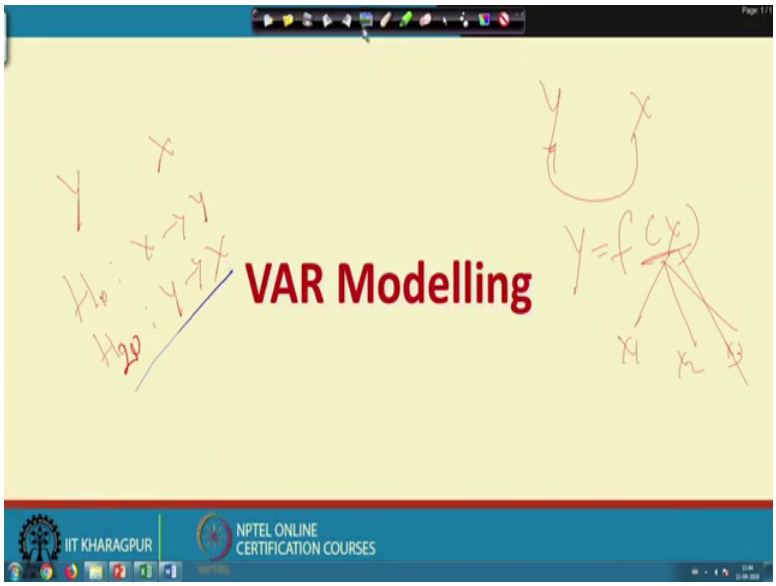
Engineering Econometrics
Prof. Rudra P. Pradhan Vinod Gupta School of Management
Indian Institute of Technology, Kharagpur

Lecture – 51
Time Series Modelling – VAR modelling

Hello everybody. This is Rudra Pradhan here, welcome to Engineering Econometrics that to the issue of time series modelling. In the earlier lectures we have already discussed various time series modelling, starting with autoregressive, moving average, ARIMA then arch cluster, GARCH clusters. And today in the similar flow, we will continue with another cluster that is what technically called as VAR modelling.

And it is simply back the scheme of vector autoregressive, we have already discuss a concept called as auto regressive modelling in the case of ARIMA cluster. But here the additional component is the vector part and as a result the particular structure will be better autoregressive modelling. So, that means, it is the cluster of at least, 2 different variables with time series you know structure that is what the technical things to start with this particular engineering problem.

(Refer Slide Time: 01:52)



So, what we like to do? We like to highlight actually the concept of VAR modelling and that to that to the where the situation arise when there are 2 variables and we like to predict whether they are predicting each other or not.

Of course we have discussed this issue, but now we like to discuss in a kind of, time series framework. For instance, 2 kinds of variables that is Y and X then, we are usually interested to know whether X is causing Y that is what we have already discussed in the framework of simple regression modelling. So, by default we link like this Y equal to function of X, where X maybe a vector and that way it represents like series of independent variables like X 1, X 2 and X 3. So, X 1, X 2, X 3 are series of independent variables and that to why is the dependent variables.

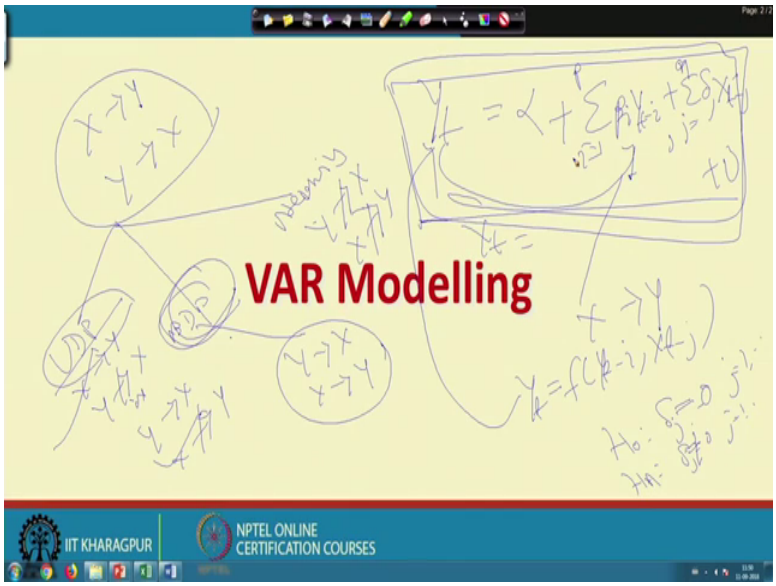
But many occasions many occasions many occasions what do you find actually. So, the Y and X can be connected with the same casualty approach, but with a times series angle. So here, the basic idea is that we like to check whether X is influencing Y and at the same times, Y can influence the concept X. So, technically in the VAR setup, we like to test 4 different hypothesis. For instance, fast hypothesis whether X causing Y and second hypothesis will be let us call it, 1 H 1 0 and this is H 2 0, where Y causing X so; that means, technically we start with the first with issue that X can influence Y and the same times, we can say that Y can influence X.

