

**Econometric Modelling**  
**Prof. Rudra P. Pradhan**  
**Department of Management**  
**Indian Institute of Technology, Kharagpur**

**Lecture No. # 18**  
**Matrix Approach to Econometric Modelling**

Good afternoon, this is Dr. Pradhan here. Welcome to NPTEL project on econometric modelling. So, today we will discuss matrix approach to econometric modelling. In the last couple of lectures, we have discussed the fundamental issues of econometric modelling and we have discussed econometric modelling with respect to bivariate systems, trivariate systems and multivariate systems, and we have also discussed various problems and assumption, which we like to use in the case of bivariate system, trivariate system and multivariate system. So, up to till now, we have discussed the fundamental issues of econometric modelling, which is one part of the **one part of the** econometric modelling.

In the next part of the econometric modelling, we have to discuss various issues and problems with respect to the econometric modelling, whether the problem is in a bivariate setup or in a trivariate setup or in a multivariate setup. The moment one you have the estimated econometric modelling then or you can say once we have the estimated econometric model then obviously next step is to check for best fitted models whether the estimated model will be considered as the best model or not. So, the way we will check the existing estimated model to be the best then there are certain steps and there are certain formalities we have to maintain or we have to go through all these formalities.

This **this** requires enough knowledge, enough structures, various techniques through which we have to make a judgment that the existing econometric model **model** will be the best and it can be used for forecasting and policy use. So, before we proceed to the second part of the models econometric model, where we have to discuss various issues. So, we like to highlight the same issue, which we have discussed earlier that to bivariate approach trivariate approach and multivariate approach. So, here our idea is to apply

matrix and to simplify the existing system, which we have discussed earliest. So that means, whether the problem is bivariate setup or trivariate setup or multivariate setup, we can **we can** start with matrix and also we can solve with matrix. So, up to till now we have discussed basic steps, then in between we used matrix little bit, so that the problem will be very simple one and through that simplicity we will get the estimated model.

Today we will discuss very beginning of this econometric systems and we will have the estimated model by the entire first to last we have to see how matrix can help, so that you can say we will get the estimated models without you can say with you can say less effort or you can say without any complexity.

(Refer Slide Time: 04:02)

MATRIX APPROACH TO ESM

$$Y_j = \beta_0 + \sum_{j=1}^n \beta_j X_{ji} + U_j$$

$$Y = [Y_1, Y_2, \dots, Y_n]$$

$$X_1 = [X_{11}, X_{21}, \dots, X_{n1}]$$

$$X_2 = [X_{21}, X_{22}, X_{23}, \dots, X_{2n}]$$

$$X_3 = [X_{31}, X_{32}, X_{33}, \dots, X_{3n}]$$

$$\vdots$$

$$X_n = [X_{n1}, X_{n2}, \dots, X_{nn}]$$

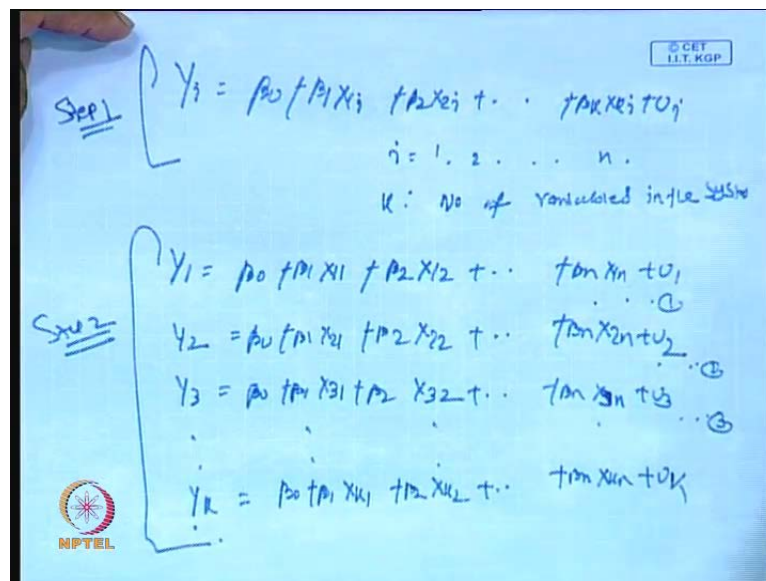
Let me start or let me highlight how is this particular setup through which you have to proceed. So, we have discussed in the last class couple of back couple of lectures back, so that **Y equal to beta 0 equal to** Y equal to beta 0 plus summation beta j X j i plus U j. So, this is the initial setup i equal to 1 to n. So, this is how the model is all about we can set here let us say Y j here.

Now the basic formality here is that. So, we **we** will assume that Y structure is like this Y 1 Y 2 up to Y n and X 1 consist of; that means this will be X 1 X 2 like this, so X 1 is equal to X 1 1, X 2 1, X 3 1 up to X n 1. So similarly, X 2 equal to X 2 1, X 2 2 and X 2 3 then it will go to X 2 n. Then similarly, X 3 is equal to X 3 1, X 3 2 then X 1 1, X 2 1,

$X_{31}, X_{11}$ , so then it should be  $X_{13}$  it should be  $X_{31}, X_{32}, X_{33}$  up to  $X_{3n}$ . So, this is how you have to proceed.

Similarly, we have to find out  $X_n, X_{n1}, X_{n2}$  up to  $X_n$ . So, this is how the entire structure is all about. So, by the way we have to transfer the entire system into matrix format. So, how we will do that?

