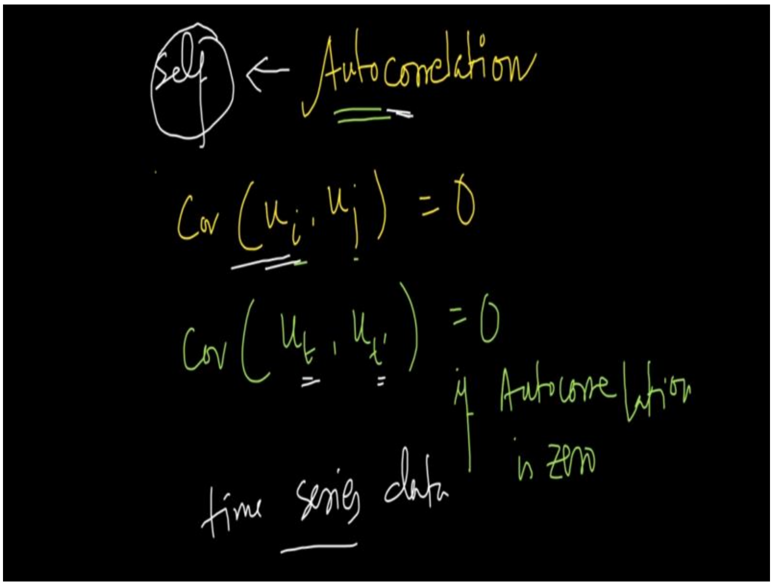
Applied Econometrics
Prof. Tutan Ahmed
Vinod Gupta School of Management
Indian Institute of Technology – Kharagpur

Lecture – 70
Autocorrelation

Hello and welcome back to the lecture on applied econometrics. So, in this lecture, we are going to talk about autocorrelation.

(Refer Slide Time: 00:31)



Now, autocorrelation is a problem, where we have our regression model, it may, if the autocorrelation is present. Then, it is we need to actually address that like we have seen in the previous cases heteroscedasticity, multicollinearity. If those issues are present in our regression model, then we need to actually do the diagnostics and actually, you know, find a way out to sort of get rid of those errors or those problems.

Similarly, if autocorrelation is present in our model, so, we sort of need to find a way to actually get rid of autocorrelation. Now, what is autocorrelation in the first place? Now, we know for trauma, the Gauss Markov assumptions. If we need to have sort of the best if Linear Unbiased Estimator for the regression equation.

So, what you need to have is that, we need to kind of fulfill few criteria and for that we need to have one criteria fulfilled that is covariance among the error terms. Let us say, I will write i and U j let us i and j be 2 different observation has to be 0. So, essentially it means that the error term related to observation i and the error term related to observation j are uncorrelated there is no correlation.

Now, particularly for autocorrelation we use the term t instead of i and j. So, how we actually write that, we write it as covariance of U t and U t prime let us say is equal to 0. So, that is what we have to have, if autocorrelation is 0. Now, why do we have this different notations? Well, essentially autocorrelation is something that we will see in time series data mostly, it is related to time.

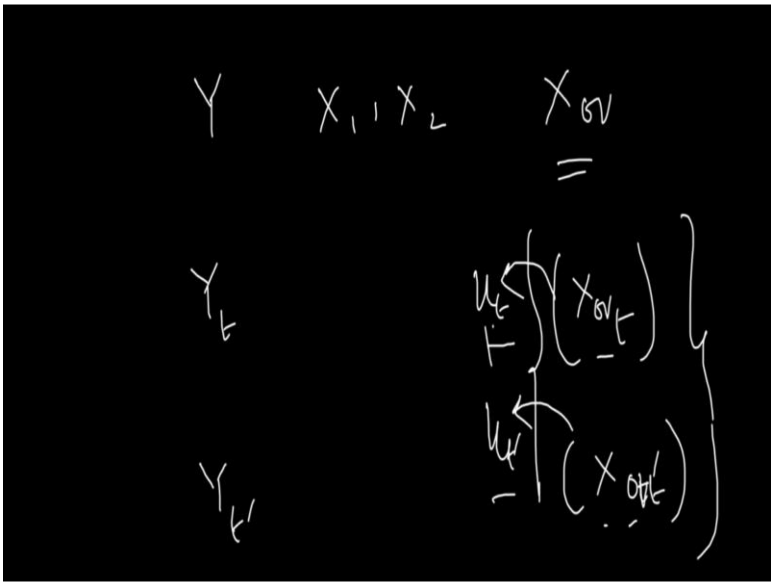
The very term auto, is from the Greek word an auto means self. So, basically I am correlated with myself. So, if I am need to correlate, I need to be correlated with myself. The data actually the data generating process would require me to have some sort of time dimension. So, that is why autocorrelation is a problem that is mainly seen in case of time series data.

