Econometric Modelling

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

Lecture No. # 24
Heteroscedasticity Problem

Good after noon, this is Doctor Pradhan here, welcome to NPTEL project on econometric modeling. Today, we will discuss the Heteroscedasticity problem. In the last couple of lectures, we have discussed the basic framework of econometric modelling and last couple of lectures, we have discussed various aspects of econometric modelling the aspect means problem aspects. Like in the last two to four lectures, we have discussed the problem of multicollinearity and problem of autocorrelation. Actually these problems are extra problems related to what we have discussed in the fitness of the model, where we have highlighted briefly or you can elaborate two way specification test or the goodness fit test. So having specification test or having goodness of fit test is not sufficient to say that the model is a best fit test.

Even, if the model is a true specification test and true goodness of fit test. Still it is or it cannot be considered as a best model for forecasting or policy use. The way we highlighted few problems already, so like multicollinearity, autocorrelations and today‘s discussion is all about Heteroscedasticity. First of all, what is Heteroscedasticity? The way, we have discussed multicollinearity into autocorrelation; similarly, we will discuss the issue of Heteroscedasticity; this is a very interesting issue, and you know very important for the econometric modelling.

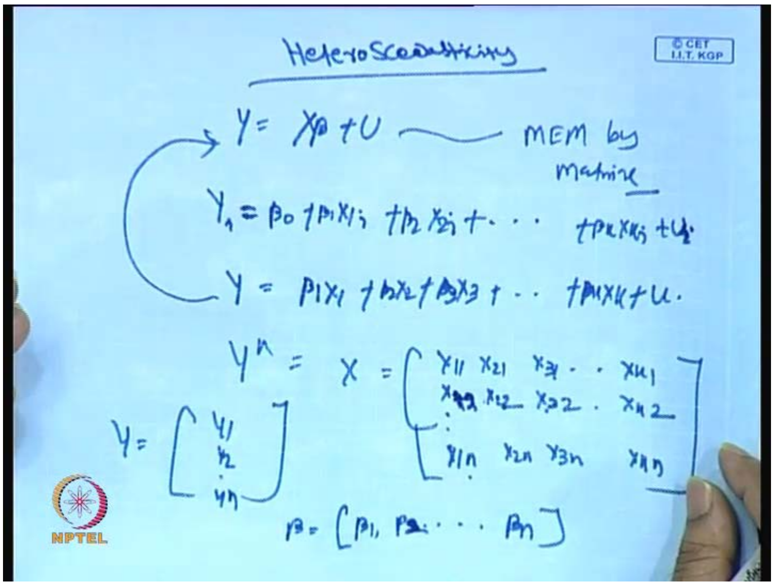
Because this particular component maybe you know very helpful for cross sectional analysis, also very useful in times the series analysis. We have very beautiful model like (( )) etcetera. So during that times we will discuss details about the Heteroscedasticity issue. Today we will briefly highlight the structure of Heteroscedasticity. Heteroscedasticity basically, the counter part of Homoscedasticity; Homoscedasticity is always positive for the econometric modelling. If anything is attach directly attaching the homoscedasticity issue then obviously it is good for the modelling or you can say it is very good sign for positive sign for the fitness of the models. But, if there is any question of Heteroscedasticity coming into that particular angle then it is very problem for the modeling or it is a negative aspects of the econometric modelling.

So, to transfer this Heteroscedasticity; homoscedasticity is the core objective of the econometric modelling means this particular discussion. Now we will highlight what is all about the Heteroscedasticity issue how we will bring that particular components? You remember in the last lectures I had mentioned, so far as econometric modelling is concerned there are three aspects; first aspects the dependent classification, independent classification and error classification. Till now, we are discussing one dependent variables, several independent variables; and in the last two lectures, we have discussed several error terms in fact we start with error term one you what within you will create so many use.

If there is any such game you know with respect to you, then obviously this type. This is very interesting problem or extra problem related to this econometric modelling. Today this particular issue Heteroscedasticity is purely on error terms like you know autocorrelation issue. The thing is that whether it is a bivariate model or multivariate models. Heteroscedasticity is always issue because it is once you have error terms then you will create several error terms and now there is several if different modelling you can build or we can you can put a structures so that it can be a further analyzed or you can say and go for a interesting issues etcetera.

Now we will see first how it is exactly coming; this Heteroscedasticity issue. Then we will discuss with respect to its nature, consequences and its causes and its detection rules and also its solution rules. First of all what is all about the Heteroscedasticity? Let me briefly highlight or bring the issue of Heteroscedasticity then we will discuss the details about the Heteroscedasticity.

(Refer Slide Time: 05:10)