(Refer Slide Time: 06:13)



Let us start with a simple model here. So, like  $Y_i$  equal to  $\beta_0$  plus  $\beta_1 X_{i1}$  plus  $\beta_2 X_{i2}$  plus  $\beta_k X_{ik}$  plus  $U_i$ . This is the multivariate model with  $k$  number of variables and  $i$  is sample observation. So, here  $i$  equal to 1 2 up to  $n$  that is sample observations and  $k$  is the number of variables in the systems that way that to independent variable, so because  $Y$  is already 1.

How we will write this particular structure? If this is a actually this is in a implicitly format, so we will try to put it in explicit format, so that all the equations can be very it can be very structure and it can be easily understood. Let us see here, so  $Y_1$  equal to  $\beta_0$  plus  $\beta_1 X_{11}$  plus  $\beta_2 X_{12}$  plus  $\beta_n X_{1n}$  plus  $U_1$ . So, this is first equation. So, let it be call it 1. Then  $Y_2$  equal to  $\beta_0$  plus  $\beta_1 X_{21}$  plus  $\beta_2 X_{22}$  plus  $\beta_n X_{2n}$  plus  $U_2$ .

Similarly,  $Y_3$  equal to  $\beta_0$  plus  $\beta_1 X_{31}$  then plus  $\beta_2 X_{32}$  then plus  $\beta_n X_{3n}$  plus  $U_3$ . This is we will call it 2, this we will call it 3, it will

continue like this. So, then finally we will get  $Y_k$ , because  $k$  number of variables are there,  $Y_k$  equal to  $\beta_0$  plus  $\beta_1 X_{k1}$  plus  $\beta_2 X_{k2}$  plus  $\beta_n X_{kn}$  plus  $U_k$  because it is in  $k$ -th series. So, this is how the structure is all about.

Let us assume that this is the step 1 of this matrix approach and this is step 2 of this matrix approach. Should I start once again? So, the step 1 process is to highlight the multivariate setup. So, where  $Y_i$  equal to  $\beta_0$  plus  $\beta_1 X_{i1}$  plus  $\beta_2 X_{i2}$  plus  $\beta_k X_{ik}$  plus  $U_i$ , where  $i$  represents number of sample observation in the systems and  $k$  represents number of independent variables in the system and  $Y$  is the dependent variable.

This is the basic framework of multivariate model. So, once you will expand that particular model, then we have series of equation like  $Y_1$  equal to  $\beta_0$  plus  $\beta_1 X_{11}$  plus  $\beta_2 X_{12}$  up to  $\beta_n X_{1n}$  plus  $U_1$ . So, similarly,  $Y_2$  equal to  $\beta_0$  plus  $\beta_1 X_{21}$  plus  $\beta_2 X_{22}$  up to  $\beta_n X_{2n}$  plus  $U_2$ . So, it will continue. So, in the  $k$ -th series we will get  $Y_k$  equal to  $\beta_0$  plus  $\beta_1 X_{k1}$  so, up to  $\beta_n X_{kn}$  plus  $U_k$ . So, this is how the entire system all about. So, what you have to do? We will transpose this particular system in to a matrix approach.

(Refer Slide Time: 10:27)

© CET  
I.I.T. KGP

$$Y = X\beta + U$$

where,

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_k \end{bmatrix}_{k \times 1} \quad X = \begin{bmatrix} 1 & X_{11} & X_{12} & \dots & X_{1n} \\ 2 & X_{21} & X_{22} & \dots & X_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ k & X_{k1} & X_{k2} & \dots & X_{kn} \end{bmatrix}_{k \times n}$$

$$\beta = [\beta_0, \beta_1, \dots, \beta_n]_{n \times 1}$$

$$U = [U_1, \dots, U_k]_{k \times 1}$$

$$Y = \frac{k \times 1}{k \times 1} = \frac{(k \times n)(n \times 1)}{k \times 1} + k \times 1$$

NPTEL

Now, I will write simply here I will write  $Y$  equal to  $X\beta$  plus  $U$ . So, the entire system whatever we have discussed in the step 2, so I can represent or I can immediately transfer into this particular format so that means, instead of writing so many equations, so many variables and so many parameters at a time. So, it

is better you put in one equation and in a compact format, so that in a very small space and very small way you can represent the entire issues.

Let say here  $Y$  equal to  $X\beta + U$ . So, how we have to represent? Where,  $Y$  equal to the column vectors so,  $Y_1, Y_2$  up to  $Y_k$ . So, this is how you can say  $Y$  series all about. That means this particular format here so  $Y$  equal to  $X\beta + U$  is nothing but in a vector format. So that means, here we are writing  $Y$  and here we are writing  $X\beta$  and  $U$ . So that means,  $Y$  has a series of component,  $X$  has a series of component. So,  $\beta$  has a series of component and  $U$  has a series of component.

Let us first put it in a particular structure then we will expand it how we will go for this estimation? So now, the moment we will transpose the entire step 2 equation into step 3, where  $Y$  equal to  $X\beta + U$ , so where  $Y$  represents  $Y_1, Y_2, Y_k$ , so this is one matrix. So  $y$  is the variable, which consists of  $Y_1, Y_2$  up to you can say  $Y_k$ . So, that means, it is the order of  $k$  into or  $k$  cross  $1$ .

Similarly,  $X$  represents here  $X_{11}, X_{12}, X_{1n}$  up to  $X_{k1}, X_{k2}, X_{kn}$ . So, this is  $1$ , then  $X_{11}, X_{21}, X_{31}$  then it will come to  $X_{k1}$ . So similarly,  $X_{12}, X_{22}, X_{32}$  then it will continue  $X_{k2}$ , so  $X_{1n}, X_{2n}, X_{3n}$  then it will continue then finally  $X_{1n}, X_{2n}$  then we will have  $X_{kn}$ . So, this is order of  $k$  into  $n$ . So that means  $k$  number of variables with  $n$  observation.

