilon_{t_0}] - [\tau + \rho y_{t_0-1} + \varepsilon_{t_0-1}] = \varepsilon_{t_0} - \varepsilon_{t_0-1} \\ y^{ii}_{t_0+1|t_0} &= \tau + \rho y_{t_0} + [\varepsilon_{t_0}], \quad e^{ii}_{t_0+1|t_0} = [\tau + \delta + \rho y_{t_0} + \varepsilon_{t_0+1}] - [\tau + \rho y_{t_0} + \varepsilon_{t_0}] = \delta + \varepsilon_{t_0+1} - \varepsilon_{t_0} \\ y^{ii}_{t_0+2|t_0+1} &= \tau + \rho y_{t_0+1} + [\delta + \varepsilon_{t_0+1}], \quad e^{ii}_{t_0+2|t_0+1} = [\tau + \delta + \rho y_{t_0} + \varepsilon_{t_0+2}] - [\tau + \rho y_{t_0+1} + \delta + \varepsilon_{t_0+1}] = \varepsilon_{t_0+2} \end{aligned}$$

Model (iii)

$$\begin{aligned} y^{iii}_{t_0|t_0-1} &= y_{t_0-1}, \quad e^{iii}_{t_0|t_0-1} = [\tau + \rho y_{t_0-1} + \varepsilon_{t_0}] - [y_{t_0-1}] = \tau + (\rho - 1)\, y_{t_0-1} + \varepsilon_{t_0} \\ y^{iii}_{t_0+1|t_0} &= y_{t_0}, \quad e^{iii}_{t_0+1|t_0} = [\tau + \delta + \rho y_{t_0} + \varepsilon_{t_0+1}] - [y_{t_0}] = \tau + \delta + (\rho - 1)\, y_{t_0} + \varepsilon_{t_0+1} \end{aligned}$$

and

$$\begin{aligned} y^{iii}_{t_0+2|t_0+1} &= y_{t_0+1}, \\ e^{iii}_{t_0+2|t_0+1} &= [\tau + \delta + \rho y_{t_0} + \varepsilon_{t_0+2}] - [y_{t_0+1}] = \tau + \delta + (\rho - 1)\, y_{t_0+1} + \varepsilon_{t_0+2} \\ &= \rho\, [\tau + \delta] + \rho\, (\rho - 1)\, y_{t_0} + (\rho - 1)\, \varepsilon_{t_0+1} + \varepsilon_{t_0+2} \end{aligned}$$

Now let's compare the models.

Model (i) corresponds to a model where t_0 is large so parameters τ and ρ are well estimated before the break, but since the break date is not known, the forecasting model becomes bad for as long as the break is not noticed: the forecast bias is δ.

Now Model (ii) is simply the same as Model (i) but where we add the forecast error made in the previous period. In general (when there is no break) this does not affect the forecast bias which stays at zero, but now the variance is double since the forecast error is $\varepsilon_{t_0} - \varepsilon_{t_0-1} = \Delta\varepsilon_{t_0}$ instead of ε_{t_0}. But this model becomes useful when there is a break since it self-corrects: the variance is always larger but from $t_0 + 2$ onwards, the forecast bias is back to zero, even if we do not know that a break has occurred!

Finally, for model (iii) the forecasts do not take at all into account the values of τ and ρ, so it can be used e.g. when the sample size is very small so that we cannot estimate (τ, ρ) accurately. (It is a model which is optimal when $\Delta y_t = \varepsilon_t$.) Then if ρ is small (say, close to zero) then the forecast bias for $t_0 + 2$ onwards is small even if we do not know there has been a break. When ρ is large, the bias becomes large, in this case if might be better even to combine Model (iii) into Model (iv) $y^{iv}_{t+1|t} = y_t + \Delta y_t$ which will eventually self correct (this is actually a model which is optimal when $\Delta^2 y_t = \varepsilon_t$)