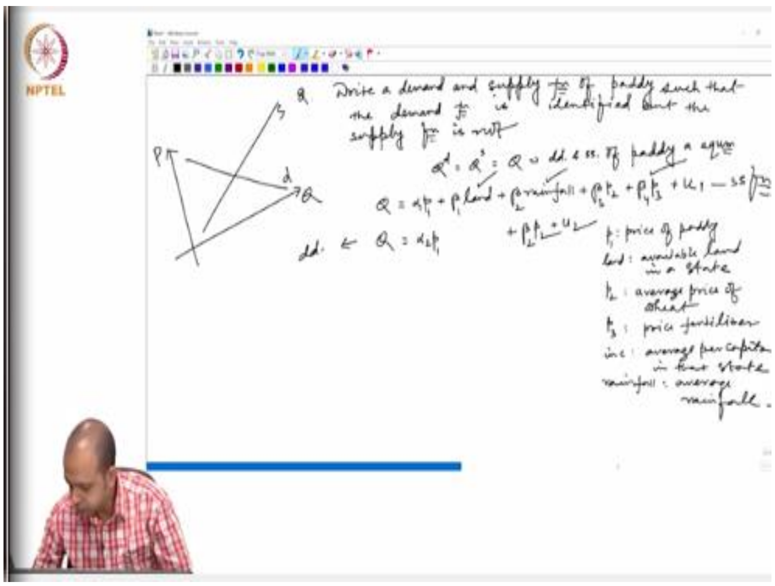
Applied Econometrics
Prof. Sabuj Kumar Mandal
Department of Humanities and Social Sciences
Indian Institute of Technology-Madras

Lecture-20
Simultaneous Equation Model-Part VIII

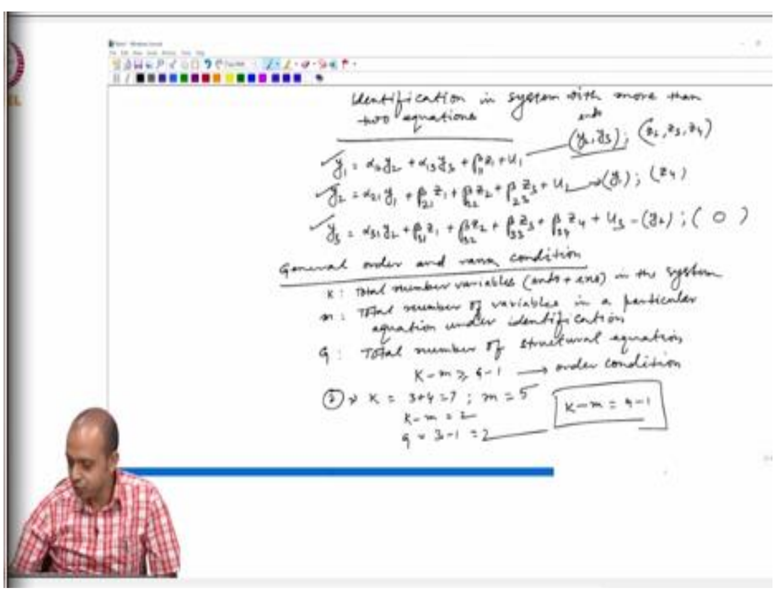
Now what we will do?

(Refer Slide Time: 00:16)



We will try to generalize this system by including more than 2 equations, let us see.

(Refer Slide Time: 00:44)



So, this is let us say identification in systems with more than 2 equations. So far whatever system we discussed in all those systems there are only 2 equations. So, now let us say that this is my equation.

$$Y_1 = \alpha_{12}Y_2 + \alpha_{13}Y_3 + \beta_{11}Z_1 + u_1$$

Here the first subscript indicates the equation number and second subscript is the variable. So, alpha 12 that means that is the first equation I am denoting by 1, 1, 1 always 1.

Similarly,

$$Y_2 = \alpha_{21}Y_1 + \beta_{21}Z_1 + \beta_{22}Z_2 + \beta_{23}Z_3 + u_2$$

$$Y_3 = \alpha_{31}Y_2 + \beta_{31}Z_1 + \beta_{32}Z_2 + \beta_{33}Z_3 + \beta_{34}Z_4 + \quad u_3$$

. So, this is a system with more than 2 equation, we have 3 equations sometimes we have more than 3 equations also. Now when we have 3 or more than 3 equations in a system it is difficult to tell immediately which equation is identified, which equation is not?

So, we need to find out some general condition of identification. First of all if you look at the first equation how many endogenous variables are there? You have 2 endogenous variable y 2 and y 3. And how many variables are excluded? We have z 2, z 3 and z 4, z 2, z 3 and z 4 they are excluded, so we have 2 endogenous variable but 3 excluded exogenous variables, so first equation is identified.

So, that means in this equation y 2 and y 3 they are endogenous variable, these are all endogenous variables. And what are the exogenous variables? Z 2, z 3, z 4, z 1 is also exogenous but that is included in the model, so I am not considering that, so excluded variable are z 2, z 3 and z 4 they are excluded. So, when you have 2 endogenous variable, 3 variables excluded obviously order condition is satisfied.