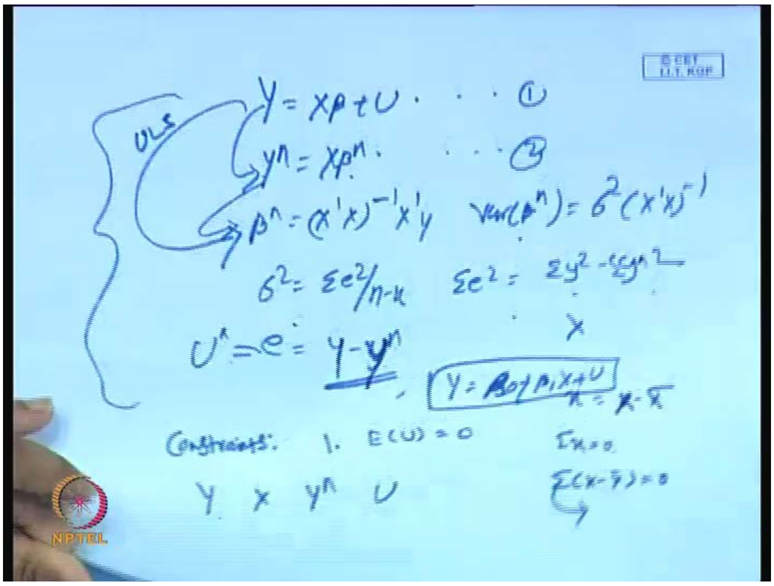
Let us take a case of econometric modelling for Y equal to say X beta plus U. This is you know generalized multivariate models, multivariate econometric models by matrix approach. This is multivariate econometric model by basic econometric approach or else you can write like this way. Y equal to beta 0 beta 1 X 1 X 1 i plus beta 2 X 2 i plus continue beta k X k i plus U i all right. Now obviously this is U i. Last class we have discussed in fact U i it can continue like U 1 U 2 U 3 like this way but, first you start with this one then if first have U after that you can go for extension of U.

Now this particular model can be written this way simply beta 1 you know beta 1 X 1 beta 2 X 2 beta 1 X 2 plus beta X 2 plus beta 3 X 3 plus you can say beta k X k plus U. I can write you also like this removing the intercept terms. So intercept terms in fact in the deep labels it may very less impart on dependent variable. When you have a multivariate models and when there are k number of independent variable in the system. That time intercept has a minor rule in fact it will go for you know difference methods like first difference method of second difference methods.

That times by default it will be removed, so in the long run the impact of intercept is a not intercept on the impart of you know slope coefficient is very important like beta 1, beta 2, beta 3, etcetera. Obviously we can have this type of model now, we have Y hat equal to we have Y hat similarly, you can create that term matrix here if will transfer this particular equation into this particular format then obviously, where X equal to X 1 1 X 2 1 X 3 1 up to X k 1 then similarly, X 2 1 X 2 2 X 3 2 then X k 2 this is X k 2.

This is X 1 2 X 2 2 then X k 2. Similarly, it will continue up to X 1 n X 2 n, then enter this is X 1 2. This is X 1 2 X 2 2 then X 3 2 like this way. So X 1 n X 2 2 to X 2 n then X 3 n then it will continue X k n. This is how the variable X is all about Y is obviously like this Y 1 Y 2 up to Y n and a beta equal to you know beta 1, beta 2, beta 3 so, beta will be this beta will be like this beta 1, beta 2 it is in problem format like beta 2 beta up to beta n. Similarly, U 1 U n you can create. This is the basic frame work of econometric modelling with respect to you know multivariate model or you can say multivariate models by matrix approach.

(Refer Slide Time: 08:47)



Now I little bit highlight here, so how this error terms will be coming into the picture. So, briefly once again I will start with the X beta plus U now, this is a cell set up. Ultimately our n is to get Y hat Y hat equal to X beta hat. Now what is x beta hat here? beta hat equal to X transpose X inverse X transpose Y. This is beta hats and variance of beta hat is a variance of beta hat beta hat is equal to sigma square X transpose X inverse then, sigma square equal to summation e square by a n minus k summation e square equal to summation y square minus summation y hat squares. This is how you have to proceed but, you remember to transfer this equation to this equation. That means y means particular this first to this equation this is equation number two means it is the extension of this one.

That means Y hat equal to X beta hat square beta hat equal to this much and you know followed by next step is to test the significance of the parameters means all beta for that you need to have a variance of beta hat and for that need you have to sigma squares and followed by summation e square y square y hat square. It is chain how you have to proceed, we have discussed long back. I am not highlighting in details but, one thing I will like to clarify here is that, the moment the moment you transfer this true model to estimated models and by the way you are considering that it is the best models which can the used for which can be used for forecasting and policy use.

That times there is a small small issue you have to take care before you directly proceed to forecasting or policy use. For instance you know the moment you will transfer from Y X beta plus U 2 Y hat equal to X beta hat. Now this is the estimated equation and for that we till now we have discussed about the o l s technique. By the application of o l s we transfer this model Y X beta plus u 2 Y hat X beta hat. In fact this is a simple regression model and this is you know estimated model.

Now by the way once you will make difference Y minus Y hat then you will get the error term all right. Now e equal to Y minus Y hat or sometimes Y minus Y hat. Sometimes it can be called as you hat also. Sometimes it can also be called as you hat. You hat or e equal to Y minus Y hat that is the reference of error term.

Now, with this basic setup all these things are absolutely ok or can be represented here is in the form of the modelling for the thing is the moment is since we are utilizing or using o l s to transfer this original equation to estimated equation and that model we are considering that is best fitted model. Obviously it most go through its condition means constant sets. So, o l s has lots of constants; some of constants are directly related to X; some of the constant direct related to Y; some of the constraints are directly related to U and some of the constant is very generalized in natures.

But here our issue is the constant on error term only. What are the restrictions or conditions you have to put for error term before you using this estimated model for both forecasting or policy use? That is our agenda we have to discussion. Now the moment will transfer the equation to this equation. Then you know we have already used o l s technique here is. You have used already o l s technique. So, o l s technique is based on certain assumptions which we have discussed long back but, right now I am just briefly bringing that issue, because that issue will give you the indication of Heteroscedasticity.

