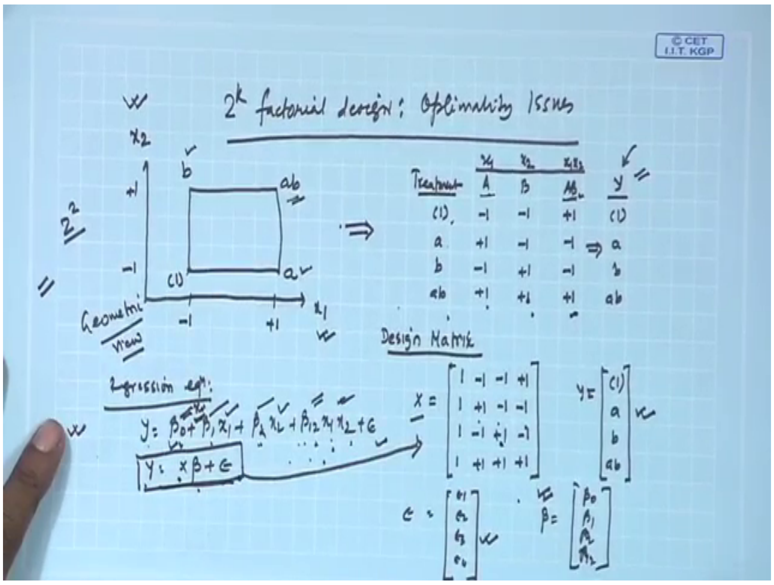
Design and Analysis of Experiments

Prof. Jhareswar Maiti
Department of Industrial and Systems Engineering
Indian Institute of Technology, Kharagpur

Lecture - 40
2 k Factorial Design Optimality Issues

Welcome. We will discuss optimality issues in 2 to the power k Factorial Design.

(Refer Slide Time: 00:25)



So, before going to the optimality issues, I will discuss with reference to 2 to the power 2 design that the regression surface as well as how the regression coefficients are estimated. You have already seen this one, but I want to repeat it. So, you see this page and I am 100 percent sure that all those terms written here is known to you.

The first one is talking about the geometric representation of 2 to the power 2 design. Here we have 2 factorials a and b, when u use the same thing in regression we denote them by x 1 and x 2. And x 1 has 2 labels, x 2 also has 2 labels; the labels are denoted by minus 1 and plus 1 as high and low respectively.

And when you conduct experiments keeping both the factors at low level or minus 1 minus 1 label, the total observations what you obtained with responds to with reference to the y, responds variable denoted by within bracket 1, when only at a at high level it will be represented by a only b or x 2 at high level, it will be represented by b and when both the factors will be high, it will be represented by y.

Y will be, total of y will be represented by ab. Now same thing when we use the geometric sign that is algebraic signs. So, the geometric view analogous representation in algebraic sign is this. What here we have 4 treatment combinations 1 a, b and ab. And x 1 or a, a is having the values minus 1, plus 1, minus 1, plus 1, x 2 minus 1, minus 1, plus 1, plus 1. And if you multiply both, you will get x 1x 2 that is AB, factor interaction A, B or x 1 and x 2. This would be plus 1, minus 1, minus 1, plus 1.

And the corresponding y total, exactly it depend combination, what the way we are basically following. We are writing in this fashion. You may right in other way, but this is what is the practice, then you see a left hand side here, I have given you the regression equation with that is basically first order regression equation with interactions, for this 2 to the power 2 factorial design.

