

# Lezione 5 - 07/10/2019

venerdì 1 novembre 2019 19:47

## Lezione 5 – 07/10/2019

### ARMA(p, q)

Let  $\varepsilon_t$  w.n.(0,  $\sigma^2$ ), then

$$Y_t = c + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}$$

is ARMA(p, q).

Stationarity of the whole ARMA(p, q) depends on the autoregressive part only: we have to check if the roots of

$$1 - \phi_1 z - \dots - \phi_p z^p = 0$$

are all outside the unit circle.

For invertibility, we require that the roots of

$$1 + \theta_1 z + \dots + \theta_q z^q = 0$$

are outside the unit circle.

#### ARMA(p, q)

Last week we studied MA model and AR. Excel file in Ariel and encourage us to practise.

For AR if I put -0.75 the today is the opposite of tomorrow.

AR(p)

$$Y_t = c + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + U_t$$

MA(q)

$$X_t = u + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}$$

How about calling  $X_t \rightarrow U_t$  and put the model on AR(p) and combine the two models.

This is a marriage of the two models and these will share characteristics of both model. We learnt that MA(q) is always stationary, but AR required more than stationarity but condition on p. It's required that are all outside the unit circle.

$\Phi(L) Y_t = c + \varepsilon_t$

To have stationarity I need  $\Phi^{-1}(L)$  exists, so that I can write:

$$Y_t = \Phi(L)^{-1} c + \Phi(L)^{-1} \varepsilon_t$$

Start from the first model ( $\rightarrow$  rossa) I rewrite using the Lag operator and line3 I collapse the polynomial with notation  $\Phi(L)$

I have this two polynomial and then I say "to help stationarity I need to be able to invert polynomial of  $\Phi(L)$ ".

To have stationarity we look at the moment, but there's another way look the formula on the red square. How to get to this formula? This is .. so I need that  $\Phi(L)^{-1}$  exists. This mean that the  $\Phi(L)$  inverting the  $\Phi(L)$ . To have stationarity I need to invert  $\Phi(L)$ .

LAG value of a constant is it's self, so we get  $\Phi(1)^{-1}$ .

We so last week that we can invert the polynomial if I can break the polynomial of order p and I can invert each one of them. So the way I will be able to invert the polynomial is the same trick we were dealing the last week. To check stationarity of ARMA is the same of AR.

I manage that the stationary depend on autoregress part only and checking the root of this guy  $1 - \dots$  are all outside the unit circle. We now know how to check for stationarity. Another was INVERTIBILITY.

How do I check for invertibility??

I do the three red line. And to have invertibility I want instead of  $Y_t$  as a function of innovation, I want  $Y_t$  in the function of the past.

$$Y_t = \sum \pi_j Y_{t-j} + \varepsilon_t$$

I want  $\Pi(L) Y_t = \varepsilon_t$

Put  $C=0$  so we don't worry about it. How do I move these guys to obtain ESP? I need that  $\Phi(L)^{-1}$  exist.

On past observations only. I cant take away  $\Phi(L)$  and past epsilon. To do it I need to invert  $\Phi(L)$  so that  $\Pi(L)$  will be  $\Phi(L) * \Phi(L)^{-1}$

How do I know if  $\Phi(L)^{-1}$  exist? Will be the same as we did with the MA(q).

Can I get the moment? Yes, it will be painful. At least we can get the mean to know if the process is stationary. And it's the same formula of mean of AR model.

For Autocovariances: we have to establish a couple of preliminary result. 1° preliminary result we know that  $\varepsilon_t$  is correlated with this time and not the past time.

2° preliminary I get the expansion in the formula of ARMA(1,1) and then I look at  $\varepsilon_{t-1}$  and the covariance. The process is stationary so if we have  $\varepsilon_t$  or  $\varepsilon_{t-1}$  it's the same.

**So I can now get the gamma values  $\rightarrow$  that are variance**

For  $\gamma_2$  will be very easy then

It's just  $\phi \gamma_1$ .

I saw the same thing on the autocovariances of AR(1). This ARMA(1,1) will look like an AR(1) for all the step after the first one. At the first steps only will look differently.

This show you this is actually correct. So dependence structure will combine both the effects.

If we go down getting autocorrelation we simply get it. If we are curious just stick some number on the excel file.

If I put -0.5 and -0.7 it will amplify the effects. So look like AR but MA will amplify the effects of AR.

In general if I have an ARMA(p, q) the autocorrelation and covariance will look like a car crash of AR and MA. But after 2 will look like and AR. It's like AR with a bit of more story.