Now, if will that will bring the problem of Heteroscedasticity, so will like to now what is exactly the error terms and what are the restriction we can put or o l s have in the case of error terms. In fact we are not bothering about the general conditions or the condition related to Y and the condition related to X.

We are very much interested for this particular Heteroscedasticity issue, what is the condition or constraints for this error term U? Now what are the constraints; first constraints first constraints is E of U must be equal to 0 that means mean of error term must be equal to 0; that is as usual you know basic of the statistics or you can say beginning of a univariate analysis. So mean of U means its expected value of U, where we use probability then we are discussing about the mean issue.

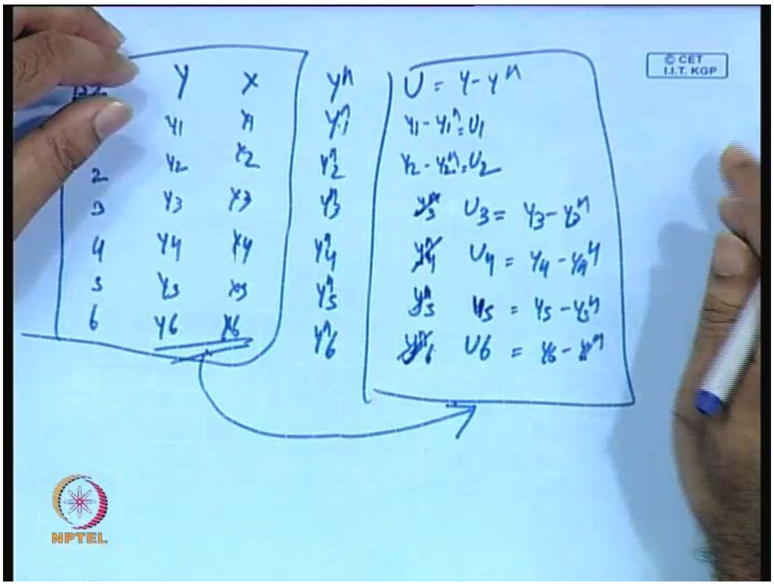
You know when a simple variable say X and Y then obviously its deviation is always equal to 0 that deviation is mean in fact. That deviation may be with respect true value or with respect to with the attachment of probability value. If it is with respect to attachment of probability label then we use the expectation obviously the expectation will give you the mean variance etcetera. But, you know like you know if you go by original issues the mean component which is put it in is equal to or you can site in the form of average. When you will calculate average say mean is one of the standard examples.

One of the standard properties of mean is that is some of the deviation of items from the arithmetic mean equal to 0 like this for instant if will expire able then you know we have X standard deviation upon X. X equal to X minus X bar that means summation X should be exactly equal to 0. X minus X bar should be exactly equal to 0. So summation X equal to 0 means summation X minus X bar equal to 0 so that, there is a statistical proof through which can get to know deviation of some of the deviation of item from the arithmetic mean must be equal to 0. Since it is a mathematical proof you can verify very easily. The way we are highlighting the issue with respect to a standard variable; say X. Similarly, it can be also applicable for U because ultimately U is another variables here.

We start with Y and X. Lets assume that this particular model is based on two variables only. So, Y minus Y hat lets a say Y equal to simply beta 0 plus beta 1 X. Simple this much of information plus U this is the means for simplicity we are putting only bivariate models.

Because, the Heteroscedasticity problem can be discussed on under bivariate condition and can be discussed under trivariate condition and that can be discussed on the multivariate condition. Now we are whatever conditions were there so we have to discuss the structure or you can say strategy concepts are all most all more or less same in fact. That is why instead of going a multivariate problem it is better to raise this issues in a univariate setup. Ultimately that will give you path you how discuss at Heteroscedasticity in a multivariate angle.

(Refer Slide Time: 16:58)



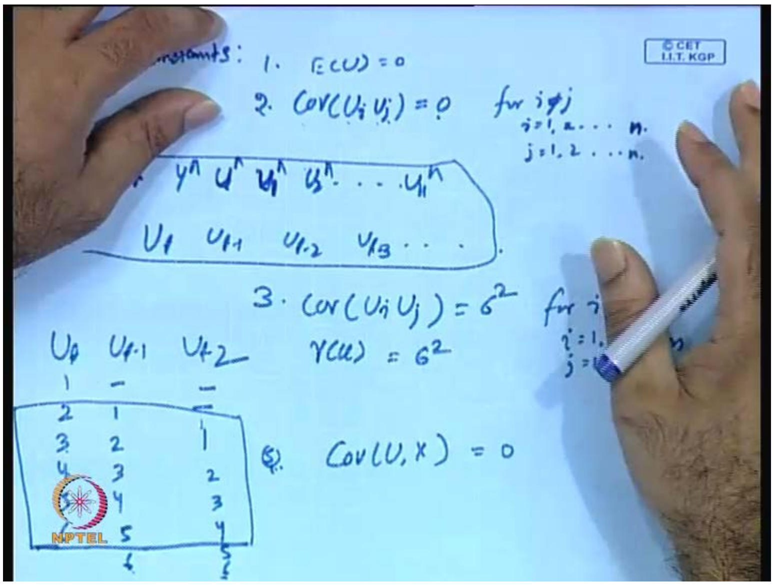
Now what is U hat? like you know Y and X. We will get Y hat similarly, will get U that means all together there are four variables in the system. I will highlight put it in other way, let us see here is Y then X then you know Y hat then you know U. This is sample information; Sample information says 1 2 3 4 5 6, let us 6 only then this is Y 1; this is Y 2; this is Y 3; this is 4; this is 5; this is Y 6. But, remember Y 1, Y 2, Y 3 are not variables; these are all sample information only.