There is a high possibility the first hypothesis will it true, the second hypothesis not or the other issue will be the second hypothesis will it true. We can validate this second hypothesis, but first hypothesis will not be in the picture and in the third instance both can go simultaneously and the final one is both may not go together. That means, they may not cause each other so; that means, technically we have 3 different inference, the first one there is no such causality between Y and X. Second instance there is casuality but, it is the approach of unidirectional where either X causing Y, while Y does not cause X and in the third case, we can have both Y can influence X and X can influence Y, that is what the technical a technical concept called as bidirectional causality.

So technically, we like to check which one is the correct, ones whether it is actually X causing Y or Y causing X or they cause each other or they do not cause each other, that is where the basic framework of VAR modelling. So, what we used to do? That we like to know the technical issues of VAR modelling, while addressing the causality between 2 variables Y and X, that to check whether X causing Y while Y does not cause X and at the same times. We like to check whether both cause each other or they do not cause each other, that is what the basic background and that to we have a kind of time series frame work.

So be, the beauty of the VAR modelling in a bivariate framework is like this. So, technically it can be set equation like, you know in a such a from; that means, with the help of lag modelling because, this ARIMA cluster arch cluster, GARCH cluster or then VAR cluster, all are started with the kind of lag involvement. So now, we like to check in the framework of VAR modelling, how is this particular linkage? So, the linkage will be like this.

(Refer Slide Time: 07:49)



So, we can start with say Y t is equal to function of, let us say alpha then plus summation beta i Y t minus i, i equal to 1 2 P plus summation delta j X t minus j and j equal to 1 2 q plus the error term..

So, that is what the first model and so that means, we like to check 2 different hypothesis all together, the first one is a whether X causing Y and the counterpart will be Y causing X and this framework will take you into 3 different paths; so the unidirectional path, bidirectional path and then, neutrality path neutrality path ok. So, then the thing is that in the unidirectional path so if X causing Y. So, Y does not cause X sometimes or Y can cause X, but X does not cause Y, that is what the kind of structure and in the same times in the case of bidirectional path, where Y influence X and at the same time X influence Y both will go parallelly.

But in the case of neutrality expectations Y does not cause X and X does not cause Y that is what the kind of gain all together. So now, we like to check which one is the correct;

that means, all cannot go simultaneously, if it is a neutrality then bidirectional is not possible or even a unidirectional is not feasible, if unidirectional feasibles then either this one's or this one's, I have worried both will be cause each other then the; that means, technically we are here. So, that mean in total we may here or we may here or we may here, that is what our the expectation and VAR is a such a technique, which can derive such kind of inference the variable may be any kind of engineering variables that is not big deal, but the structure of this problem is like this.

And to test this we have to design a framework like this and this model is to test whether X causing Y. So, technically this kind of models Y t as a function of lag Y t n lag of another variable X t. So that means, in simple form we can call Y t as a function of Y t minus i and X t minus j, that is what the framework. Now transferring into a particular, explicitly for the model can move into like this, where the null hypothesis is that the delta j should equal should not equal to 0 ok. So, we can start delta j equal to 0 against delta j not equal to 0, where j equal to 1 2 v and the alternative hypothesis will be delta j equal to 0 ok, so you can start delta j equal to 0 against delta j not equal to 0, where j equal to 1 2 so on like this.