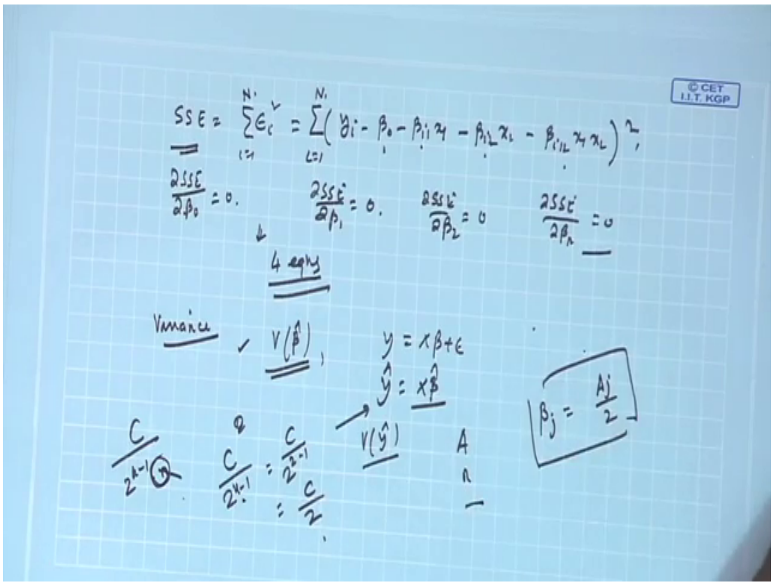
So, here y is equal to beta 0, beta 1 x 1, beta 2 x 2, beta 1 2 x 1 and x 2. So, beta 0 is the intercepta, beta 1 and beta 2 are the main effect, beta 1 2 is the interaction effect and error. So, the same can be written in matrix form; y equal to x beta plus epsilon where y is the responds values at different experimental runs, x is the design matrix, beta is the regression coefficient to be estimated, epsilon is the vector representing the error terms ok.

So, then will come to this side. Here what happens, in this particular equation with respect what are the different term like x. First here, there are 4 independent observations; independent observations in the sense am saying independent treatment settings. And either the average or the total at each independent observation, each independent treatment settings is considered here. We are considering here 2 terms, sometimes average also can be used.

Now the first column is actually contains all 4 1's, this is for beta 0; second, third and fourth column with reference to beta 1, beta 2 and beta 1 beta 2 x. That is for from estimation point of view otherwise I can write that these first column is beta 0. Here one x 0 is multiplied, who is taken all the time 1 and second, third, fourth column with respect to x 1, x 2 and x 1 x 2. So, this is known as design matrix.

This is the responds vector. This is the error vector and there is another vector called beta which is beta 0, beta 1, beta 2 and beta 1 2. This part is known to you. Now when we ask you that you estimate the parameters, what you will do.

(Refer Slide Time: 05:35)

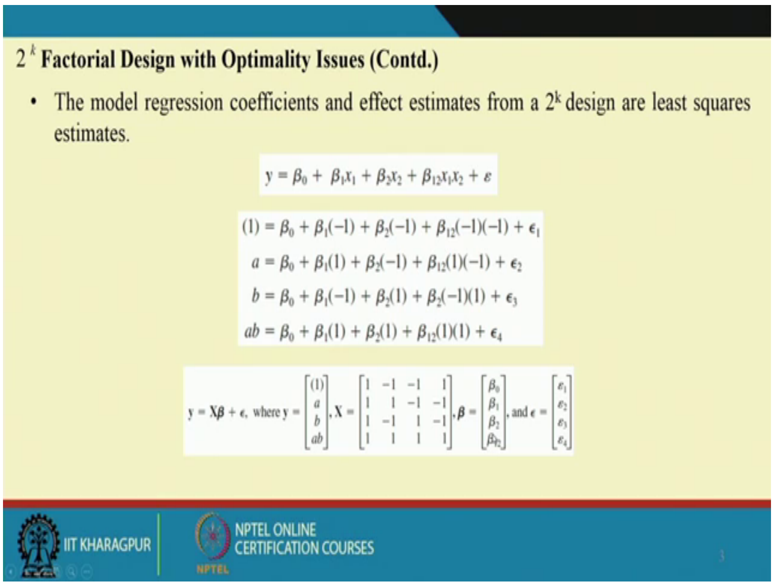


You will find out SSE which is epsilon i square, sum of i equal to 1 2, 1 2 here number of observations say, that may be let it be n. And then what we do, this will be nothing but y i minus beta 0 minus beta i 1 x 1 minus beta i 2 x 2 minus beta i 1 2 x 1 x 2, this square and obviously, I equal to 1 to n; number of observations. Then you will find out del SSE by del beta 0, put 0. del SSE by del beta 1, put 0 or del similarly del SSE by del beta 2 put 0, del SSE by del beta 1 2, put 0. And then get you how many equation you will get, 4. How many parameters, 1, 2, 3, 4; 4 equation you will get. Solving this 4 equations, you will be getting all the estimates ok.