Variable is Y. Y 1 is the sample for one unit; Y 2 is the sample units of Y for second unit; Y 3 is the sample unit of Y for sample 3. Y 4 as the information of Y on sample for like this will go to. Let us say there are 6 items under Y. Similarly, X can be analyzed X 1 X 2 X 3 X 4 X 5 and X 6. Initially you have this much of information only, by the process then we will get Y hat obviously the way we are writing Y and X. Obviously Y hat 1 Y hat 2 Y hat 3 Y hat 4 Y hat 5 Y hat 6 this is a how you have to derive U equal to Y minus Y hat.

That means it is Y 1 minus Y 1 hat then Y 2 minus Y 2 hat. Similarly, this is y hat 3 y hat 4 y hat 5 y hat 6 that is in a not small, y minus Y hat we are calling it small; it is better not to say y hat y 3 y 4 you can quality this one. This is U 1; this is U 2; this is U 3; this is U 4; this is U 5; this is U 6. This is error term only so that means Y 1 equal to this much; Y 2 is a this much; Y hat equal to Y 3 minus Y 3 hat. This is equal to Y 4 minus Y 4 hat; this is equal to Y 5 minus Y 5 hat; this is equal to Y 6 minus Y 6 hat. This is how you will get this particular series so that means Y is itself is a variable here.

We have four different variables Y X Y hat and U. So our constant is; Y X there is a relationship. How this relationship will be getting affected by this error observations? This is how we will get the error since this particular issue.

(Refer Slide Time: 19:55)



Now, we are discussing what are the conditions or constraints against the error terms, that means we have already discuss the constraints; its first constraint say first is E of U equal to 0 that is mean of error term must be equal to 0; second covariance of U i U j a equal to 0 for i not equal to j and i equal to 1 2 up to n. Similarly, j can be 1 2 up to n.

This is how the entire structure is; that means covariance of U i equal to U i Uj is equal to 0. You see here Y X and Y hat then U hat. Once U hat then you will get that let us say that is original U 1. You then get Y 1 hat U 1 hat U 2 hat, similarly you can continue you can say U k hat like you know we have discussed little bit about time series in the last class.

Once you have Y then you will go y t minus 1; y t minus 2; y t minus 3 and so on. This particular structure can be analyzed similar way. Once you have U t, you can go with the U t minus 1 U t minus 2 U t minus 3 U t minus 4 like this way or you can go with respect to U t U t plus 1 U t plus 2 U t plus 3 like this way. Either in say past collection or you say present collection, future collection.

But right know, we are discussing about the future collection. There are sample observations like this way let us say like this U t U t minus 1 U t minus 2 U t minus 3 like this. It will continue like this. But remember one thing when you will create additional U for given U then obviously one of the problem you will be face is that number of sample size will get reduced, because to maintain the consistency of all the variables at a time.

You have to find out the appropriate log length and provided it has to satisfied with the degrees of freedom. Once until unless you have that then it is problem for you that means what I like to say one of the interesting condition for this error correction is that covariance of U i U j equal to 0 for i not equal to j. In fact this particular problem we have discussed in the last class that is nothing but, the autocorrelation problem. Autocorrelation means there is no such covariance between the error terms.

Now third assumption is that covariance of U i U j is equal to sigma square for i equal to j; i equal to 1 2 up to n; j equal to 1 2 up to n that means covariance of U i upon U j for i equal to j mean, it is a nothing but, variance of U it is nothing but, variance of U. So variance of U must be constant.

Once you have to used so first check it whether mean of that error term is exactly equal to 0; this is step one. If it is satisfied then you have to go to step two. Step two means you create additional U t. U is like U t minus 1; U t minus 2; U t minus 3; like this. That means your sample point will be coming like this is U t; U t minus 1; this is U t minus 2.

If it is coming 1 2 3 4 5 6 then obviously this till 1 2 3 4 5 6 then obviously this will be 1 2 3 4 5 and 6.

So 5 and 6 that means you are loosing here one sample you are loosing here two samples and there is addition of extra sample. This side you have to remove and this side you have to remove. So, that means now the model can be fitted with 1 2 3 4 5. 5 observations you have to take into considerations before using you know second log length.

Similarly, if you will add another variable then you're going to loose the degree of freedom one that means instead of 5 then you will take only four sample points only. Once we add one after another variable then obviously means with respect to log introduction then obviously we will loose sample observations. That is how you must be very careful how much log length you have to consider for a particular problem.

That means the moment you will use log length that is nothing but, the indication of time series modelling. One of standard assumption of the times series modelling is that your sample observation should be substantially very high. If the sample observation is substantially very high then obviously if will create number of the U(s) within a given setup then it will not serious problem.

But, if you have a very small size means sample size then obviously to have more number of U is very difficult so that is why before you handling this type of problems like (( )) and Heteroscedasticity. One of the standard entry point is that your sample observation should be exclusively very high. If the sample observation is very low then it is very difficult to have the best fitted models and that means it is very difficult to through this autocorrelation and Heteroscedasticity so typically.

That is how you have to be very careful so, that means variance of you must be here sigma square. It has to be satisfied then third forth condition is that covariance of U upon X that means it is a linkage with the independent variable. Covariance of U upon X also is equal to 0. Otherwise it will automatically give an indication of like multicolinearity problem because what is multicollinearity? So, having linear relationship among the regressors that means U is also one type of regressor here, because U is situated you can say included in the right hand side of the equation.