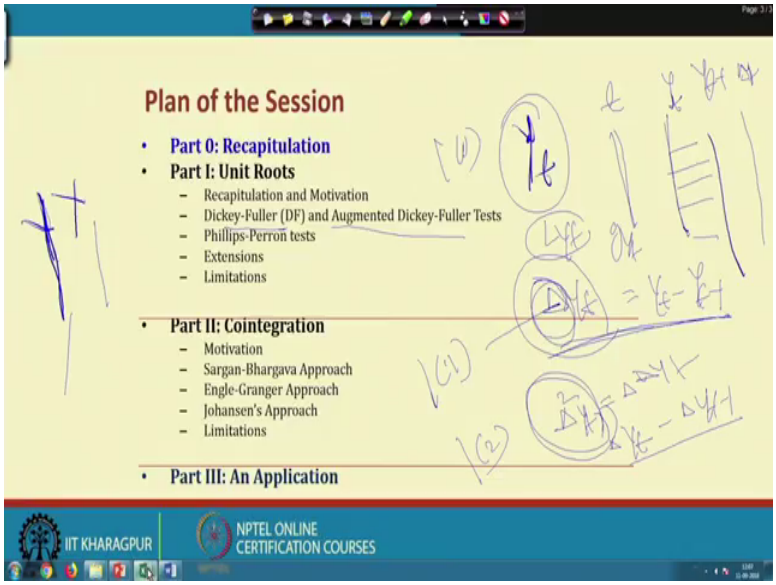
That means technically, we like to check whether these coefficients are coming ok. So, this should be this should be equal to 0 against this not equal to 0, that is what the null hypothesis. So now, in the similar way we can write here another equations, where X t as a function of X t minus i and Y t minus j, if it will write like this, then that brings the causality between Y to X. Now, we like to check whether both are correct, where we are rejecting the delta j for both the models, if that is the case that is what called as bidirectional causality.

If not then either, the first one is true, the second one is rejected and second one is true, first one will be rejected and then finally, both can be rejected that is what the case of neutrality hypothesis. So now, before you come to this VAR setup to bringing or to know the influence of X on Y or Y on X by the framework of lag modelling, we like to check some of the basics, that is what called as stationary check and the kind of cointegration check usually in the times series framework, we have a concept called as short run structure and the long run structure and VAR will bring that particular structure to highlight, whether there is a relationship it in Y and X.

And that to whether in the short run or in the long run. So, to start with this VAR modelling, we have to first highlight the concept called as a unit root and the concept called as co-integration. So, before we analyse the VAR modelling. So, we should know what is actually the unit root concept? And what is the co-integration concept? And how it is link with VAR modelling? So, until unless the unit root and co-integration, you can not specifically use or utilize the VAR model to address the causality between 2 variables with reference to this kind of structure, let us say Y and X. So, let us see how is that case?

So, in such a situation what do you what do you like to do? So, we like to bring the concept of unit root and co-integration.

(Refer Slide Time: 15:10)



So, let us know what is actually unit root and co-integration? And then we will discuss you know the VAR modelling. So somehow, unit root is a slightly torch in the case of in a ARIMA models, where we use the term autoregressive integrated moving average, the term integration there it is influence of unit roots only. So, unit roots literary means to check the stationarity of the variables and stationarity represents the stability of particular variables, usually we like to check or we like to use a particular variable to study the influence, where the variable by default is a stable one.

But in any sense we are not going in deep about this theoretical understanding about the unit root and co-integration. However, we just like to know, what is the concept of unit

root? And how we can actually check the unity root? And that to as per the requirement of VAR modelling; so the test which we can actually do for unit root, that is called as unit root test and the objective of unit root test is to check whether, the variable are stationary or non stationary. If non stationary at what level the variable will be stationary? So that means, here we have to bring various levels, where the variable will be stable; so starting with the original form the first difference form, the second difference form and so on.