Now when you come to second equation, what is your endogenous variable only y 1, **y 1**what is your excluded variable? z 4. z 1, z 2, z 3 is there, z 4 is exclude that is also identified. Third equation, y 2 is the endogenous variable and what is your excluded variable? See z 1, z 2, z 3, z 4 everything is there, so excluded variable is 0, there is no variable which is excluded from the third equation, so order condition is not satisfied.

So, you had to think in this way to tell which equation is identified, which equation is not? But here I have 3 equations, I can look at and compare and check which is identified. In a system let us say there are 50 such equations, sometimes you may have to work with 50 equations. And it is very difficult in a system of 50 equations to see a particular equation, what are the endogenous variables, which is excluded, which is included, so many things.

So, a general rule general rank and order condition, so this is general rank or I would say order and rank condition. Let us assume that K is total number of variables endogenous plus exogenous in the system. Let us say that m, m is total number of variables in a particular equation under identification. So, if we are thinking of whether second equation is identified or not, m is basically total number of variables in second equation.

If we are thinking whether the first equation is identified or not, then m is the total number of variable in the first equation like that. And G is total number of structural equation then what is the condition? K - m should be greater than equals to G - 1, this is the general rank order condition, so this is the general order condition. Now what is the meaning of this? When I am saying let us say we are talking about equation 2, let us think about equation 2.

In equation 2 K is the total number of variables in the system itself, so we have y 1, y 2, y 3 these 3 endogenous variable z 1, z 2, z 3, z 4, 4 exogenous variable, so K = 3 + 4 = 7. What is m? For equation 2, that means when you are considering identification per equation 2. In equation 2 how many variables are there? 1, 2, 3, 4, 5. y 1, y 2 endogenous, z 1, z 2, z 3 exogenous, so total 5, so K - m = 2.

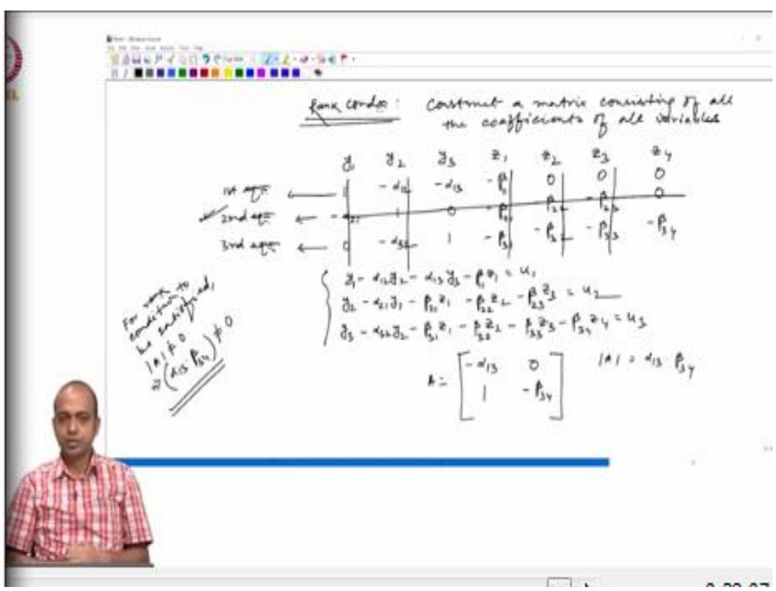
And what is G here? G equals to total number of structural equation 1, 2, 3, so G - 1 = 2. So, that means I can say that K - m = G - 1, so second equation is exactly identified because it is equal. Now if you think about the second equation individually as I said there is 1 endogenous variable and look at 1 excluded variable only, so that is why from this equation I can say that easily second equation is exactly identified.

So, this is the general formulation and same thing is happening when you individually look at the equation also. Basically what this equation says? When I am saying K is the total number of variable and m is the total number of variables appearing in the particular equation that means basically K - m indicates how many variables are actually excluded. So, how many variables are excluded?

And since G is the total number of structural equation how many structural equation you can get in a system? You will get only those many structural equations, as many endogenous variable you have in the system. See here I have 3 endogenous variable y 1, y 2, y 3 for each endogenous variable I have 1 structural equation, for y 1 I have 1 structural equation, for y 2 I have 1 structural equation, for y 3 I have 1 structural equation.

That is why G - 1 is basically a number of endogenous variable minus 1. So, that is how it is making sense, greater than or equals to. So, that means G - 1 would be then the total number of endogenous variable minus 1. And this side you have excluded exogenous variable, so this is how the general order condition is satisfied. But the rank condition is little involved that we need to do it separately.

(Refer Slide Time: 13:21)



So, now we will discuss about rank condition. For rank condition what we will do? We will construct a matrix. So, construct a matrix consisting of all the coefficients of all variables. So, if you think about the first equation, so what I will do? First I will write all the variables, so what is my variable y 1, y 2, y 3, z 1, z 2, z 3 and z 4. Now think about the first equation, in the first equation I have y 1, so coefficient is 1.