The moment something is in right hand side of the equation that can be considered as a independent clusters. Obviously there may be some colinearity between U and x but, the model requirement is that or model the system of modelling is that. Whatever variables in your right sides, whatever in the independent clusters these variables should be independent to each other. If there not independent to each other an obviously there is a multicolinearity problem; first you solve that particular problem, then you have to go for other problem, because it is a step wise process.

One problem once you will solve you will go to another problem; once another problem will get solved then you go to another problem but, in the same times you must be very careful when you will be solve a second problem the first problem which already clarified or cleared that can be again added to another problem.

When you have a series of test or series of constraints in your front and obviously to clear all these constraints it is very difficult job that means you have to go for very optimum condition or you can say perfect combination, where everything can manage in a typical way. Otherwise it will get affected by the systems. Covariance of U upon X must be equal to 0. These are the standard assumption we are sighting in the o l s technique to have the true regulation; true regression and true you can say estimated regressions.

That means the model which we have the shift o l s technique application o l s technique can be used or can be considered as the best after conformation from the specification test and also diagnosis test and let us assume that it is also confirmed the autocorrelation and multicolinearity. Still that model cannot be used for you can say forecasting or policy wanted unless take the Heteroscedasticity, because heteroscedasticity is a serious problem. Heteroscedasticity very critical meaning is that it is a heterogeneous in nature but, generally so far as a cluster analysis requirement is that objective. The model will be perfectly ok if your sample units are very homogeneous in nature. That is how clustering analysis has lots of utility in this econometric modelling. You must be very careful about that particular issue.

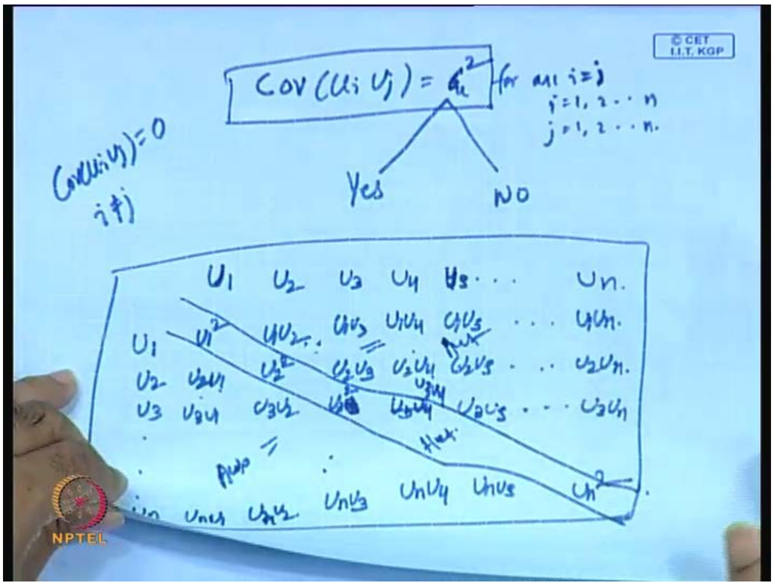
This is the structure how to means these are the conditions you have to site before using the estimated model for goodness fit or specification test or also for use of forecasting or policy use. Now here the game of Heteroscedasticity will start. Generally if this is the case there is no such problem; this is you know by definitional problem; one this is the problem of autocorrelation; this is problem of Heteroscedasticity and this is another way of representing the multicolinearity issue.

This particular problem is more dangerous and this particular problem is also more dangerous. In the time these multicollinearity issues you have already discussed. This is also we have discussed. Now this particular problem we have to highlight in a broad way. That means with a given setup once you apply ols technique you will get the estimated model. Once you get the estimated models so you have to go for lots of test before you assume that this model or before you site that is model is considered as the best.

For that you have to clear through specification test; goodness of fit test; then again to solve the multicollinearity issue to solve the autocorrelation issue and then you will come down to Heteroscedasticity issue. Once it will pass then you have to go for any other problems but, these are the standard tricks you have to follow before you move to any other problems directly.

Now you see here how is that particular setup that means you see here what is Heteroscedasticity issue out of we have sited all together high different assumptions. Now if you if you go through that 5 sited assumptions then one such assumption is very typical you know integration with this particular component that is Heteroscedasticity issue.

(Refer Slide Time: 31:08)



That means we have sited one of the condition is that covariance of U i U j equal to sigma square, for all i equal to j and i equal to 1 2 up to n and j equal to 1 2 up to n. This is how we have a highlighted. Now you see here this is sigma square means it is a constant this is not sigma this is called as a sigma square. Generally we use sigma square instead of writing sigma square we can use sigma square U, that means it is the moment will put sigma square U it is an indication that it is error variance, otherwise I can write sigma square X sigma square V. It is a variance of standardization of sigma square X means is variance of X; sigma square y means variance of Y.

Similarly, sigma square U it is a variance of U. So the rule is that the variation should be constant over the error terms. There may be some variations in the sample points but, here there should be you know constant variance. If the variance is not constant then obviously it will go against the system of econometric modelling, that too typically for this Heteroscedasticity issue. Now if by any chance if it is done or if it is obtained then that means there are two questions here: yes questions and no questions.