For instance, if a variable say Y then it is original form, we declare as Y t. So, it will be like this, it will be called as Y t so; that means, in dataset, we have a t corresponding to Y information. Now we have the term called as a delta Y t, which is nothing but, called as a Y t minus Y t minus 1 that is the difference; so t minus 1 means if I use Y t and this Y t. So, we can call as Y t then by default Y t minus 1 will be just 1 lag behind. So, if this series starts here corresponding to this Y t, then Y t will be just 1 lag behind, this is how the series will be generated then, we can create a component called as a delta Y t by making the difference between the 2.

And then we will get a new variable called as a delta Y t. So, delta Y t equal to Y t minus Y t minus 1. Now in the similar way, we can create delta square Y t, which is equal to delta upon delta Y t; that means, we have already delta Y t here so; that means, technically it is a nothing, but actually delta Y t minus delta Y t minus 1 ok. So, then we will we will simplify you refine Y t minus to Y t minus 1 plus Y t minus 2 like this, we will have a similar kind of structure. So, like this you can increase the order. So, the delta is called as faster first order differential difference operator and that I will like to bring original form, first difference form and second difference form.

This is what is called as variable, in it is original form this is what called as variable in it is first difference form this is called as variables with reference to second difference form, we can also equally like you can write L Y t. So, log of Y t we can call g Y t growth of Y t. So, like there are various ways of addressing the variable that is called as transfer form and we like to check actually what level the variable is bringing actually, it is stationarity.

So, that is how we are checking, the unity root, the standard in a standard form is actually to check Y t delta Y t delta square Y t and so on.

So; that means, the degree of moment. So, if the variable reaches stationarity at it is original form that that is a 2 Y t then by default the order of integration is 0. If the variable reaches stationarity, at the first difference; that means, technical here then the order of integration, we called as I 1 here, it will be I 1 if the variable reaching stationarity at the second difference; that means, here. So, then the declaration is equal to as I 2. So, like this we like to know what is the bracketed term is called as a order.

So, what is the order of integration through which we can declare that variable is a stationarity variable is a stationary or reaching stationarity so; that means, technically the form where the variable is actually the reaching stationarity, that is the form we can use in the VAR setup, that is what that is what the big deal in VAR modelling. So, we have plenty of test that to conduct the unity root starting with Dickey-Fuller, that is called as a D F and then little bit advance that is called as a Augmented Dickey-fullers then, we have plenty of other test like Phillips –Perron, KPSS, Andrew driver test.

So, likewise we have actually series of unit root test and more or less every test will give you similar kind of inference, if you use and if we and declare the variable is having order of integration 1 1, I 1 that is the order of integration and then the variable will be declared as stationarity, where the order of integration is 1, again if the variable reaching stationarity at I 2; that means, the order of integration is a 2. So, this same inference you can also get through other unit root test statistics like, you know Phillips- Perrons, KPSS and Andrew driver like there are so many other test as well. So, in any case we can use at least, you know 1 or 2 test minimum to check, what level the variable is a reaching stationarity?

So this is the first part of the game and in the second part of the game, we like to know whether the variables are co-integrated so; that means, the starting of the VAR model at least with 2 variables like, we start with you know simple linear regression modelling that tools, you know linear regression modelling with respect to Y and X here, the starting is also with respect to 2 variables Y and X compared to simple Y still we can have regression, what is we have discuss in ARIMA cluster and arch cluster and GARCH cluster, but in the VAR set up, we need actually at least 2 variables and then we bring their lags and by default like ARIMA cluster and arch cluster and GARCH cluster, we need to fast optimise the lag length.

So like, we have already highlighted the big model, where Y t as a function of Y t minus j or X t minus j, the lag creation so obviously, you have to first specify, what is j? That is the optimum lag if say j equal to 1, then the model will be restricted to Y t as a function of Y t minus 1 only, if not let us say j equal to 3 then, Y t will be influenced by Y t minus 1, Y t minus 2 and Y t minus 3 so; that means, at a test stress 3 lag can influence the current Y and sometimes it is true for instance, the investment in infrastructure, investment in transportation. So, the fast year transportation will bring transportation output today.