That means the matrix  $X$  means it is the matrix of order  $k$  cross  $n$ , so  $k$  number of rows and  $n$  number of columns. So that means, we have to represent again  $\beta$ , so  $\beta$  is equal to here, so like  $\beta_0, \beta_1$  up to  $\beta_n$ . So, this is one particular format. So, then finally,  $U$  represents  $U_1, U_2$  up to  $U_k$ . So, for  $\beta$  the order is  $n$  cross  $1$ , so  $n$  number of rows into  $1$  column. So, this is also  $k$  cross one. So, it is  $k$  number of rows into  $1$  column.

So now, in the left side, the moment I will write  $Y$  equal to  $X\beta + U$ , so then obviously,  $Y$  here is  $k$  cross  $1$ , so that means, the entire structure  $X\beta + U$  should be  $k$  cross  $1$ . So that means, here  $X$  represents  $k$  into  $n$ , so then  $\beta$  represents  $n$  cross  $1$ , so then plus  $U$  represents  $k$  cross  $1$ . So, if you will target then by matrix structure this becomes automatically  $k$  cross  $1$ . So that means,  $X$  is a vector,  $\beta$  is vector. So, when  $X$  is order  $k$  cross  $n$  and  $b$  has a order  $n$  cross  $1$ , so obviously, in the first case right side and the second case left side should be equal then

obviously, this will be cross automatically. So obviously, the entire multiplication will be order  $k \times 1$  only. So, this is  $k \times 1$  and this is  $k \times 1$ . So, this is **this is** systematic means efficient one, the reason is that you have to apply this addition property, because this is order same **this is order same**. So, addition can be possible.

(Refer Slide Time: 15:47)

The image shows a handwritten derivation on a blue background. The equations are as follows:

$$Y = X\beta + U \quad \dots \textcircled{1}$$

$$Y^{\wedge} = X\beta^{\wedge} \quad \dots \textcircled{2}$$

$$e = Y - Y^{\wedge} = Y - X\beta^{\wedge}$$

$$e = Y - X\beta^{\wedge}$$

$$e' = (Y - X\beta^{\wedge})'$$

$$\sum e^2 = e'e$$

$$= (Y - X\beta^{\wedge})' (Y - X\beta^{\wedge})$$

$$e' = [e_1 \ e_2 \ \dots \ e_n]$$

$$\sum e^2 = [Y' - (X\beta^{\wedge})'] (Y - X\beta^{\wedge})$$

There is a diagram showing a vector  $e$  with components  $e_1, e_2, \dots, e_n$  and a matrix  $J$  with rows  $e_1, e_2, \dots, e_n$ . A bracket groups the  $e_i$  terms in the matrix  $J$ .

Now we have, so I will write it here once again, so  $Y$  equal to  $X\beta + U$ , so this **this** will put it an equation number. So that means, when we will go for matrix approach, then our starting process is from  $Y$  equal to  $X\beta + U$ , because whatever we have discussed before that then these are all basic information to come to that particular equation  $Y$  equal to  $X\beta + U$  until unless you have a complete information about the setup then you cannot understand this particular equation. So,  $Y$  equal to  **$X\beta + U$**  means it is derived from this particular setup, so where the system is a  $k$  number of variables that too independent variable and one dependent variable  $Y$  and  $n$  number of observations are there. So,  $n$  number of observations,  $k$  number of independent variable and one dependent variable that is  $Y$ . So that means, with one dependent variable and with  $k$  number of independent variable and sample size  $n$ . So, we have to build a model estimated model through the application of matrix that is the agenda, which we discussed today.

Now  $Y$  equal to  $X\beta + U$  is the starting point of this matrix approach. So, what I have to do, let us assume that my estimated model will be  $Y^{\wedge}$  equal to  $X\beta^{\wedge} + U^{\wedge}$

$U = X\hat{\beta} + U$ . So, now what I have to do here is  $X\hat{\beta} + U$ , so by the way, how do I get an error component? So, next step is we have to get the error component  $e$  equal to  $Y - \hat{Y}$ . So, this is the error component.

Now, as far as our earlier discussion with respect to bivariate, trivariate and multivariate, the moment we will get  $e$  or that is error component then obviously, what we have to do? We usually make it sum squares, then we have to minimize with respect to corresponding parameters. If it is bivariate, then we have to minimize with respect to  $\beta_0$  and  $\beta_1$ , because in the bivariate there are two parameters and in the case of trivariate, we have 3 parameters  $\beta_0, \beta_1, \beta_2$ . Obviously, we have to minimize the error sum square with respect to  $\hat{\beta}_0, \hat{\beta}_1$  and  $\beta_2$ . Similarly, if you will add one after another then obviously the equation will start adding, so obviously the complexity will start increasing.

Now, we will start with some different way; that means, we have to assume that we are handling the entire structure at a time. So,  $e = Y - \hat{Y}$ . So, that time we are making square and taking the summation. So here, there is no concept of making squares and taking the summation, here we have to apply matrix multiplication the 2 application of transpose. So, now  $e = Y - \hat{Y}$ , so that means, what is this, so it is equal to  $Y - \hat{Y}$ . What is  $\hat{Y}$ ? So,  $\hat{Y} = X\hat{\beta} + U$ , so  $\hat{Y} = X\hat{\beta}$  there is no question of  $U$  error component will be automatically removed. So that means  $\hat{Y} = X\hat{\beta}$  only. So, this is equation number 2. So that means, when we subtract  $Y - \hat{Y}$  we will get error component, so this is just subtraction. So, original equations with this particular estimated equation so that means, it is nothing but  $Y - X\hat{\beta}$ . So, this is  $Y - X\hat{\beta}$ . So, I will write it here. So,  $e = Y - X\hat{\beta}$ .