That means this is one of the assumption. Now if it is assumption then by default it will turn into a hypothesis. Hypothesis is that covariance of U i U j is a equal to a sigma; not equal to sigma; or equal to sigma square U and against the alternative hypothesis that covariance U i upon U j is not equal to sigma square U is obviously sigma square is always there.

For instance; if you will integrate U 1 with U 2 you will get sigma square. If U 1 means U 1 1 then U 2 2 like this where you know the cross studying will be in a diagonal setup like this you see here is what is error means error through variance matrix covariance matrix. So you see you have U first then U will first U equality U 1 then U 2 another variable U 3 U 4 U 5. Let us say this is U n.

Similarly, you have here U 1 U 2 U 3 up to you have U n here. Now we like to prepare a variance covariance matrix we like to prepare a variance covariance matrix. This will give you indication of means this particular matrix will give you signal for autocorrelation problem and it will give signal for the Heteroscedasticity problem or you can say Homoscedasticity problem; Homoscedasticity is the counter part of Heteroscedasticity.

Now by the moment you will say variance covariance matrix then obviously we can prepare this way so it is nothing but, U one square; this is U 1 U 2; then this is U 1 U 3; this is U 1 U 4; then this is U 1 U 5; then it will continue U 1 U n all right. This is first component and similarly, U 2 1 U 2 U 1 U 2 square U 2 U 3 U 2 U 4 then U 2 U 5 then it will continue U 2 U n. Similarly, U 3 U 1 U 3 U 2 then U 3 U 2 U 3 squares then U 3 U 4 then U 3 U 5 then continue U 3 U n. Similarly, U n 1 U n U 1 U n U 2 then U n U 3 then U n U 4 then U n U 5 U n squares.

This is how the complete variance covariance matrix where there are n number of error terms that means start with a single error terms U 1 then you will create a several error terms U 2 U 3 U 4 U 5 up to U n. Now if you will put time series aspect then it will start like this U t U t minus 1 U t minus 2 U t minus 3 up to you can say U t minus k. Now if we say k at means t equal to t minus k depending upon and the value of k and t so the error term can be evaluated. The evaluation means like this U 1 U 2 up to U n and this is U 1 U 2 up to U n.

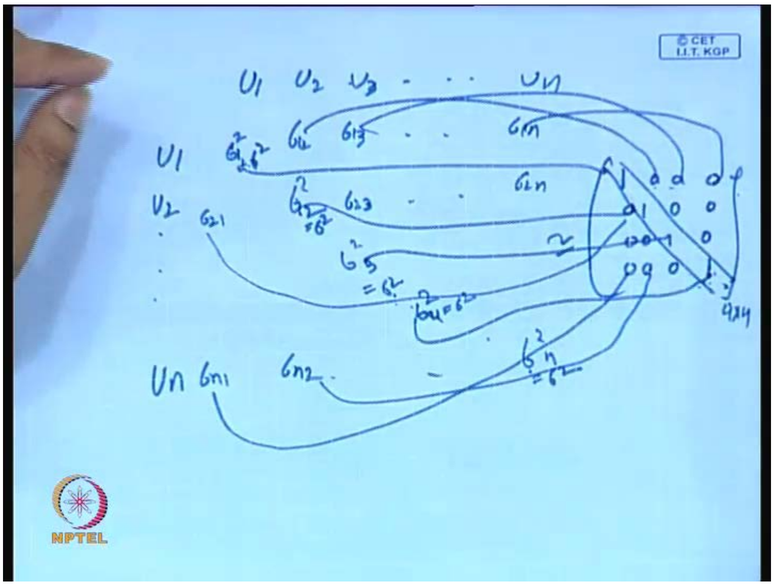
Now if will multiply vertically or horizontally then obviously you will this particular matrix. Now this particular matrix has three different steps; there are three difference steps of this particular matrix. Like this way this U 3 U 4 will come here only. This is the diagonal matrix you have here. That means, this is the diagonal elements that means you have two different matrix; one is a off diagonal and this is on diagonal; this is of diagonal; this is diagonal.

That means have 3 difference setups, one is diagonal elements then off diagonal; on diagonals and off diagonals. These are you know symmetric in nature for instance U 1 U 2 is nothing but, U 2 U y. Similarly, if we have this much of information then automatically transpose this and paste it here then you will get this particular matrix. You remember here this side off diagonals, on diagonals are more or less same because it will give you the situation, where off diagonal and on diagonal then covariance of U i U j is equal to as per the conditionality it should be 0 for i not equal to j.

When U 1 U 2 is there then obviously i not equal to j, when U one square is there then it is the case where i equal to j but, the starting is the covariance of U i U j. When i j are equal then you will get diagonal elements when i not equal to j then either you may be in off diagonals or may be in the on diagonal. That is depending upon the situation. Now if you were here or you are here then this particular problem is called as autocorrelation problem. If you are in off diagonal, on diagonal then you are in this is our autocorrelation problem and this is Heteroscedasticity problem. But we have not highlighted what is Heteroscedasticity here but, auto correlation you can directly get to know because it is the if covariance matrix, variance matrix but, you know so accurate issue of autocorrelation Heteroscedasticity is that you see here but, be careful one thing here. In fact this autocorrelation like you know there is lots of connection between multicollinearity and autocorrelation.

There is also similar and in the similar way there is also some kind of similarity in the case of auto correlation and Heteroscedasticity. One of the similarity is that the problem is discussed under the condition of error terms only. Basically there may be some it is definitely has an integration with a dependent variable or independent variable but, the root is the error term for both autocorrelation and Heteroscedasticity. Now after having the variance covariance matrix.

