Forecasting
Bonus Exercise

JY Le Boudec
Exercise: Compute the prediction

We have a times series $Y_t$. We computed the differenced time series $X_t = Y_t - Y_{t-1}$ and found that $X_t$ can be modelled as an AR process: $X_t = \epsilon_t + 0.5 X_{t-1}$ where $\epsilon_t \sim \text{iid } N(0, \sigma^2)$

1. Is this a valid ARIMA model?
2. Compute a point forecast $\hat{X}_t(2)$
3. Compute a point forecast $\hat{Y}_t(2)$
4. Compute the first 3 terms of the impulse response of the filter $\epsilon \rightarrow Y$
5. Compute a prediction interval for $Y_{t+2}$ done at time $t$
6. How would you compute a prediction interval using the bootstrap?
1. Is this a valid ARIMA model?

A. Yes because $X = F\epsilon$ and $F$ is an ARMA filter
B. Yes because $X = F\epsilon$ and $F$ is a stable ARMA filter with stable inverse
C. No because differencing is not a stable filter
D. I don’t know
Solution

The only thing to verify is whether the filter that defines the model for \( X \) is stable and has a stable inverse.

We have \( X_t - 0.5 \, X_{t-1} = \epsilon_t \) i.e.

\[
(1 - 0.5B)X = \epsilon
\]

\[
X = \frac{1}{1 - 0.5B} \epsilon
\]

The filter is \( F = \frac{1}{1 - 0.5B} \)

The zeros of the numerator polynomial are none \( \Rightarrow \) OK

The zeros of the denominator polynomial are: \( z - 0.5 = 0 \Rightarrow z = 0.5, |0.5| < 1 \Rightarrow \) OK

Answer B
2. The point predictions for $X$ are...

A. $\hat{X}_t(2) = X_t + 0.5\hat{X}_t(1), \quad \hat{X}_t(1) = 0.5X_t$

B. $\hat{X}_t(2) = 0.25X_t + 0.5\hat{X}_t(1), \quad \hat{X}_t(1) = 0.5X_t$

C. $\hat{X}_t(2) = 0.5X_t + 0.5\hat{X}_t(1), \quad \hat{X}_t(1) = -0.5X_t$

D. $\hat{X}_t(2) = 0.5\hat{X}_t(1), \quad \hat{X}_t(1) = 0.5X_t$

E. I don’t know
Solution

\[
\begin{align*}
X_{t+2} &= \epsilon_{t+2} + 0.5 \, X_{t+1} \\
X_{t+1} &= \epsilon_{t+1} + 0.5 \, X_{t}
\end{align*}
\]

We use as point forecast the conditional expectation of \( X_{t+2} \) given we have observed \( Y \) up to time \( t \). Note that \( L \) and \( F \) are invertible, therefore observing \( Y_{1:t} \) is the same as observing \( X_{1:t} \) or \( \epsilon_{1:t} \).

Take the expectation conditional to the observation up to time \( t \) of the above equations and obtain

\[
\begin{align*}
E(X_{t+2}|Y_{1:t}) &= 0.5E(X_{t+1}|Y_{1:t}) \\
E(X_{t+1}|Y_{1:t}) &= 0.5X_{t}
\end{align*}
\]

because \( E(\epsilon_{t+2}|Y_{1:t}) = E(\epsilon_{t+2}|\epsilon_{1:t}) = 0 \) and idem \( E(\epsilon_{t+1}|Y_{1:t}) = 0 \).

We can rewrite this as:

\[
\begin{align*}
\hat{X}_t(2) &= 0.5\hat{X}_t(1) \\
\hat{X}_t(1) &= 0.5X_t
\end{align*}
\]

Answer D
3. The point predictions for $Y$ are ...

A. $\hat{Y}_t(2) = \hat{X}_t(2) + 0.5Y_t(1), \hat{Y}_t(1) = \hat{X}_t(1) + 0.5Y_t$

B. $\hat{Y}_t(2) = \hat{X}_t(2) - 0.5Y_t(1), \hat{Y}_t(1) = \hat{X}_t(1) - 0.5Y_t$

C. $\hat{Y}_t(2) = \hat{X}_t(2) + \hat{Y}_t(1), \hat{Y}_t(1) = \hat{X}_t(1) + Y_t$

D. $\hat{Y}_t(2) = \hat{X}_t(2) + Y_{t+1}, \hat{Y}_t(1) = \hat{X}_t(1) + Y_t$

E. I don’t know
Solution

We use as point forecast the conditional expectation of $Y_{t+2}$ given we have observed $Y$ (hence $X$ and $\epsilon$) up to time $t$.

$$Y_{t+2} = X_{t+2} + Y_{t+1}$$
$$Y_{t+1} = X_{t+1} + Y_t$$

Take the expectation conditional to the observation up to time $t$ and obtain

$$\hat{Y}_t(2) = \hat{X}_t(2) + \hat{Y}_t(1)$$
$$\hat{Y}_t(1) = \hat{X}_t(1) + Y_t$$

Answer C
4. What is the impulse response of the filter $\epsilon \rightarrow Y$?