Now what is  $e^T$  we will call it  $Y - X\hat{\beta}$  into its transpose. So, before handling this particular problem you must have sufficient information, so sufficient knowledge about this matrix. Matrix is a simply there are many ways you have to represent the matrix. The simplest definition of matrix is to arrangement of elements in row wise and column wise, but that is in order to the case here, so we are not arranging the elements row wise and column wise.

In the meantime, we are making systematic approach. That means, here we are applying the application part of matrix, so how this matrix can be utilized properly to make the complex problem into simplex? So, that is the basic agenda here. So, this is how we have calculated  $e$  equal to  $Y$  minus  $X$  beta and  $e$  transpose equal to  $Y$  minus beta hat to the power transpose.

The moment you will get this one, then we **we** like to get summation  $e$  square, so summation  $e$  square so that is error sum square is equal to  $e$  transpose  $e$  only **e transpose e only**. So, this is how summation  $e$  square, so that means, what is  $e$  transpose  $e$ ? So, let us say  $e$  equal to  $e_1 e_2$  up to  $e_n$ , so  $e$  transpose  $e$  means, so this structure will be like this, so  $e_1 e_2$  up to  $e_n$  this is 1 column, so then another column is  $U_1 e_2$  up to  $e_n$ . So that means, the multiplication will be like this  $e_1 e_1 e_2 e_2$  then  $e_n e_n$  the cross product will be automatically 0, because this is our standard assumption. So, as a result we will get  $e$  transpose  $e$ . So,  $e$  transpose  $e$  is nothing but  $Y$  minus  $X$  beta hat to the power transpose into  $Y$  minus  $X$  beta **X beta** hat. So, this is our entire structure. So, this summation  $e$  square means  $Y$  minus  $X$  beta hat into  $Y$  minus  $X$  beta. So that means, summation  $e$  square is equal to  $Y$  **Y** transpose minus  $X$  beta hat to the power transpose into **into**  $Y$  minus  $X$  beta **X beta** hat, so this is how the entire structure is all about.

(Refer Slide Time: 22:23)

$$\begin{aligned}
 SSE &= (Y' - \beta' X') (Y - X \beta) \\
 &= Y'Y - Y'X \beta - \beta' X'Y + \beta' X'X \beta \\
 &= Y'Y - 2\beta' X'Y + \beta' X'X \beta
 \end{aligned}$$

$$\frac{\partial SSE}{\partial \beta} = -2X'Y + 2X'X \beta = 0$$

$Y'X \beta = \beta' X'Y$

$$\frac{\partial (X'X)}{\partial X} = 2X = 2X'$$

$$\beta = (X'X)^{-1} X'Y$$

$$(X'X)^{-1} X'Y = (X'X)^{-1} (X'X) \beta$$

$$(X'X)^{-1} X'Y = I_n \beta$$

Now we have to **we have to** you can say we have to diversify it, so what you have to do? So, summation  $e$  square equal to  $Y$  transpose minus beta hat transpose  $X$  transpose into



$Y - X\beta$ , so this is how, because as per the property of transpose  $X\beta$  hat transpose means  $\beta$  hat transpose  $X$  transpose. So, this is how the property all about. So, I am not going to teach all these details about matrix, just if that you have to learn yourself.

Here I have to give the only structure, how matrix can be used to estimate this econometric model that to multivariate analysis? So, accordingly we have to proceed. So that means, I am assuming that you must have some information about matrix until and unless you have complete information about matrix it is very difficult to touch this one or you cannot handle this particular component. So obviously, you have to first learn the basic of matrix then you have to apply. Means by any chance if you have no matrix knowledge then you have to go by the simple structure of the econometric modelling, which we have discussed in the last class. So, if you like to make it is very attractive and you like to save time or you can say make it very **(( ))** then obviously, you have to learn matrix and you have to apply the matrix to get the system done in a proper way or systematic way. So, this is how the summation  $e^2$ .

Now what you have to do? So, this is nothing but, we have to go for multiplication, so  $Y$  transpose  $Y - X\beta$  hat transpose  $X\beta$  hat minus  $\beta$  hat transpose  $X$  transpose  $Y$  plus  $\beta$  hat transpose  $X$  transpose  $X\beta$  hat, so this is how the entire multiplication all about. So, now **now** we have to simplify this one, so now this is nothing but  $Y$  transpose  $Y - X\beta$  hat transpose  $X\beta$  hat minus  $\beta$  hat transpose  $X$  transpose  $Y$  plus  $\beta$  hat transpose  $X$  transpose  $X\beta$  hat. So that means, this particular component this and this will be equal because it is symmetric in nature. So that means,  $Y$  transpose  $X\beta$  hat is equal to  $\beta$  hat transpose  $X$  transpose  $Y$ . If we will apply this particular structures then obviously the equation will be reduced to this much.

