Lezione 2 – 24/09/2019

venerdì 1 novembre 2019 19:40

Lezione 2 – 24/09/2019

[some examples]

Alcuni esempi che discuterà nel futuro. We discuss this model and we will be able to come back. When we revise we say “ah this go for this”.

Second model is called M average 1. We will find why is the first stationary. 1) is egotic because depends on something happening in before.

Last model is A R 1. We will be able to say that 3 is stationary. At this stage i’m not proving that is stationary but we will go over it in a week.

Pretend that they are normally distributed as well. The join density is the big formulas. Depend on the mean and the variance and the correlation between the two variable that is ‘p’.

The join density became the product of the two marginal. So that means that the two variables are independent. They are asymptotically independent. Yt does not depend on the past or the dependence is very low.

3. examples Yt does depends on the past and go away to -j. Asymptotically are independent. How I we establish that? Is not so easy, we’re there is nothing coming from the past. Density is the product of the two marginal and the probabilistic condition of asymptotic independent. That means they looks like they are independent. The dependence from the past is very very small.

We don’t know anything about this model.

CAPITOLO 2

[SAMPLE MOMENTS]

LLN and CLT for stationary and ergodic.

This model works with epsilon -1. The sample averages

There’s a pattern: if I sum it 3 time we have 3 times mu, if I some it n times I will have n times mu.

LLN I have to divided to the T.

The porpuse is that we have a road map where we have

CLT we just put the root of T. So N(0 , var ^2)

The ratio if T goes to infinity, the ratio goes to 0. If we multiply for Teta, the variance go multiply for teta^2.

MA1 allows me to establish the low larger number. We were able to do inferential statistics.

I will be able to do the same, but we have teta1 … +teta2…

There’s not particular reason. Beveridge-nelson composition to define the central limit theory (CLT).

I can in fact allow ax

(il simbolo con u incrociata da |) to go to the infinity providing that is going very very near to zero. This is convenient for 0.5 to the power of j. It’s never 0 but get closer every time.

The study for the more general NO? I have to study for the CLT [slide]

If I have a process that depends on the past provided that these dependence is not very strong because go down fastly.

Interesting question is how to estimate N (0 , …)? We want to know how precise the measure is. We understand how to read variance in the contest of random variable. Sigma square is different from the variance of the sample average. We gave a different name anche we call it long rule variance??

We can also establish the lowest larger number?? For the estimate to be consistent we need sample autocovariance. We need the element of the sum to go to infinity. When I look at the data we don’t normally presume that we know sigma square is normalized. We estimated it. I have to estimate sigma square and the sum. The is 2 ways to do this.

Autocorrelation robust inference on the sample mean

I will have to way o estimate 1. Estimate covariance j individually or sij individually. No we can’t use to estimate it. I need a infinite number of elements.

When I have T = 100 and estimating gamma1 = 1/ 100 SUM j = 2 j <100 (Yt – $\underline{Y}$) (Y_{t-1} – $\underline{Y}$).

The second covariance will be : gamma2 = 1 / 100 SUM j =3; j < 100 (Yt – Y) (Y_{t-2} – Y)

So I use 88 observation.

Covariance of 99:

gamma3 = 1/100 (Y1 -Y) (Y100 -Y) → we have one pair useful for estimation and the estimation will no be consistent. I use 1 out of 100 estimation so is not consistent. If sample go bigger I can grow up the number of gamma.

Autocovariance will go fast to zero, so I can treats them all like 0.

There’s two way to estimate gamma. IT’s simple to say we have 100 and we use first 10 covariance . The second way is to have 100 observation and I go all the way to 10, and the I will use weight that decreases.

Gammaj nell’esempio dovrebbe essere 0!!

I will gave a weight of 9/10 , 7/10 so the farther away I go the consistent of estimation. This way is called **KERNEL.**

They are both fine, but the rectangular kernel has a problem because you must know that covariance is not negative. SO triangular is better.

The rule is to user square root of T is a optimal choice. And this is one way to estimate covariance and it’s cool because it works every time.

Why I don’t just estimate teta? I have just to estimate FI. If I do it I have an alternative way to estimate and this is called **parametric estimation** → assume another piece of information.

Both techniques work, they give us a consistent estimated that we can plug in to the formula of the covariance. Which one we should prefer? The parametric estimation will be more precise. The unappalled part is the parametric required that we know the correct model. We don’t necessary require this assumption. It’s safer to use non-parametric estimation. It’s safer!

CAPITOLO 3 FORECASTIN AND IMPULSE RESPONSE FOR STATIONAY PROCESSES

This was an introduction to this part. It’s a fact that we have time series, but when I have a time series why am I interested in studying time series: we are here to know the answer to all these questions. If I know the temperature of today, yesterday ecc.. to define a future prevision.

The fact that last week was raining change our temperature? In a heuristic way we can say that makes sense. 3 months ago was raining? No sense. So we are interested to define what are consequence of today from events happened in the past. This Is called IMPULSE RESPONSE.

FORECASTING: I want to forecast today the future value. To answer this questions we need statistical theory. LLN(low large number) and CLT (central limit theorem).

To answer first question we get to wold decomposition

Wold decomposition said that any stationary process may be represent in the form [SLIDE] where epsilon t sono gli errori compitui in forecasting. Yt is depending on the past. It’s a linear formula and it’s just sums. This means that if I have a stationary model I have a simple model that fits the data. If I have a process that is not linear (this guys impose linearity) you could give it a linear representation. THAT’S THE HOT THING

Model 1) is not linear and they are multiplying each other. Even thought is not linear we may thing linear model that will generate Yt. So I can use a linear model to describe Yt even if this is not a linear function. The long if I have stationarity, then I will find linear model that fits the data. Using linear model is much easier than using power 2 or power of ¾.

I can take any stationary I can estimate the contribution of yesterday, of day-2, day-3 ecc.

Impulse response

The derivate of Y respect to t-j → this Is exact the PSI j.

PSI j is the effect of what happen yesterday that define the effect of today (PSI).

Linear filter

PSI(L) epsiolont = (PSI0 + PSI1L + PSI2 L^2 …) = PSI0 EPSt + PSI L EPS2

Where L is the lag operator. I can rewrite the formula in PSI (L) EPS t.

The interesting thing is that I can combine this filter so what happen when I take Yt-1 or I sum them? I generate another filter and I new process will still have the property of Yt and has the same condition. When I have stationary series I take a linear transformation of these series I still have stationary series. The meaning is when I have stationary time series and I transform it I will obtain another stationary time series.

Invertibility

If I have these impulse response function and I have Yt and the weight to all the bast(??) shock I know that the effect of the shock of 3 days ago is PSI. Can I measure the shock of PSI 3 days ago? To do this, we rewrite our formula using the polynomial operator. Typically, I will be interesting of the input compulse function. I will be able to answer the question of how will be shocked today from yesterday.

What is the forecast of today with the information that we have at this point? NEW WEEK.