So, this is what you will be discussing. Now content is basically the optimality issues with and when you are talking about optimality issues, it is related to the variance optimality, variance because you are what you are doing. You are estimating beta. So, what would be the best estimate that which is which is basically having minimum beta variance beta estimate. So, using this equation y equal to x beta plus epsilon; suppose that you can use it prediction purpose, y will be x beta cap. Now you want to minimize the variance of predicted portion. So similarly, there are different estimation issues here and for every such cases.

So, depending on the situation, you may, your design may help you to optimize the situations like the variance, minimize the variance for the beta estimate, minimize between prediction amount, this is what is known as optimality issues in 2 to the power k factorial integers as such this is true for any regression model. And this is true for all kind of regressive design, only the mathematics portion will change little bit depending on the design and the data available. So, these issues will be discussing here and again we are basically taking the information from this book that Montgomery book, chapter 6.

(Refer Slide Time: 08:45)

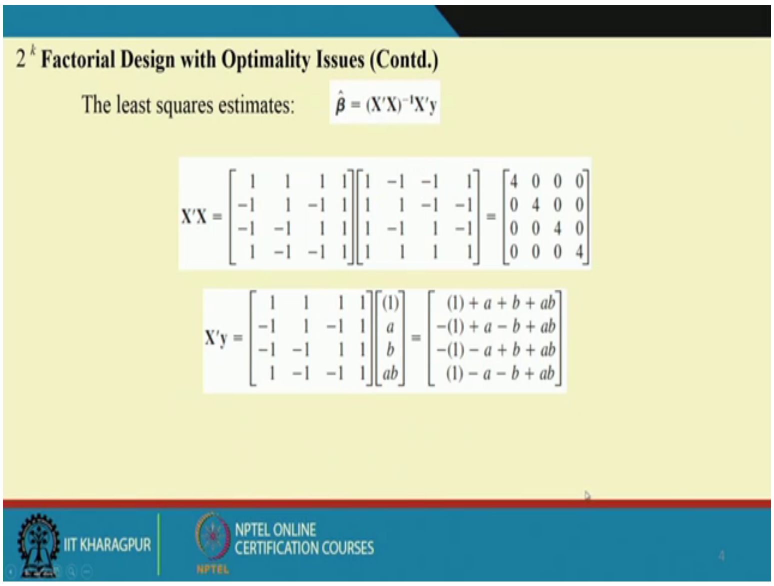


So, the explanation I have given you that, if you find out the SSE and then basically if you just go for SSE by beta 0, beta 1 all those things, you will actually estimate the beta values this part we have seen earlier. I have shown you earlier also.

So now, here what happen, in addition to that we are giving actually what I am giving you because as x 1 and x 2 takes only minus or plus 1 values. So, if you put these into these all those minus 1 and plus 1 and this then you are basically getting this matrix. What is the matrix, what is the matrix basically where matrix as am giving you here in all those things this matrix that beta 1, beta 2 all those things, this is the application.

So we have 4 y values, 1, a, b and ab. And x 1 column x 1 column is minus 1 plus 1 plus 1 see minus 1 plus 1 minus 1 plus 1, x 2 minus 1 minus 1 plus 1 plus 1. And this one, minus 1 minus 1 plus 1 minus 1 minus 1 plus 1 and these. So, if you if you just write down separately, what you will get? One is beta 0, minus beta 1, minus beta 2 plus beta 1 to plus epsilon 1 ok.