Now  $\beta$  hat transpose is our requirement, so we have to minimize with respect to  $\beta$  hat, but remember here  $\beta$  hat does not mean only one  $\beta$  **beta** hat indicates so many  $\beta$ s are there, so that means, it is **it is** this set of  $\beta_0 \beta_1 \beta_2$  up to  $\beta_k$ . So, there are series of  $\beta$ s are there, but we are putting  $\beta$ . So, here  $\beta$  is not an individual component, rather it is a vector. So  $d$  summation  $e^2$  by  $d\beta$  hat, which is nothing but **nothing but** which is coming like this, so the structure will be like this. So, I can directly write here  $2 X$  transpose  $Y$   **$2 X$  transpose  $X$**  plus  $2 X$  transpose  $X\beta$  hat,

because **because** there is a property of matrix is that  $d(X^T A X) = d(X^T) A X + X^T d(A) X + X^T A d(X)$  is equal to  $2 A X$  or  $2 X^T A$ .

This is how by applying this particular component if you differentiate this particular structure then obviously we will get this particular equation. So that means, this should be equal to 0. If we make equal to 0 then obviously if you will simplify then  $X^T A X = X^T A X$ . So now,  $X^T A X = X^T A X$ . So, we need what is the value of  $\beta$ ? So  $\beta$  had so that means what you have to do? We have to multiply  $X^T X^{-1}$  in both the sides, because we need  $\beta$ . So, if we will go to the right side then obviously  $X^T X$  is the extra component.

Now if we will multiply  $X^T X^{-1}$  **inverse** then obviously, we can have only  $\beta$  in the right side, because just like if  $A$  is a matrix then by properties of inverse matrix,  $A^{-1} A = I$  so obviously, we have to apply that particular properties of inverse matrix then we have to find the solutions.

Now if we will multiply  $X^T X^{-1}$  on the both the sides then that will be like this  $X^T X^{-1} X^T A X = X^T X^{-1} X^T A X$ . So now, this and this is unit, so if I will take this is  $A$  then this is  $A^{-1}$ , so that will be unit matrix, so that means this is called as unit matrix of order  $n$   $\beta$ . So that means, it is equal to  $X^T X^{-1} X^T A X$ , so that means, what we will conclude here **what we will conclude here**, so  $\beta = X^T X^{-1} X^T A X$ . This is the conclusion we have or this is the result we have received from this particular analysis.

(Refer Slide Time: 28:30)

$$Y = X\beta + U$$

$$\hat{Y} = X\hat{\beta}$$

$$e = Y - X\hat{\beta} \quad e' = (Y - X\hat{\beta})'$$

$$\hat{\beta} = (X'X)^{-1} X'y$$

$$(X'X)^{-1} = \begin{bmatrix} x_1 & x_2 & \dots & x_n \\ x_1 \begin{bmatrix} \Sigma x_1^2 & \Sigma x_1 x_2 & \Sigma x_1 x_n \end{bmatrix} \\ x_2 \begin{bmatrix} \Sigma x_2 x_1 & \Sigma x_2^2 & \Sigma x_2 x_n \end{bmatrix} \\ \vdots \\ x_n \begin{bmatrix} \Sigma x_n x_1 & \Sigma x_n x_2 & \Sigma x_n^2 \end{bmatrix} \end{bmatrix}$$

$$X'y = \begin{bmatrix} \Sigma y \\ \Sigma x_1 y \\ \Sigma x_2 y \\ \vdots \\ \Sigma x_n y \end{bmatrix}$$

That means **that means** our structure is like this. Our structure is here our starting point is  $Y = X\beta + U$  then by the way we **we** have  $\hat{Y} = X\hat{\beta}$  then by the way we have  $e = Y - X\hat{\beta}$  and by the way we have  $e' = (Y - X\hat{\beta})'$  equal to  $Y - X\hat{\beta}$  to the power transpose then we have the **we have the** multiplication  $e'$  into  $e$  then obviously, we have to differentiate this sum squares with respect to  $\beta$  then we will get this equation.

Now by the process we have received by the simplicity we have received  $\hat{\beta}$  equal to  $(X'X)^{-1} X'y$ . So that means, what is  $(X'X)^{-1}$ ? Let us see what is that  $(X'X)^{-1}$ ? So that means,  $(X'X)^{-1}$  means we have to apply the inverse matrix formula then you have to do the job accordingly. So,  $(X'X)$  is basically like this. So, let us say  $X_1, X_2$  up to you can say  $X_n$  this side and this side  $X_1, X_2$  up to  $X_n$ . So, then we have to go for this, this into this, this into this, this into this then this into this, this into this and this into this like this. So that means, this particular item will be summation  $X_1^2$ , this particular item is equal to summation  $X_1 X_2$ , this item equal to summation  $X_1 X_n$ .

Similarly,  $X_2$  into  $X_1$  then summation  $X_2 X_2$  square then  $X_2 X_n$ . So similarly,  $X_n$  into  $X_1$  summation  $X_n X_2$  and summation  $X_n$  into  $X_n$  that is  $X_n$  square. So that means,  $(X'X)^{-1}$  means only this much. So, this is the  $(X'X)^{-1}$  matrix. Similarly, we get to know what is **what is** the  $X'y$ ? So,  $X'y$

that means, we can write it here also  $X'Y$ . So that means, we have to multiply the  $X'$  vectors then  $Y$  vectors the **the** product will give you  $X'Y$ , which is nothing but summation  $\sum Y$ , then summation  $\sum X_1 Y$ , then summation  $\sum X_2 Y$ , then continue equal to sum summation  $\sum X_k Y$ . So, this is how the entire structure. I will write it once again here.

