Univariate Analysis Prediction

Applied Econometrics

Prediction

- Once we have characterized the trend and the cycle of a variables we can use this information to different purposes.
- It is possible to draw some economic conclusions (persistence, unit root consequences…)
- It is also possible to predict the future values of the series.

Persistence

- Half-life is the time required for a quantity to reduce to half its initial value.
- It was initially used in Physics, but this concept can also be used in Economics.
- A commonly used measurement is:

$$\text{HL} = \frac{\ln(0.5)}{\ln(\rho)}$$

Unit Root

- In some situations the presence of a unit root may invalidate some economic restrictions.
- This is the case of the purchasing power parity or the Fisher effect
- Then, if we may reject the unit root null hypothesis in real exchange rate or in the real interest rate, these economic restrictions hold.

Prediction

- We can use the ARIMA information in order to predict the future values of a variable.
- To that end, we need a sample size (T), to determine the number of periods ahead we want to predict (m) and the set of available information ($\Im_T$)

Prediction

- Then, it can be proved that the best predictor can be stated as follows:

$$\hat{y}_T(m) = E(y_{T+m} | \Im_T)$$

- Where the set of information is composed by all the observavions of the variable (y_1, …, y_T) and also by all the perturbations (u_1, u_2, …, u_T)

Prediction

- Imagine we have an AR(1) and we need to predict the future value of the period T+1.
- Then, the best predictor is the following one:

$$\hat{y}_T(1) = E(y_{T+1}|\Im_T) = E(\phi_1 y_T + \delta + u_{T+1}|\Im_T) = E(\phi_1 y_T|\Im_T) + \delta + E(u_{T+1}|\Im_T) = \phi_1 y_T + \delta + 0 = \phi_1 y_T + \delta$$

- If the parameters are unknown, we should substitute them by appropriate consistent

Prediction

- Now, let us imagine that we have an MA(1). Then, the best predictor is the following one:

$$\begin{aligned}\hat{y}_T(1) &= E(y_{T+1}|\mathfrak{I}_T) \\ &= E(\,\delta + u_{T+1} + \theta_1 u_T\,|\mathfrak{I}_T) \\ &= \delta + \; E(u_{T+1}|\mathfrak{I}_T) + E(\theta_1 u_T|\mathfrak{I}_T) \\ &= \delta + 0 + \theta_1 u_T = \delta + \theta_1 u_T\end{aligned}$$

- Again, the unknown values should be substituted by appropriate estimators

Prediction Error

In order to assess the goodness of the prediction, we should take into account the prediction error. This is defined as follows:

$$e_T(m) = \hat{y}_{T+m} - \hat{y}_T(m)$$

And, assuming that the classical hypotheses hold, it can be proved that:

$$\frac{e_T(m)}{\sqrt{\widehat{Var}[e_T(m)]}} \sim t_{T-k}$$

Prediction Error

For those cases where it is not possible the use of a linear estimation (presence of MA component), then it is easy to prove that:

$$\frac{e_T(m)}{\sqrt{var[e_T(m)]}} \overset{As}{\rightarrow} N(0,1)$$

Then, we can use this result asymptotically and substitute the unknown parameters by their corresponding consistent estimators.

Prediction Error

For an AR(1), we have that:

If m=1

$$
\begin{aligned}
e_T(1) &= y_{T+1} - \hat{y}_T(1) \\
&= \delta + \ \phi_1 y_T + u_{T+1} - (\delta + \phi_1 y_T) \\
&= u_{T+1}
\end{aligned}
$$

Consequently, it can be proved that:

$$Var[e_T(1)] = Var(u_{T+1})$$

Prediction Error. AR(1)

If m= 2

$$
\begin{aligned}
var[e_T(2)] &= y_{T+2} - \hat{y}_T(2) \\
&= \delta + \ \phi_1 y_{T+1} + u_{T+2} - \left(\delta + \phi_1 \hat{y}_T(1)\right) \\
&= u_{T+2} + \phi_1 y_{T+1} - \phi_1 \hat{y}_T(1) \\
&= u_{T+2} + \phi_1 [y_{T+1} - \hat{y}_T(1)] \\
&= u_{T+2} + \phi_1 u_{T+1}
\end{aligned}
$$

Prediction Error. AR(1)

Then, the variance of the prediction error is

$$var[e_T(2)] = var(u_{T+2} + \phi_1 u_{T+1})$$
$$= var(u_{T+2}) + var(\phi_1 u_{T+1})$$
$$+ 2\phi_1 cov(u_{T+2}, u_{T+1}) = \sigma^2 + \phi_1^2 \sigma^2$$
$$= (1 + \phi_1^2)\sigma^2$$

Prediction Error. MA(1)

For a MA(1), we have that:

If m=1

$$e_T(1) = y_{T+1} - \hat{y}_T(1)$$
$$= \delta + u_{T+1} + \theta_1 u_T - (\delta + \theta_1 u_T) = u_{T+1}$$

Consequently, it can be proved that:

$$Var[e_T(1)] = Var(u_{T+1})$$

Prediction Error. MA(1)

If m= 2

$$
\begin{aligned}
e_T(2) &= y_{T+2} - \hat{y}_T(2) \\
&= \delta + u_{T+2} + \theta_1 u_{T+1} - (\delta) \\
&= u_{T+2} + \theta_1 u_{T+1}
\end{aligned}
$$

Prediction Error. MA(1)

Then, the variance of the prediction error is

$$var[e_T(2)] = var(u_{T+2} + \theta_1 u_{T+1})$$
$$= var(u_{T+2}) + var(\theta_1 u_{T+1})$$
$$+ 2\theta_1 cov(u_{T+2}, u_{T+1}) = \sigma^2 + \theta_1^2 \sigma^2$$
$$= (1 + \theta_1^2)\sigma^2$$

Forecast accuracy

We dispose some statistics which can help us to assess the goodness of the prediction fit.

The most popular ones are MSE and the Theil’s U.

Forecast accuracy. MSE

MSE measures the average prediction error of a model.

$$\text{MSE} = \frac{\sum_{i=1}^{m} e_T(i)^2}{m}$$

Alternatively, we can also use the MAE

$$\text{MAE} = \frac{\sum_{i=1}^{m} |e_T(i)|}{m}$$