(Refer Slide Time: 10:29)



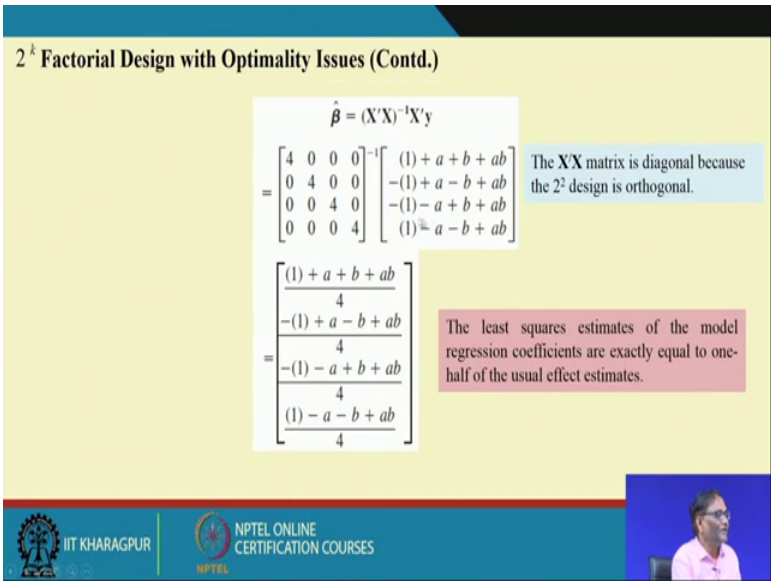
And then, the estimation point of view, this is the, this is what is the formula least square; what is the least square formula? X transpose for X inverse, X transpose y. So, if your data x is this, x is this one, then x stands for you will be these and then when you multiplied these 2, you will get a very interesting matrix here. You see that diagonal elements are 4 of diagonal elements are 0, getting me.

So here, you come back, come back to these you have 1, 2, 3, 4; 4 independent distinct observations here. Means at each experimental settings, you may get many observations, but your taking total or average of these and then you are making them as distinct observations. So, there are 4 because of this it is 2 to the power 2 designs; if it is for 2 to the power 3, it will be 8; 2 to the power 4, it will be 60 like this.

So, here this 4 things are, now if you want to include that no each observations also will be a that application should be there, then it will be 4 n, 4 n, 4 n, 4 n, 4 n, 4 n, it will they are like this. Now if you multiplied X transpose into y, another interesting one you are getting these. This also I have explained earlier.

So now what happened, you are basically getting the, this if you think that this; that means, the second, third and fourth is nothing but contrast for that is a, b and ab effects. So, left hand side that mean now, beta cap is multiplying these 2 will give you what. So, if you take the inverse of these, will become diagonal element will become 1 by 4 and half diagonal element will become will be will remain 0.

(Refer Slide Time: 12:38)



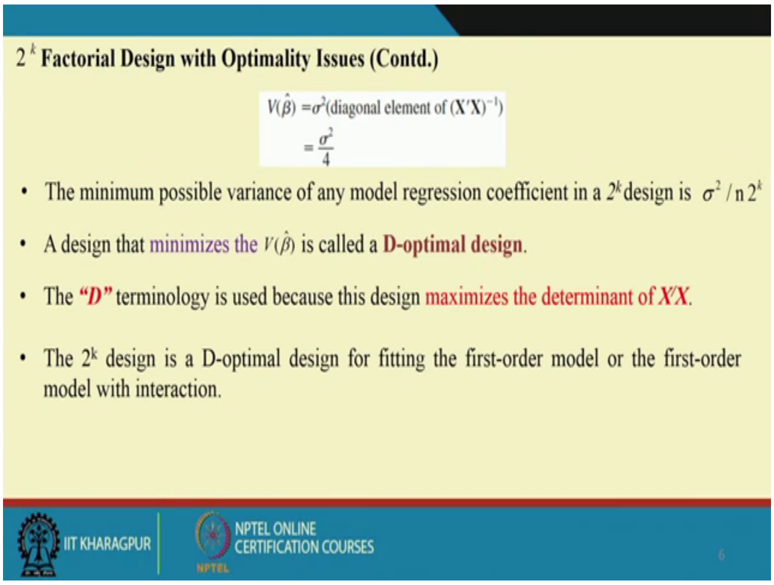
Then your calculation will be like these. So, all the contrast this will be divided by 4. So, the X transpose X matrix is diagonal because 2 to the power 2 design is orthogonal, this is the key things here we wanted to show you.