(Refer Slide Time: 31:17)

$(X'X)^{-1} = \begin{bmatrix} \Sigma X_1^2 & \Sigma X_1 X_2 & \dots & \Sigma X_1 X_n \\ \Sigma X_2 X_1 & \Sigma X_2^2 & & \Sigma X_2 X_n \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma X_n X_1 & \Sigma X_n X_2 & & \Sigma X_n^2 \end{bmatrix}$   
 $\hat{\beta} = (X'X)^{-1} X'Y$   
 $Var(\hat{\beta}) = \sigma^2 (X'X)^{-1}$   
 $\sigma^2 = \frac{SSE}{(n - k)}$   
 $SSE = SST - SSR$   
 $\Sigma y^2 = \Sigma y_n^2 + \Sigma \epsilon^2$

That means, here  $X'X$  inverse means  **$X'$  transpose  $X$  inverse means  $X'$  transpose  $X$  inverse means** this is summation  $\sum X_1^2$  summation  $\sum X_1 X_2$ , then summation  $\sum X_1 X_n$ , then summation  $\sum X_2 X_1$ , summation  $\sum X_2^2$ , summation  $\sum X_2 X_n$ , then continue then summation  $\sum X_n X_1$ , summation  $\sum X_n X_2$ , then summation  $\sum X_n^2$ . So, this is  **$X'X$**  transpose  $X$  inverse and  $X'$  transpose  $Y$  is equal to summation  $\sum Y$ , summation  $\sum X_1 Y$ , then summation  $\sum X_2 Y$ , then summation  $\sum X_k Y$ , so this is how the entire systems.

Why **why** we are writing all these things? because we have the structure means by the way when you have  $X\hat{\beta}$ , then  $Y$  hat equal to  $X\hat{\beta}$ , then  $\hat{\beta}$  equal to  $(X'X)^{-1} X'Y$ , so that means, if we will multiply this and this then obviously, you will get this particular product. This **this** is nothing but this is  $n$ -th series  $n$  into  $k$  and this is  $k$  into  $1$ , so you will get  $n$  into  $1$  matrix. So that means,  $Y_1 Y_2$  up to  $Y_n$  that is the  $n$ -th observation.

Now once you have the estimated means if will go by as usual procedure then first **first** objective is to set the equation then you transfer into mathematical form of the model, so

that means, you set the theory then you transfer the theory into mathematical form of the model, then you have to transfer into statistical form of the model, then you have to apply some technique to get the estimated model. The moment you will get the estimated model then obviously next objective is to go for checking the fitness of the model. So that means, as usual we have two different structures that to test the significance of the parameter and to test the overall fitness of the model.

Now to check the significance of the parameters we have to means we need **we need** to calculate variance of all these parameters, then standard error of all the parameters, then followed by t statistics and then you have to compare with the tabulated accordingly we have to look, whether it is significant and if it is significant at what levels? So, whether it is 1 percent level or 5 percent level or 10 percent level and whether it is through 1 tailed test or 2 tailed test, which we have already discussed all these detail.

Now in the time being, in this particular matrix, so the moment you will transfer the entire complex problem into simplex by representing one simple equation like  $Y$  equal to simply  $X$  beta plus  $U$  and obviously, by the way means by the way when we start estimating then we have received  $\hat{Y}$  equal to  $X$  beta hat where beta hat equal to  $X$  transpose  $X$  inverse  $X$  transpose  $Y$ ,  $X$  inverse into  $X$  transpose  $Y$ . So, we have already **already** mentioned what is  $X$  transpose  $X$  inverse and we have already  $X$  transpose  $Y$ . So, now if we will again multiply then obviously the product will give you the coefficients of beta; that means, the estimated coefficient of beta that to beta 0, beta 1, beta 2 up to beta k. So once, you have this beta coefficient estimated beta coefficient that is from beta 0 to beta 1 beta k hat.

Next step is to check whether all these **all these** parameters are statistical significant, so that means, we like to test whether beta 0 is significant or beta 1 is significant, beta 2 is significant up to beta k is significant. So, we have to set the null **null** hypothesis accordingly corresponding to each particular parameter separately then you could test one by one. So, there may be possible, because there are k number of variable **some some** parameters may be significant, some parameters **some parameters** may not and there may be possibility that no parameters will be significant and there may be possibility that all the parameters will be significant.

It may be any case that **that** does not matter, but we have to **we have to** now justify what are the ways or what is the procedural measure, how to check all these details? So, now for that the standard component we need is called as a variance of beta hat, because here we are just getting the beta hat coefficient that is  $X^T X^{-1} X^T Y$ . So, now, when I write beta hat equal to  $X^T X^{-1} X^T Y$  it looks like one component only, but beta is a vector here, so it **it** involves beta 0, beta 1, beta 2 up to beta k.

Since, we are representing in a vector format as per the vector format, so we like to know whether this particular beta is statistically significant. So obviously, if you like to know whether this particular beta is significant then we like to know whether this you know this particular means we like to know what the variance of beta hat **beta hat** first and obviously we like to know what is the standard error then we have to set the null hypothesis then accordingly, we have to calculate the t statistic and can compare with the tabulated statistic with particular level of significance, so as usual you have to proceed accordingly.

In the meantime, we need to know, what is the variance of beta hat? So here **so here** beta hat **beta hat** equal to  $X^T X^{-1} X^T X^T Y$  this is how beta hat. So now, we need variance of beta hat **variance of beta hat** is equal to here sigma square  $X^T X^{-1}$ . So, there is actually there is mathematical derivation here, so how do we get this variance of the beta hat from the given beta hat, that is  $X^T X^{-1} X^T Y$ . So, I am not deriving all these details, because we have lack of time, so that is the reason I am omitting lots of steps here.