For example, if you take let us say GDP, you will find if, you will actually need to look for, if the GDP of one period is related to the previous period, which usually is the case. And let us say, if you look at temperature data, for example, which we are going to see in this talk. So, you will need to actually see, if the temperature data of this period is actually correlated with the previous period.

So, essentially autocorrelation now, you have learned previously heteroscedasticity and we know for heteroscedasticity, the formula is something like this autocorrelation heteroscedasticity is pretty close. But in case of autocorrelation, we have the connotation of self is involved. Heteroscedasticity you can have 2 different variables. But, in case of autocorrelation, it is only with the self.

And the time dimension is important in case of autocorrelation. Now, if we want to actually see further what happens because of autocorrelation, or why autocorrelation is present? So, let us first talk about why autocorrelation present.

(Refer Slide Time: 04:11)

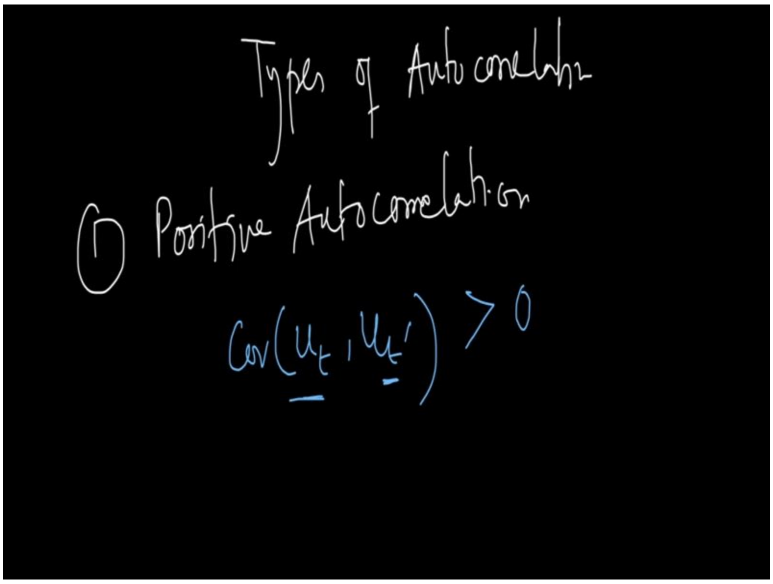


Autocorrelation usually, it happens because of some omitted variable. Let me explain that. So, let us say, Y is actually determined by some you know, X 1 and X 2, but there is another omitted variable X ov, let us say. Now, when I have and that is actually out of my model, now, when I have Y at time period Y t. Now, the X ov will have a value of X ov t, let us say now, this X ov t is actually out of this model, it is not included.

Now, Y t prime when I measure Y t in Y t prime time. So, I have my Y t prime and then there is an X ov t prime. Now, because you are omitting this X ov t and X ov t prime. So, whatever pattern is there in X ov t, or X ov t prime, or v t prime. So, then that part will be reflected in the error term. So, the U t and U t prime will actually capture this X ov t and X ov t prime will contribute to U t and U t prime.

And that is the reason, because the U t and U t prime is capturing this X ov t and X ov t prime, it will actually show that there is some pattern in the X t U t and U t prime.

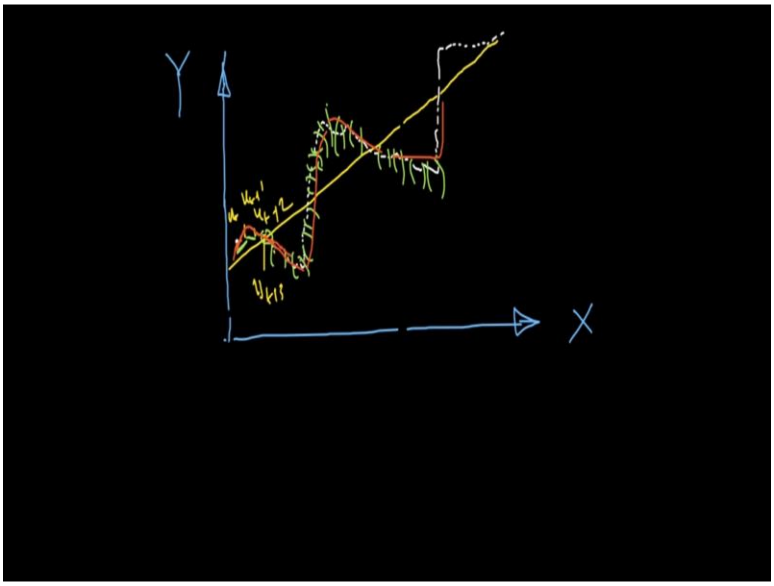
(Refer Slide Time: 05:36)