So, because of orthogonal design, you are getting half diagonal elements 0; there is no covariance, independency between the and here when the least square estimates of model regression come exactly equal to 1 half of usual effect estimates, that also we have seen earlier. If you estimate the effect, you will what you will write down contrast by 2 to the power k minus 1 into n, contrast by 2 to the power k minus 1 into n, that is sense you do.

So, 2 to the power now if, as we have taken total or average, this part will not come; so, it will be by 2 to the power k minus 1 and then contrast by 2 to the power 2 minus 1. So, it will be contrast by 2 that will be the effect of A, B like this. But here is the regression coefficient we have shown you that beta j regression coefficient and effect is A j and this are relationships earlier we have shown you ok.

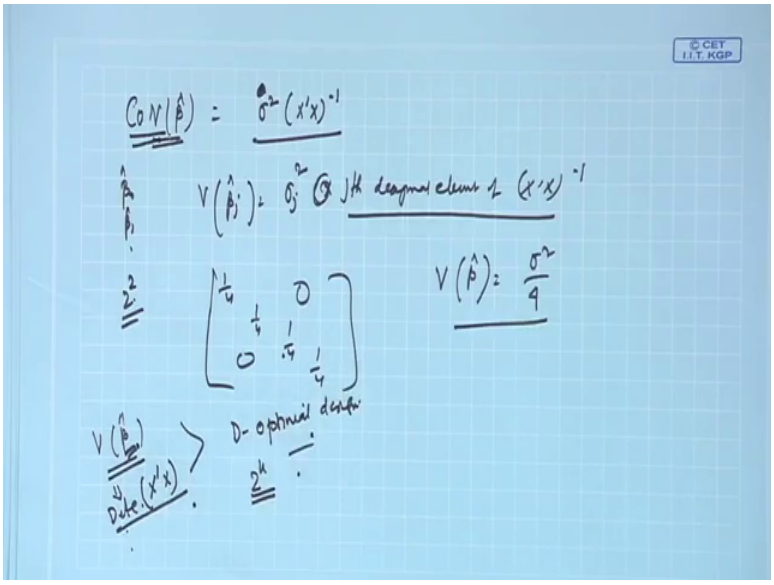
So, with reference to this x transpose x beta and the prediction part y cap, then we will we will explain now the optimality issues.

(Refer Slide Time: 14:18)



First of all, what will be the estimate of variance of beta?

(Refer Slide Time: 14:22)



If you recall the regression lectures, there we have shown you that variance of beta cap is sigma square, X transpose X inverse. Actually the estimated, if you know sigma square, you write down sigma square. Suppose I write this sigma square, this. Now that is because u have beta 0, beta 1 like this. But if I want to know the, this is other way we can, we have written as co variance. Now if I write variance of beta j cap, then this is sigma j square, the j th element of j th diagonal element of x transpose for x inverse.