What you have to know? So, we just fit the entire system into the matrix approach; that means, you put a single equation, which can describe the entire structure. So, then by means by the application of matrix then we come to a conclusion, so that means, we have the estimated beta coefficient and we like to now check whether this particular beta coefficient is significant or not. So, for that we need to have variance of beta hat and variance of beta hat is calculated by this particular formula sigma square  $U X^T X^{-1}$ .

Now what is sigma square U? So, sigma square U, we have discussed the entire structure is almost all same, just we have to put it is, matrix approach is just like old wine in a new

bottle. So, we are just putting something in a very systematic way and more attractive way that is **that is** all about this particular lecture. Otherwise; these are the component, which we have already discussed. So, we are just representing in the different format. So, now variance of beta hat equal to sigma square X transpose X inverse. So, what is sigma square? Sigma square equal to summation e square by n minus k. So, n represents number of sample observation, k represent number of involvement of variable that is independent variable.

Now what is summation e square? Summation e square equal to summation Y square **summation Y square** minus summation Y hat square. Summation e square to summation Y square minus summation y hat square. So, this is how we have to write. So, now, we know there is a concept called as a total sum square is equal to explained sum square plus residual sum square. So, this is **this is** noting total sum square means this is summation Y square, this is summation Y hat square and this is summation e square this is residual sum squares.

Now how do we get all these detail? Because this is all about matrix approach, so, it is not as usual standard structures. So, it may be similar line, but there is little bit different. So, we do not like to dig all these details right now. So, we like to know, what is the simple formula, through which the job can be done; so that is our objective here.

(Refer Slide Time: 40:08)

Handwritten mathematical derivations on a blue background:

$$\sigma^2 = \frac{\sum e^2}{(n-k)}$$

$$\sum e^2 = \sum y^2 - \sum \hat{y}^2$$

$$TSS = \sum y^2 = y'y - n\bar{y}^2$$

$$ESS = y'y - \frac{(\sum y)^2}{n}$$

$$ESS = \frac{\sum y^2 - \frac{(\sum y)^2}{n}}{\sum y^2 - \frac{(\sum y)^2}{n}}$$

$$R^2 = \frac{ESS}{TSS} = \frac{\sum y^2 - \frac{(\sum y)^2}{n}}{\sum y^2 - \frac{(\sum y)^2}{n}}$$

$$\bar{R}^2 = 1 - \frac{(1-R^2)(n-1)}{(n-k)}$$

Logos for NPTEL and IIT KGP are visible in the bottom left and top right corners of the slide respectively.

Here  $\sigma^2$  equal to  $\sum e^2$  by  $n - k$ . So,  $\sum e^2$  equal to  $\sum \hat{Y}^2$  minus  $\sum$ , so the  $\sum Y^2$  minus  $\sum \hat{Y}^2$ . Now here we can write total sum square  $\sum Y^2$  equal to  $\sum \hat{Y}^2$ . So, it can be also you can say other way we can write it here, so that is nothing but  $Y^T Y - n \bar{Y}^2$ . So, this is that means, what we are doing? This  $\sum e^2$  means; so the way we have received  $\sum e^2$ ; that means  $e^T e$ . So, similarly, we have to get the  $\sum Y^2$  by multiplying  $Y$  with  $Y^T$ . So, the moment you will go for multiplication  $Y^T Y$  then you will get  $\sum e^2$ . So now, if you will simplify again then the entire structure will be  $Y^T Y - n \bar{Y}^2$ . So, this is TSS.

Similarly, RSS equal to  $Y^T Y - \beta^T X^T Y$ . So, then similarly ESS equal to  $\beta^T X^T Y - n \bar{Y}^2$ . So, this is TSS, this is RSS this is ESS. So, now we can check it whether TSS equal to RSS plus ESS. So, you see here. So, if you will club all these details then obviously this and this will be cancelled, so obviously we have this particular structure only. So, this is nothing, but this one. So that means,  $\sum Y^2$  equal to  $Y^T Y - n \bar{Y}^2$ .

Now our job is almost all done, because we need only what is estimated beta coefficient, its variance. So, once you get the variance then obviously we set the null hypothesis then you can get the standardize of beta hat means once you have a variance of beta you make it square root you will get the standard error of all these coefficient means that to beta hat here then you have to set the null hypothesis. Accordingly, we will calculate the t statistics by beta divided by standard error beta hat then ultimately we have to go accordingly. So that means, we have to check with the tabulated statistic and also significance level. So, this is one part of the story in that to specification test and that to significance of the parameter.

The next part of the problem is to know this over all fitness of the model. So, far as the overall fitness of the model, so as usual we have to we like to know what is the value of  $R^2$  and  $\bar{R}^2$  that is adjusted  $R^2$ . So, what is  $R^2$  here?  $R^2$  as usual means the formalize the ratio between explained sum square by total sum square.



Similarly, we can also apply here, because R square equal to ESS by TSS, but here ESS is equal to  $\beta' X' Y - n \bar{Y}^2$  divided by and divided by TSS that is  $Y' Y - n \bar{Y}^2$ . So, this is how the R square value is all about. So, the moment you will get R square value then accordingly you can get the adjusted R square, adjusted R square is nothing but  $1 - \frac{1 - R^2}{n - k}$ . So, this is not a problem. So, once you have R square then you can easily get R bar square provided you must have knowledge on n and you must have knowledge on k, but here we are all most all in the particular structure; that means, there are n observation and k number of variables. So, that is why there is no problem here means there is no confusion at all. So, when we will handle a particular practical problem that times n must be in a particular figures exact figures here also particular figure that to exact figure. So, accordingly you have to go for the adjusted R square.