Now, going by this, we can actually see that autocorrelation mostly are of 2 types. Types of autocorrelation, if we look at mostly of 2 types. First, it is a positive autocorrelation. Let me explain what I mean by positive autocorrelation. So, in case of positive autocorrelation, we will the essentially this value, let me use a different colour, this covariance U t and U t prime, they will lead to a value which is greater than 0, let us say equal to greater than 0.

So, if it has to be greater than 0. So, in at what condition we will have the covariance between 2 variables is actually greater than 0. So, if they both have a positive, the same sign, either they both are positive or both negative. So, what I mean by they both are positive, and they both are negative. So, to illustrate that, let me actually draw this.

(Refer Slide Time: 06:42)



So, let me actually draw. Let us say, this is my independent variable along the X axis, and my dependent variable on the Y axis. So, let us say this is my X, and this is my Y. Let us say this only X is what is determining the Y. And then I have my regression line, let me draw my regression line with a different colour. And then let us say we have our error terms. Let me actually plot this errors. So, let us say these are my error terms.

And you can see that I am actually plotting something perhaps that is a pattern. And of course there has to be, if I actually want to illustrate the autocorrelation. So, let us say we have something like this, and this and this, and this and this and this. So, essentially, if I take U t and U t. So, let us say, if we take this error terms in successive period. Let us say, this is U t, and this is my U t +, or I should not write it here is X axis.

Let us say this is U t, this is U t + 1, this is U t + 2, this is U t + 3, and so, forth. So, essentially, if I take correlation between U t and U t prime U t + 1. So, then I will have a positive value, because they both are positive. Similarly, U t + 1 and U t + 2, they are positive, they are above the regression line. The both the errors are wearing the same symbol, they are both positive.

But perhaps when I, you know change the direction, when I move from positive to negative. Then I will have my correlation coefficient that will be a negative sign, because there are terms of that opposite signs. But then, here again, if I just repeat the same exercise all throughout this lines. So, I will have all these values are going to be positive. Because they are on the same side of the regression line.

So, both the error terms are negative. So, when 2 negative terms, I take correlation coefficient, they are going to basically give me a positive term. Similarly, on the other side, all positive error terms and they are going to give me a positive correlation coefficient. And this is precisely, why we call this this autocorrelation to be positive autocorrelation.

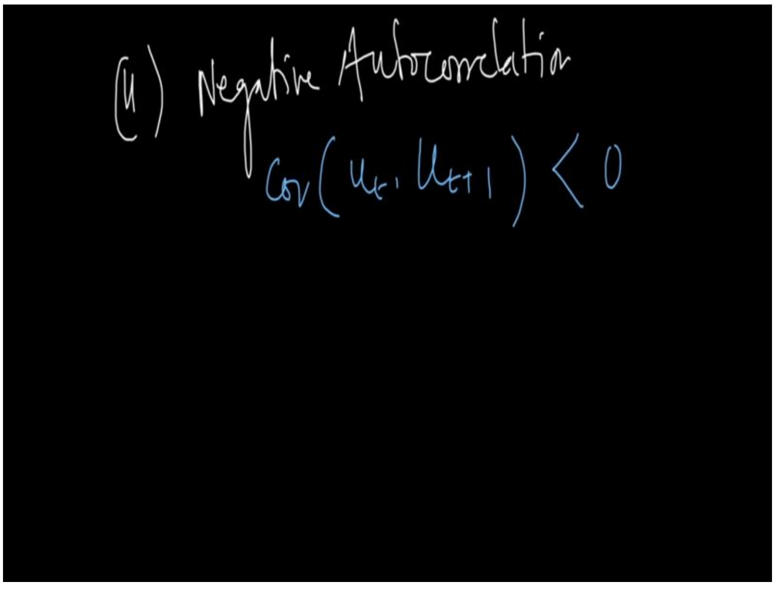
Because whenever I take the error terms the correlation coefficient of the error terms, as long as they are on the same side of the line, we are going to get a positive value. So, it does not matter, if my line is on the positive side or the negative side. All I care is, if 2 successive values are bearing the same side. So, in this kind of examples, the autocorrelation we see is a positive autocorrelation of course, we have autocorrelation because we can see the residual term that we have that actually is bearing a pattern.