So now the question is a how many fast you have to consider for today's transportation outcome? Mistake any simple in a civil engineering problem. So, the past investment will bring the current performance and in this kind of framework. So, the VAR is the supreme and VAR can do that and detect upto what lag? Or upto what position the first behaviour or pass reaction of the variable will influence the present behaviour?

And for that we need to know the stationarity issue and the co-integration issue. So, we first do the unit root check, since the game starts with 2 variables Y and X. So first thing is you check what is Y order, order of integration and X order of integration? Since, it is a vector autoregressive. So, you can have a more number of variables into the system, but let us understand the bivariate structure then, you can go for a multivariate structure. So, coming to the co-integration you first need to check, what is the unit root position between Y and X? We need to report, whether they are having same order of integration to this a stationarity or different order of integration to this the stationarity that brings the difference and that brings difference in modelling.

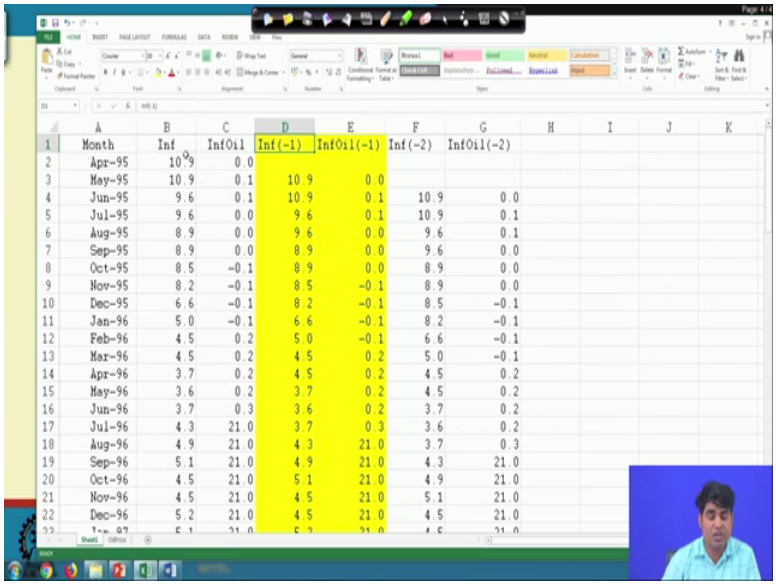
And we need to have that information before you start the process, co-integration simple represents the issue of longer on relationship between the 2 variables, there is a high chance in the short run in, short span of time the variables may not have any kind of linkage relationship, but over the time they can converge each other, they can cause each other that is technically called as co-integration. So, they are in a different position in the short runs, but ultimately they only they only converge each other in the long span of time.

And the technique co-integration we like to like to check whether these 2 variables will be converge each other over the period of time, if that is the case then, we can use say

vector autoregressive VAR is a 2 form for simple vector structure vector auto regressive, if not then there is a counterpart called as vector error correction modelling. So, there will be error correction term, which will be representative to on the long running part and the difference between VAR and VSM will be coming into the picture with reference to co-integration only.

So; that means, it is a mandatory requirement, that you should know what is unit root position? And what is the co-integration position before you start estimating the VAR modelling? And look for the kind of inference. Until unless, the structure of unit root and co-integration. You are not in a position to go estimate the VAR modelling, but we should know, how is this set up, actually in the case of in a unit root?