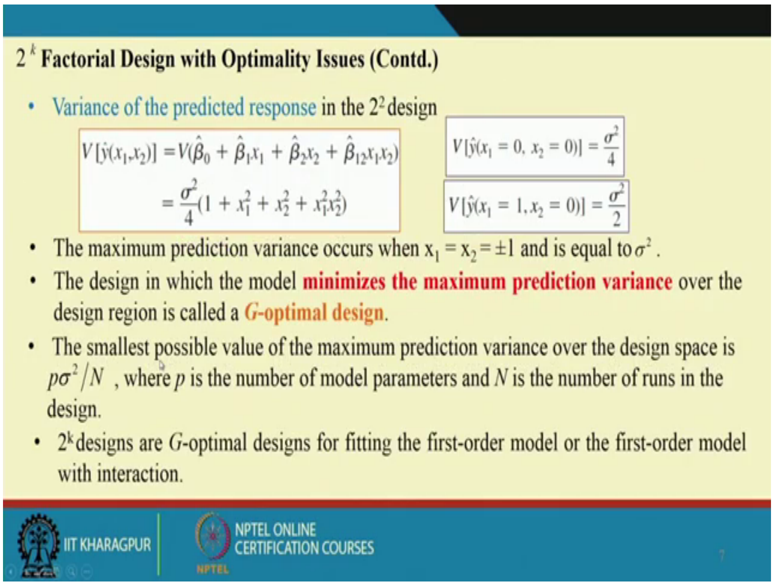
So, X transpose X inverse of 2 to the power 2 design is 1 by 4, 1 by 4, 1 by 4, 1 by 4 and other part is of diagonals are 0. So, that means, everything divided by 1 by 4. So, actually what happens then, so variance of any beta component, any beta cap will be sigma square by 4. Now, and this is what is the minimum variance, this is what is the minimum variance.

Now let us read out the points here, the minimum possible variance of any model regression coefficient in a 2 to the power k design is sigma square divided by 2 to the power k into n. Now a design that minimizes variance of beta cap is known as D-optimal design. So, this is it is mathematical issue, that D-optimal design, but from the practice point of view, but from our application point of view.

So, a D-optimal design is one which minimizes the variance of the estimated parameters here. Now the D terminology is used here because it maximizes the determinant of X transpose X. The D is coming from the determinant point.

So, that is why it is known as D-optimal design. And 2 to the power k design, the way we have explained, the actually that the, that is what the 2 to the power k design also. The 2 to the power k design is a D-optimal design and it for first order model and first order model with interactions.

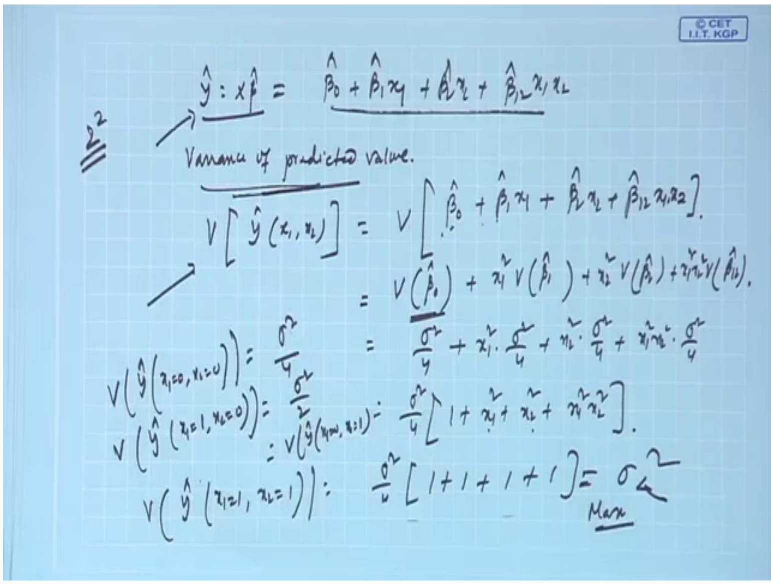
(Refer Slide Time: 17:26)



So, essentially what we have done, we have explained from optimality point of view that variance of beta, this would be minimized and it is done in 2 to the power k design using the least square estimates in the manner that it optimizes the determinant of these or maximizes these determinant of these and collectively these of design is known as D-optimal design.

So, what do I mean that D-optimal design is one which minimizes the variance of beta cap and 2 to the power k design is D-optimal design, 2 to the power k design with a first order a with interact and with interaction also it is D-optimal design. So, this is the first. Second we will see another variance.