**Will be a question of this** 😞 😞 😞

How I derive the IRF: we did it the same for AR(1).

La pagina con tante somme si può saltare. E' uguale a quella di prima ma in modo differente.

### COMMON FACTORS

Excel

Put 0, 0. This is a white noise process. So let's focus on this.

$Y_t = \varepsilon_t \rightarrow$  so this is a white noise. So autocorrelation does not depend on the past.

Look at ARMA(1,1)

$$Y_t = -0.5 Y_{t-1} + \varepsilon_t + 0.5 \varepsilon_{t-1}$$

Write -0.5 and 0.5

And I get the same and it's the autocorrelation of white noise. It's easy to establish that

$$Y_t = -0.5 L Y_{t-1} + \varepsilon_t + 0.5 L \varepsilon_{t-1} \text{ and we get } (1 + 0.5L) Y_t = (1 + 0.5L) \varepsilon_t$$

We see that these factors are the same so we may simplify them.

What we get?? We get exactly the white noises that we started.

ARMA(1,1) that will look and act as it was white noises. It's because we cancelled the factors.

As long as we have a factor that are the same in the left and right we can cancel them.

So we can get the smaller model because the bigger is called **over parametrized**. Would I want to cancel them? YES if I want to obtain white noise.

If I get ARMA(2, 1) I can simplify in AR(1). It's simpler to manage AR(1) instead of ARMA(2,1) model.

If you do not know the parameter of the model we will not be able to estimate these guys.

These are equally valid representation of the same objects. All model are good but depends on what we want to do. MA if past innovation, representation that use the smallest number of parameter I use ARMA. But all are all the same.

### Stationarity and Ergodic ARMA.

They are all stationarity and ergodic so are identically distributed white noise.

Why I like ARMA models??

- First one is sum of ARMA model we still get ARMA model. What goes for the individual goes for the community
- Second one are forecasting! The problem is that we needed to know all autocovariances, and if I had 100 observation I don't want to estimated 100 covariances. I need  $\Phi$  to get all these autocovariances.

### BUT I GETS BETTER

Imagine that I have AR(1).

$$Y_t = \dots$$

What AR(1) tells you is that  $Y_t$  depend on the past ( $Y_{t-1}$ ) and  $\varepsilon_t$  is the new information that came right now. Let's think of  $Y_{t+1}$ .

$$Y_{t+1} =$$

Depends on today and new information that is coming tomorrow ( $\varepsilon_{t+1}$ ). Actually, I can observe  $Y_t$ , but I will never observe  $\varepsilon_{t+1}$ . The best option is to forecast  $Y_{t+1} | t$  and the value of today.

So the big formula over there will simplify in this easy formula:

$$Y_{t+1} | t = \Phi Y_t$$

AR is good when we want to forecast because a complex formula to a very simple.

How about MA(1)? It's different.

In the MA(1)

$$Y_t = \varepsilon_t + \theta \varepsilon_{t-1}$$

$$Y_{t+1} = \varepsilon_{t+1} + \theta \varepsilon_t$$

I got the  $\varepsilon_t$  but not the  $\varepsilon_{t+1}$ . I don't observe  $\varepsilon_t$ , I only observe  $Y_t$ . If I had  $\varepsilon_t$  I just forecasting it right way. The problem is that I don't have it, because I have the observation ( $Y_1, Y_2, \dots$ ) but not  $\varepsilon_t$ . But if the process is invertible I can derive  $\varepsilon_t$  from the past. This suggest that I could get away with this, but I have to go back infinite to past. But what is the weight that we give to infinite past? A power of  $j$ .  $0.5^{100}$  is nearly 0, so instead of going all the way to infinity, after big observation we truncated it.

We assume that  $\varepsilon_0 = 0$  so we assume that nothing happened before our first observation.

Setting  $\varepsilon_{t+1}$  to zero will be a very small mistakes.

Example:

AR(2) with 3 values and we see how forecast will look like.

For MA(1) let's pretend that I observe 10 observation for  $Y_t$ . This will be not the true value of  $\varepsilon_t$  but it's faster. It's gives me the same forecast with a error of 0.0002. The first usually is better because the second it's huge. When I have ARMA model I can put the in matrix because it's heavy. We should use the formula to get the approximate forecast knowing that estimation is pretty good. ARMA model are cool because forecast is much much easy than we saw at the very beggins. Answer the two question an give me a flexible react to difficult example.