Similarly, the way we are testing the significance of the parameters through t statistics here also we have to test the overall fitness of the model with the F statistics. So, now as usual we have to we like to know what is F statistic here. So, F statistic again the ratio between explained sum square with a residual sum square. So, accordingly you will get the F statistic. So, again you have to check the tabulated statistic with corresponding degrees of freedom that is  $n - 1$  to  $n - k$ . So obviously, accordingly you have to see whether the overall fitness of the model that too R square is statistical significant or not if it is so, then you have to proceed accordingly.

That means, what we have discussed till now is that we just transpose the entire structure into particular format, so complete in a matrix approach. So, in the earlier versions of our modelling that to bivariate and trivariate and multivariate we just start deriving all these equations then we apply summation accordingly we will get the all estimated equations. That means, we have to apply the wireless technique, then we have to differentiate with respect to this error sum squares to that estimated parameters and ultimately with respect to particular modelling system that too bivariate, trivariate and multivariate .We estimate the parameters or you like to differentiate the sum square with respect to all these parameters.

Then once you have standards in equation we transpose the standard equation to some particular either in matrix format or you can say principle or you can directly solve

this simultaneous equation if you are very comfortable. If the problem is bivariate or trivariate, it is very easy to solve very quickly without having the touch of matrix even (( )) etcetera etcetera, but when the problem is more complicated that means, when the system will be more than three, more than four or you can say five, six seven, like this and that times it is very difficult to handle manually.

Either you go with software technique or you can say there are standard so many software(s) are there, so you just enter the data and give click click you will get the results it is not difficult at all. But, if you have not statistical software or if you are not allowed to use statistical software then you you cannot handle this multivariate problems within a time timeframe and that too particularly for exam point of view or you can say class point of view.

In that context to transfer this complexity to simplicity, so what you have to do? You have to apply the matrix approach then automatically you can get these things easily. So, it is not such a serious issue which is very easy to tackle the particular problem. That means what we can conclude that matrix has a lots of application, lots of utility that to use in econometric modelling. Particularly, when there is multivariate analysis then the matrix has lots of typical it has a lots if beautiful beautiness so, that we can transfer the entire things into a simplicity format. So, this is how the entire structure all about.

Now whatever we have discussed the matrix approach for multivariate modelling we can we can examine or we can justify the same thing with respect to a particular problem. So, it does not mean that the matrix approach will be always applicable to multivariate problems, so even within the bivariate setup you can you can also we can also or you can say apply matrix approach directly. But, matrix approach it is a very simple technique, but when the system is very consist means very the structure is very small then obviously that times the matrix approach looks like a complex, but when the problem is bigger and very complex then that time the matrix approach will make the thing very simpler. That is why you have to take a decision and you have to you have to proceed accordingly.

But, in the meantime, it is not possible to handle many variables and that application of matrix, because ultimately whether using the simple simple analysis or matrix analysis ultimately you have to handle the entire means all these variables at a time. So now, when we will apply the matrix approach then that to multiple variable that times your

size of the matrix will be bigger, so when the size of the matrix is bigger that means, with respect to its orders then matrix multiplication will also be not much difficult, but it will take lots of time and it will have little bit complexity.

To transfer all these **all these** structure into simple, so what you have to do? You have to look into very carefully then apply accordingly. So, my suggestion is that when the system is very simple that to we can say bivariate and trivariate then you need not require to use matrix, you can go ahead with the simple calculation, simple estimation process and when the problem will be more complicated like means particularly in the multivariate cases, so that time you have to apply matrix approach. But, if you will apply if you are very smart and you have a very smart knowledge in matrix, then even if you handle this particular bivariate system itself. So, you use matrix very beginning and to **to** till end.

But, in generally this particular technique is useful just to get the estimated model. So, now once the estimated models you have then the things will be in other way around, that means, after estimation model the next step is just to check whether this model is completely reliable or not; that means, it is free from all these errors. So, that is how we are doing for all diagnostic tests and all these significance test etcetera **etcetera**.

This means after estimation whatever these steps are available to get the respected model, this means in that case matrix may not be very useful, but to get the problem means at the starting point to get the estimated model, so within that process matrix has a lots of applications, means after that there can be use of matrix, but that usefulness is not so important because, having the estimated models then lots of things can be done easily without having the application of matrix further, but to get the estimated model that to means this particular component that matrix approach to multivariate is specifically very useful or very handy for classroom only or that to exam point of view. But, when you are handling big **big** problems particularly in a research oriented problems or the research projects, so that time obviously there is no point to do all these things in the room, so you have to directly handle with software. So, that time matrix has no utility at all. But, specifically matrix is utility when you are in the class and your job is or your job is how quickly you can evaluate this particular model without having the use of software(s) only.

Then in that case, even if without the use of you can say excel sheet. So, you have to everything you have to do through you can say matrix. So, that times it is very easy, but when you have the opportunity to handle excel sheet and you can say statistical software that times it is not required at all to apply matrix approach. That means, what is my conclusion is that matrix approach is basically very useful when there is no access to internet or you can (()). If there is access to internet and excel sheet that times it is not required at all use of matrix approach.

With this, we can conclude this particular session and in the next class, we will discuss the same matrix approach to multivariate with particular problems. We we we are just theoretical discussing these issues, but practically how we have to design all these structure into in the matrix format that is more interesting and more attractive, so that we will discuss in the next class. Thank you very much, have a nice day.