Then second equation is - $\boldsymbol{\alpha_{12}}$ that is the coefficient of y 2 and then coefficient of y 3 is - $\boldsymbol{\alpha_{13}}$ then - $\boldsymbol{\beta_{11}}$ and then we have no z 2, no z 3, no z 4. Then for the second equation I have - $\boldsymbol{\alpha_{11}}$ alpha 21 then 1, then I have no y 3, then - beta 21, - beta 22, - beta 2 3 then I do not have any set for there. Then for the third equation I have no y 1, I have only -alpha 32, then I have y 3, then - beta 31, - beta 32, - beta 33, then - beta 34.

So, this is for the first equation, this is for the second equation, this is for the third equation. And how I am writing this equation? Simply I am doing for example I am for the first equation how I am rewriting y 1, - alpha 12, y 2, - alpha 13, y 3, - beta 11, z 1 = y u 1. So, all the equations are written in this format. Similarly y 2 - alpha 21 y 1 - beta 21 z 1, - beta 22 z 2, - beta 23 z 3 = u 2. For the third equation what I am writing y 3 - alpha 32 y 2 - beta 31 and z 1 - beta 32 z 2 - beta 33 z 3 - beta 34 z 4 = u 3.

So, I have converted the equation in this format to construct this matrix. Now what we will do? For rank condition, let us say that I am considering second equation whether the second equation is identified by rank condition is not. If you are considering second equation then you just cross the second equation like this. And then you cross all the columns in which you have non zero coefficients.

So, you cross out all the columns in which second equation have non-zero coefficient. So, first equation, yes, I have non-zero coefficient alpha 21, so this is I will cross out, yes, this second column I have non-zero coefficient 1, so I will cross out. Now here I do not have anything, this is 0, so I will not cross out this, this I will cross out, this I will cross out and this I will cross out but this I will not cross out because this is 0, is this clear?

Second equation I will cross out then I will cross out columns in which the second equation is having non zero coefficient. This is crossed out because non zero coefficient, this is not crossed out because this is 0, this is crossed out, this is crossed out, this is crossed out, this is not because again 0. Now after this we have to see what are the elements left? So, I will have a matrix of alpha 13; see this is not crossed out.

Then I have 0 here, then I have 1 here and I have -beta 34 here, this is the matrix what I got after crossing let us say this is A. So, -alpha 13 0 1, -beta 3, so determinants of these is alpha 13 into beta 34. So, for rank condition to be satisfied what we need? Determinants of a should be not equals to 0 which implies alpha 13 into beta 34 should be not equals to 0. So, that is the general rank condition which is easy to identify, is this clear?

So, that means what we are doing? By crossing out these rows and columns essentially what we are trying to get what are the variables excluded from the model. See I have crossed out this column, so that means some variable is here, I have crossed out this column because I am excluding. Because this is already z 1 is included, z 2 is included, z 3 is included but z 4 is not included, that is why I am not crossing out this.

If you look at the second equation z 1, z 2, z 3 they are actually included. So, in the mechanism the way I am crossing out this column I crossed out because y 2 is appearing, this column I am not crossing out, this is again crossing out because z 1 is there, this is again crossing out z 2 is there, this is again crossing out z 3 is there, I am not crossing out this. So, these 2 columns I am not crossing out just to get the variables which are not appearing in that equation.

And then I am getting this matrix and that determinants of that matrix would have non-zero. Then essentially I am getting this alpha 13 multiplied by beta 34 should be not equals to 0, that is the general rank condition. And general order condition I have already decided K - 1 m less than equals to G - 1. So, with this we are closing our entire discussion on simultaneous equation model.

We, now know what is a simultaneous equation model, what is not a simultaneous equation model. From the beginning again if you recall 2 variables must be jointly determined but that is only necessary condition, sufficient condition is that the equation should have their standalone meaning. That means in a system of equation every equation should be a structural equation, only the structural equation will have standalone meaning.

Otherwise we saw the example of saving and housing expenditure even though they are jointly determined they do not form a simultaneous equation model. Then we discussed about a rank and order condition of identification, why identification is required prior to estimation? Because without identification you do not know which equation you are estimating. As an example you can think about the demand and supply equation, if no other variables are included both demand and supply is a function of price only.

So, I do not know whether I am estimating a demand function or supply equation, it does not make any sense. That is why if you want to estimate a demand function you need to construct of your own the system of equation with availability of data of course, such that the demand function is identified. That means you should specify some variables in the supply function which should not appear in the demand function.

And that variable which will appear only in the supply function will act as an observed supply shifter and when the supply curve will shift leftward or rightward, upward or downward I can easily identify or trace out the demand function from each and every intersection. But the rank condition says, rank condition demands something more than the order, order condition says that the variable should be only excluded from the equation.

But the rank condition says the variable should not only be excluded but also have non zero population coefficient. That means it should be significant in the other equation otherwise we cannot say that this is identified. And then lastly we discussed about a system of equation with 3 or more than 3 equations, the general rank and order condition with this matrix algebra. So, with this we are closing our entire discussion of this simultaneous equation model. Thank you.