A. 1.000 -1.000 -1.500 ...
B. 1.000 -1.500 -2.250 ...
C. 1.000 1.500 1.750 ...
D. None of the above
E. I don’t know
Solution

Answer C

We have $Y_t - Y_{t-1} = X_t$, i.e. $(1 - B)Y = X$

Further, $X = \frac{1}{1 - 0.5B} \epsilon$

Therefore $Y = \frac{1}{(1-B)(1-0.5B)}$

The impulse response can be computed by power series calculus

$$\frac{1}{(1 - B)(1 - 0.5B)} = (1 + B + B^2 + \cdots)(1 + 0.5B + 0.25B^2 + \cdots)$$

$$= (1 + 1.5B + 1.75B^2 + \cdots)$$

or with matlab

```matlab
>> h=filter([1], [1 -1], filter([1], [1 -0.5], [1 0 0 0 0 0 0]))
h =
   1.0000  1.5000  1.7500  1.8750  1.9375  1.9688  1.9844
```
5. A prediction interval for $Y_{t+2}$ done at $t$ is ...

A. $\hat{Y}_t(2) \pm 1.96 \times \sigma$
B. $\hat{Y}_t(2) \pm 1.96 \times \sqrt{1.25\sigma}$
C. $\hat{Y}_t(2) \pm 1.96 \times \sqrt{2.25\sigma}$
D. $\hat{Y}_t(2) \pm 1.96 \times \sqrt{3.25\sigma}$
E. $\hat{Y}_t(2) \pm 1.96 \times \sqrt{4.25\sigma}$
F. $\hat{Y}_t(2) \pm 1.96 \times \sqrt{5.25\sigma}$
G. I don’t know
Solution

Answer D.

We have \( Y_{t+2} = \epsilon_{t+2} + 1.5 \epsilon_{t+1} + 1.75\epsilon_t + 1.875\epsilon_{t-1} + \cdots \) (eq. 1)

This is not a good formula for computing \( Y_{t+2} \) out of the complete series \( \epsilon_t \) because the coefficients become large (the filter \( \frac{1}{1-B} \) is unstable) and the error accumulates. It is better to use

\[
\begin{align*}
Y_{t+2} & = X_{t+2} + Y_{t+1} \\
X_{t+2} & = \epsilon_{t+2} + 0.5X_{t+1} \\
Y_{t+1} & = X_{t+1} + Y_t \\
X_{t+1} & = \epsilon_{t+1} + 0.5X_t
\end{align*}
\]

as we did earlier in order to compute the point forecasts.
Solution

\[ Y_{t+2} = \epsilon_{t+2} + 1.5 \, \epsilon_{t+1} + 1.75 \epsilon_t + 1.875 \epsilon_{t-1} + \cdots \quad (eq. 1) \]

However, (eq. 1) can be used to simplify the computation of prediction intervals. Observe that the red box is necessarily \( \hat{Y}_t(2) \) — to see why, take the conditional expectation given \( Y_{1:t} \).

In other words (Innovation Formula):

\[ Y_{t+2} = \epsilon_{t+2} + 1.5 \, \epsilon_{t+1} + \hat{Y}_t(2) \quad (eq .2) \]

which can be used to produce prediction intervals. Conditional to the observation up to time \( t \), \( \hat{Y}_t(2) \) is known (non random) and \( \epsilon_{t+2}, \epsilon_{t+1} \) are iid \( \mathcal{N}(0, \sigma^2) \), hence \( \epsilon_{t+2} + 1.5 \, \epsilon_{t+1} \) is \( \mathcal{N}(0, \nu) \) with

\[ \nu = \sigma^2 + (1.5)^2 \sigma^2 = 3.25 \sigma^2 \]

Therefore a 95%-prediction interval for \( Y_{t+2} \) done at time \( t \) is

\[ \hat{Y}_t(2) \pm 1.96 \times \sqrt{3.25} \sigma \]
6. Which is a correct implementation of the bootstrap for computing 95%-prediction intervals at time $t$ and lag 2?

A. A
   Compute the time series $\epsilon_s = X_s - 0.5X_{s-1}$, $s = 3: t$
   do $r = 1: 999$
   
   draw $e^r_s$, $s = 3: (t + 2)$ with replacements from $\epsilon_s$, $s = 3: t$
   compute $X^r_{1:t}$, $Y^r_t$ and $\hat{Y}_t^r(2)$ using $X^r_s = e^r_s + 0.5X^r_{s-1}$,
   $Y^r_s = X^r_s + Y^r_{s-1}$ and the formula for $\hat{Y}_t^r(2)$
   
   $Y^r_{t+2} = e^r_{t+2} + 1.5e^r_{t+1} + \hat{Y}_t^r(2)$
   
   Prediction interval is $[Y^{(25)}_{t+\ell}, Y^{(975)}_{t+\ell}]$

B. B
   Compute the time series $\epsilon_s = X_s - 0.5X_{s-1}$, $s = 3: t$
   do $r = 1: 999$
   
   draw $e^r_1, e^r_2$ with replacements from $\epsilon_s$, $s = 3: t$
   $Y^r_{t+2} = e^r_1 + 1.5e^r_2 + \hat{Y}_t(2)$
   
   Prediction interval is $[Y^{(25)}_{t+\ell}, Y^{(975)}_{t+\ell}]$
Solution

A is simulating the entire time series, therefore it is producing a sample of the unconditional distribution of $Y_{t+2}$. It is not the prediction, it is what can be said about $Y_{t+2}$ for an observer who knows the statistics of the time series but did not observe $Y_1, ..., Y_t$.

B is simulating the time series from $t + 1$ to $t + 2$ given the data up to time $t$, therefore it is producing a sample of the conditional distribution of $Y_{t+2}$ given the observed past. It is a correct implementation.

Answer B