(Refer Slide Time: 39:43)



So that will all give you clear cut indication of Heteroscedasticity issue to get the clear cut indication of Heteroscedasticity issue. Let me highlight here what is exactly that one so the transfer matrix is like this U 1 U 2 it will be U 3 it will continue like U n then this side you will continue U 1 U 2 up to U n. Now this is U know this is sigma square means U 1 square so we will call this sigma square square 1; this side we will call it sigma square 2; this side we will call it sigma square 3; so it will continue then you will get sigma square n. This is called as a sigma 1 2, sigma 1 3 and sigma 1 n then sigma 2 1, sigma 2 square, sigma 2 3, it will sigma 2 n , sigma n 1 sigma n 2 then it is a sigma n square.

This is how the matrix all about by default this should be unit say sigma square, this should be equal to sigma square, this should be equal to sigma square, this should be equal to sigma square. Let us say unit means it is just like the case of unit matrix when you will go for matrix algebra, then it is the structure means in o l s technique constraints condition is that the error variance covariance of error matrix must be must be equal to the unit matrix.

What is meant by unit matrix, the matrix unit matrix has two interesting properties: first thing is elements are square in nature means row and column will be same, like you know two into 2 into 3 4 into 4 5 into 5 n into n. That is the matrix that means matrix must be square and the value of matrix must be exactly equal to 1. How do you get value of means this type of matrix? That means in all diagonal elements value must be equal to 1 and other diagonal elements the value must be exactly equal to 0.

That is the representation of know unit matrix and in fact in econometric matrix that is the requirement of o l s application. Means one of the standard assumption of ols is that the error matrix, error term, error component matrix that is covariance matrix should follow the unit matrix. That means this will be transferred to like this way this is 1 0 0 0 0 1 0 0 0 0 1 0 0 0 0 1. This is the case like this way. This is 1 2 3 4 1 2 3 4. This is 4 into 4 matrix.

This is four into four matrix; that means we are just considering U 1 U 2 U 3 and U 4 we are not going up to n term. Now as usual I have mentioned, this particular term is a 1 1 1 and other terms in the 0 0 0 that means you can connect here. Sigma square 1 is this much; sigma 2 square is this much; sigma 3 square is this much; sigma square 4 this is equal to sigma this will connect to this one.

This is how if this transformation means value of this particular matrix will be like this, then it is called as a homogeneous issue. It is called as a homogeneous issue or it is otherwise called as a homoscedasticity problems. This particular if we mean we are directly targeting the diagonal elements after the value of that particular matrix means a quantitative figure of that particular matrix not value of the matrix; a value of the matrix when the value of the determinant must be exactly equal to 1.

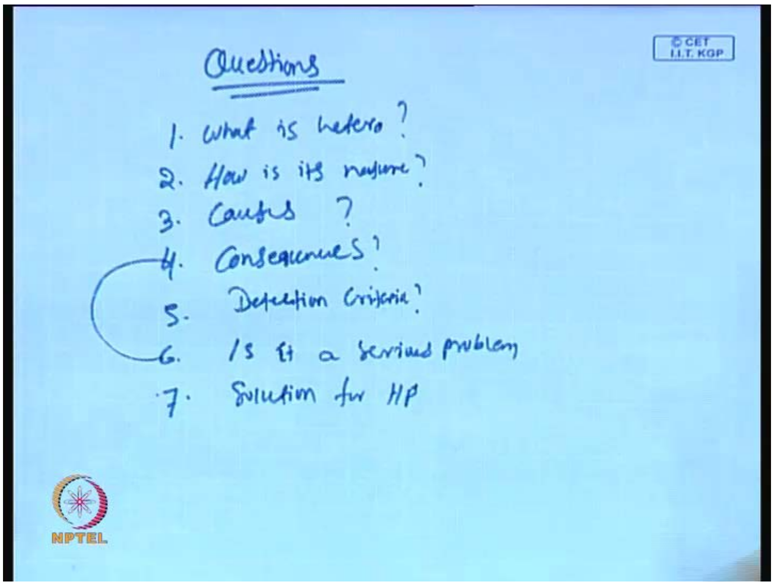
We are not obviously if you will get this type of matrix then value is always equal to 1 that is how it is called as unit matrix. That means the value one is a unit only. If that is the case then there is no such Heteroscedasticity problem but, if all are not following the unit root matrix or singular matrix or matrix having the values one unit matrix then obviously there is a problem of homoscedasticity. That will lead to Heteroscedasticity. That means our requirement is homoscedasticity. If not homoscedasticity then, obviously it is a Heteroscedasticity problem. Obviously Heteroscedasticity the presence of Heteroscedasticity is not good sign for the econometric modelling.

You got to bring the Heteroscedasticity to homoscedasticity. You may not to get exactly homoscedasticity issue but, you have to be very close to homoscedasticity that means error variance by default should be very equal for and each and every sample structure all right. Now if it is 0 0 0 then that means this will connect this one; this will connect to this one; this will connect to this one. Like this similarly, this one connect to this one then the sigma one connect to this one like this.

This is how variance covariance of matrix that means the off diagonal elements and on diagonal elements for this variance, covariance of error matrix is equal to 0 then obviously it is called as a autocorrelation problem that means it will give you indication that there is no autocorrelation problem again similarly, if the diagonal elements are giving this signal of 1 1 1 1 then it is giving a green signal that it is the homoscedasticity problem that means there is no such Heteroscedasticity problem.