So, that is actually showing some sort of sinusoidal pattern. So, it is going up and coming down and going up and so forth. So, there is a pattern so, that means they are related to each other. And the relation essentially as we understood now, is a positive autocorrelation. Mostly, we will see in problems in economics, we will mostly see positive autocorrelation, and not much of negative autocorrelation.

And I will explain why is that, but before that I need to explain what I mean by negative autocorrelation? Just one more point to note, this amplitude and the cycle of this positive autocorrelation, that is absolutely dependent on data. So, we really do not have you know that is random, we really cannot say anything. But we can say about positive articulation, that is a slow process.

So, every error term is actually you know correlated with the previous term, connect and they had same sign. So, essentially, it has to be a slow process. It cannot be very abrupt, it cannot go up and down, it cannot go up and down, it cannot do that. So, that is the positive autocorrelation part.

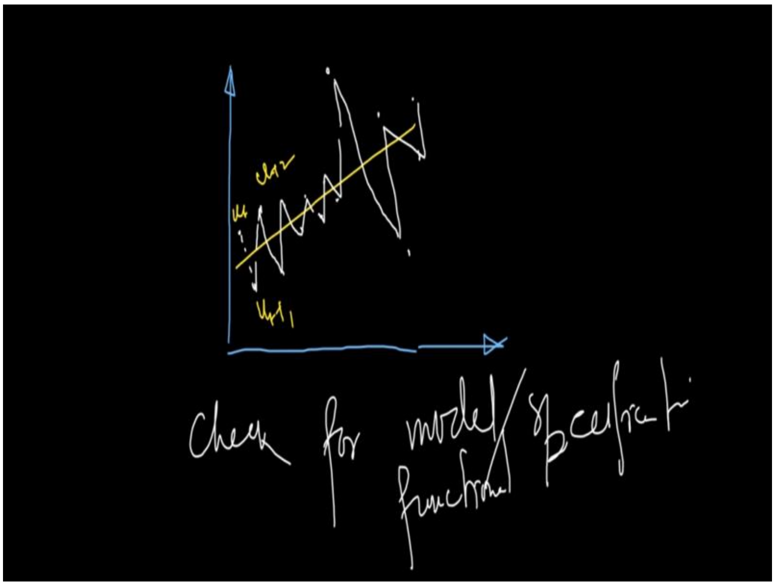
(Refer Slide Time: 11:08)



So now, let us talk about negative autocorrelation. The other type of autocorrelation that we are talking about negative autocorrelation. And what is negative autocorrelation? I hope we are maintaining the same colour consistency. Let me just check for that. So, in case of negative autocorrelation, we will have covariance U t, and U t + 1, U t - 1, it does not matter, is actually less than 0.

So, when that is possible, covariance U t and covariance U t + 1 is less than 0.

(Refer Slide Time: 11:51)



That is possible, if we have something like this. Let us say we have the same regression line, but the points they need to be on the 2 sides of this line. So that, whenever I take the successive points, are the successive error terms. They actually have different signs. So, that is all I care. Whether the successive observations are and the error terms basically are of different signs.

So, if I draw this, it is going to be something like this can go up or down and again up and down and up. So, essentially, I can see that the error term, U t here use a different colour that U t here, I have deleted some of the lines I just drew. So, U t here, U t, and U t + 1 and U t + 2, they are on the opposite side of the regression line. So, one is positive and the next is negative.

Again, the next one is positive and next one is negative, and so forth. So, as long as they are on the different sides of the regression line, my this condition is going to be satisfied. Now, that is what is called negative autocorrelation. Because of the value of the covariance among these 2 error terms. Now, that essentially means the process has to be highly sort of, you know fluctuating.

So, if the stock is going up in this period, you know, it is going to go down tomorrow and again, day after tomorrow is going to go U? So, that is really very you know sort of extremely fluctuating process. And in economics usually we do not have this kind of data you know, we can have but not usually. Mostly, I mean, in case you find negative autocorrelation, it is very important that we check for model specification or functional specification.

Because the very formula we write that might have a component that will sort of represent the model like it has a it looks the error terms are looking like this zigzag pattern. So, check for model specification or functional specification model or functional specification, if we have such kind of negative autocorrelation. So, this is about the introduction to autocorrelation and in the next lecture, we are going to give an example of autocorrelation. Thank you.