(Refer Slide Time: 29:40)



	A	B	C	D	E	F	G
1	Month	Inf	InfOil	Inf(-1)	InfOil(-1)	Inf(-2)	InfOil(-2)
2	Apr-95	10.9	0.0				
3	May-95	10.9	0.1	10.9	0.0		
4	Jun-95	9.6	0.1	10.9	0.1	10.9	0.0
5	Jul-95	9.6	0.0	9.6	0.1	10.9	0.1
6	Aug-95	8.9	0.0	9.6	0.0	9.6	0.1
7	Sep-95	8.9	0.0	8.9	0.0	9.6	0.0
8	Oct-95	8.5	-0.1	8.9	0.0	8.9	0.0
9	Nov-95	8.2	-0.1	8.5	-0.1	8.9	0.0
10	Dec-95	6.6	-0.1	8.2	-0.1	8.5	-0.1
11	Jan-96	5.0	-0.1	6.6	-0.1	8.2	-0.1
12	Feb-96	4.5	0.2	5.0	-0.1	6.6	-0.1
13	Mar-96	4.5	0.2	4.5	0.2	5.0	-0.1
14	Apr-96	3.7	0.2	4.5	0.2	4.5	0.2
15	May-96	3.6	0.2	3.7	0.2	4.5	0.2
16	Jun-96	3.7	0.3	3.6	0.2	3.7	0.2
17	Jul-96	4.3	21.0	3.7	0.3	3.6	0.2
18	Aug-96	4.9	21.0	4.3	21.0	3.7	0.3
19	Sep-96	5.1	21.0	4.9	21.0	4.3	21.0
20	Oct-96	4.5	21.0	5.1	21.0	4.9	21.0
21	Nov-96	4.5	21.0	4.5	21.0	5.1	21.0
22	Dec-96	5.2	21.0	4.5	21.0	4.5	21.0

So, let us take an example here, I will take you to the excel spreadsheet. And see here, it is a time series data and we have here data with respect to different time period and that to monthly data and we like to regress inflation with oil price.

And the first component is inflations, which is it is original form we call it inflation and or Inf, where we can call Inf t that is the original form and then we can create actually Inf t minus 1, just 1 lag behind this is the series. So, just copy and one paste, 1 lag behind and that will create a new variable called as Inf t minus 1. Similarly this is actually oil and it is called as ot and then received t minus 1 this is the lag variables, first lag and again corresponding to inflation you paste 2 points behind then you need to create the

kind of no inflation t minus 2 again, similarly oil t minus 2, that is after testing 2 point down corresponding to oil data.

So now so, you can check unity root you know at this test that is called as, first level check and that is the variables in it is original form and if you check this 2 variables at this test, this is another case I will make it you know different colours. So, this is this is the variables and we are checking these variables are there in a first difference ok. So, t minus 1s and then we can bring the difference and again we can check you know with a 2 lag ok. Now what is happening? So now, to bring you know delta Inf t so, this minus this. So, you can create delta Inf t so; that means, the first difference. So, likewise you can actually create delta Y t delta Y t minus 1.

Similarly, oil t oil t minus 1 and delta oil t like this and then you check their unity root separately and again you can do through software only particularly using the Dickey Fuller test and Augmented Dickey Fuller test, the model is very simple one ok, just you connect with you know Y t upon Y t minus 1, like a simple you know, we apply in a kind of mean equation in GARCH model and that that will bring the structure of the kind of, unity root and again, if you apply Augumented Dicky Fuller, some extension has to be added and some of the softwares by default with give you the results of DF, DFPP, Philips Perrons, KPSS like this.

So, just you know use this software and you know allow the software to dictate the order of the integration of a variable, where it reaches stationarity and that is the first hand information, we need to report before we start co-integration and then the VAR moulding. So, this is what the procedure you have to follow and then check the stationarity part first. After checking the stationarity then, you have to start linking the linking these 2 by co-integration framework. So, we should know how co-integration can be apply to explore the long run relationship between the 2 variables or group of variables.

So, first unit root then, co-integration and then VAR. So, that is what the steps you have to follow. So, I have already highlighted the unit root structure of course, I have not discuss the details about in all this test, but somehow I have name this test. So, you can use any test then check the stationarity of the variables and that is the first step of this process, once you do that then, we can move to second step and the third step.

So, in the next lectures, we like to touch upon the second step and third step, where we can know about the co-integration and the VAR structure. With this we will stop here.

Thank you very much.