So the moment you will get covariance off diagonal, on diagonal elements are all 0 then it by the default it give a signal that there is no autocorrelation problem. Then again if all diagonal elements are equal with respect to all sample size then it will give a green signal that there is no such Heteroscedasticity problem. Now, we have various questions here issue is what is all about this Heteroscedasticity means with this basic introductions now will have highlight what is all about the Heteroscedasticity issue.

(Refer Slide Time: 46:08)



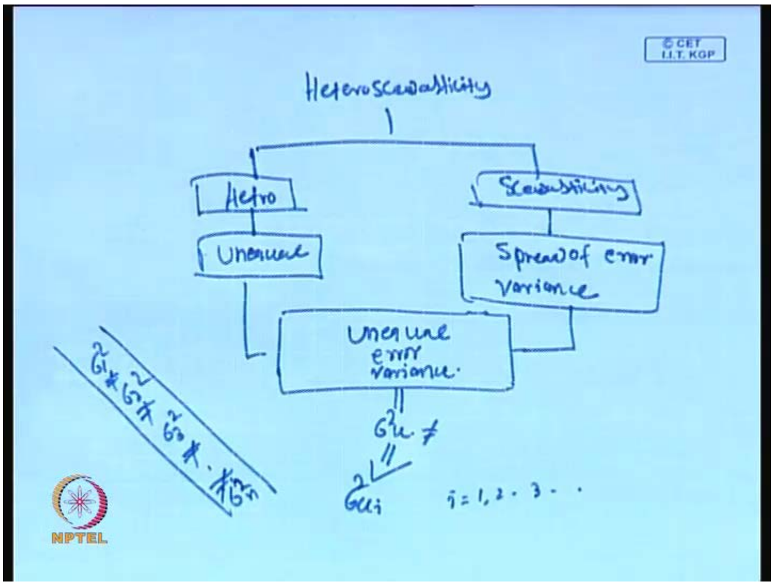
That means we have now number of questions of what are the questions. We have number of questions, first it is what is Heteroscedasticity exactly? Then second - how is its nature? This is second aspect of the problem, third: what are the causes work what are the consequences? What are the consequences then, fifth: what are the detection criteria? Sixth - is it a serious problem that means the question is the presence of

Heteroscedasticity is it a serious problem that we have to highlight, that means automatically it will give connection to consequences.

Obviously it is consequences the moment you have said consequences means what are the negative impact on the presence of Heteroscedasticity. Once if it is some positive impact then that is no problem but, if it is a negative impact then obviously you have to remove the negative impact.

So, is it a solution for this particular means is it a solution for the presence of Heteroscedasticity problems? Seventh - the solution for Heteroscedasticity problem; so that means you remember when we were talking about solution to the Heteroscedasticity problem then obviously by default Heteroscedasticity you can say means it is a negative impact on econometric modelling. You need to have its detections. That means we need to have detection and you need to have its solutions.

(Refer Slide time: 48:29)



Now next issue is we have to clearly highlight what is exactly the Heteroscedasticity issue that means first point is we have mentioned what is Heteroscedasticity. You see basically Heteroscedasticity divided into two parts one is called as Hetro then Scedasticity like homoscedasticity this is Homo and Sscedasticity. This is Heteroscedasticity Hetero and Scedasticity. Hetro means it is unequal variance.

This hetero represents unequal. Homo means it is equal homogenous means is equal for instance suppose there are 3 products are there. I like to be keen on quantities 10 10 10 that is a how homogenous in numbers only. It may be different wieght; it may be different colored; it may be different size but, number wise 10 10 10. This is one type of Homoscedasticity issue or else like this it may be ten fifteen twenty but, color wise it is same.

So, that means it is just like your sending one market this 10; another market this 10 15 and another market this 20. But the color is the same so, this is how this can be homogenous in nature. So, homogenous can be various angle and heterogeneous can be under various angles like if you put the same thing in heterogeneous issue that means let us say there are 3 different products then this is 10 greens, 10 yellow and 10 red.

That means we have to pick up something 2 from here, 3 from here, 5 from here then it will go to one particular market then rest 3 3 6 7 will go to another market and the rest of the item will go to the another market. That means number wise it is Hetero then it is color wise it is Hetro, then market wise its classification is also hetero. This is how the a heterogeneous issue is all about. Scedasticity means it is spread of error variance; it is the spread of error variance; it is the spread of error variance.

So Heteroscedasticity is two aspects Hetero unequal Scedasticity means spread of error variance. Now if you will integrate these two then, obviously it will give you the signal of it will give you the signal of unequal error variance. It will give you unequal error variance that means we need sigma square U but, now it cannot operate. It will be it cannot u it will be like that sigma square U i that means i 1 2 3 4 like this. So, that means the structure we have created here sigma square 1, sigma square 2, sigma square 3 like this sigma square n. It is not equal, not equal, not equal, not equal. This is how the Heteroscedasticity all about.

The definition says that Heteroscedasticity means the spread of unequal variance in the system econometric modelling system. If the variance error, variances are unequal over each and every sample points then, obviously this is a series problem for econometric modelling that has to be taken care first before you use this model or before you can say assume that the model is best fitted.

It is use for policy use or you can use for forecasting use. So, first Heteroscedasticity has to be removed then you have to go for using forecasting policy use with this we can conclude this particular session. Next class, we will highlight the details setup, the causes, consequence detection and its solution. So with this we have to finish this lecture thank you very much have a nice day.