(Refer Slide Time: 18:21).



Suppose, when we say y cap is x beta cap and many a times what happen, we will be interested to know the variance of predicted value. So that means variance of predicted value that may be variance of y, suppose x 1 and x 2. Here we are writing y cap x 1 and x 2 because only 2 to the power 2 design, we are x we are taken for expandation purpose. So, then what will happen, what will happen to this this will be beta 0 cap plus beta 1 x 1 beta 1 cap x 1 beta 2 cap x 2 beta 1 2 cap x 1 and x 2, this is the predicted part.

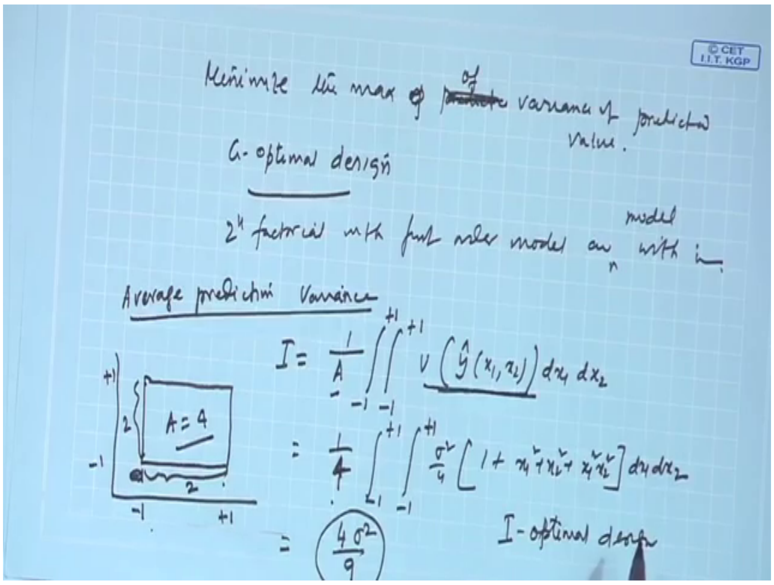
So that means, you are interested to know the variance of beta 0 cap plus variance of, these plus these plus beta 2 variance 2 plus beta 1 2 cap x 1 and x 2. Now x 1 x 2, all those things are fixed values because we have seen the design matrix. So, can it not be written that this is, now again what I mean to say that this beta 1 be the 0 this orthogonal design. So, that we have seen the co variance part of beta 0 is 0. So, because of independence, we can write that this is beta, variance of beta 0 cap, then x 1 square variance of beta 1 cap plus x 2 square variance of beta 2 cap plus x 1 square x 2 square variance of beta 1 2 cap.

So, what is variance of beta 0 cap? That is sigma square by 4. We have seen earlier then x 1 square into sigma square by 4, x 2 square into sigma square by 4 plus x 1 square x 2 square into sigma square by 4. So, it is sigma square by 4 into 1 plus x 1 square plus x 2 square plus x 1 square x 2 square. So, this is variance of predicted response, this is known as variance of predicted response. Suppose a, if all x 1, x 2, 0 0 what will happen. So, these value will be, suppose the variance of y cap given x 1 equal to 0, x 2 equal to 0, this value is sigma square by 4.

If variance of y cap x 1 equal to 1, x 2 equals to 0. So, these will be. So, these, what will be happen, this one will be sigma square by 2 or other way this equal to variance of y cap, x 1 equal to 0 and x 2 equal to 1. If we put variance y cap x 1 equal to 1 and x 2 equal to 1, what will happen this one, sigma square by 4 into 1 plus 1 plus 1 plus 1. So, this will become sigma square.

So, this is what is the maximum variance that can occur in the in the response when you predict because when all x 1 and x 2 both are the high level 1, x 1, 1; x 2, 1. So, that in that the corner point basically where the, everything is the square term whether the x 1 or plus 1 or minus 1, this occur in a maximum variance occur. So, what we require then, we require a design that will minimizes the maximum of predicted variance.

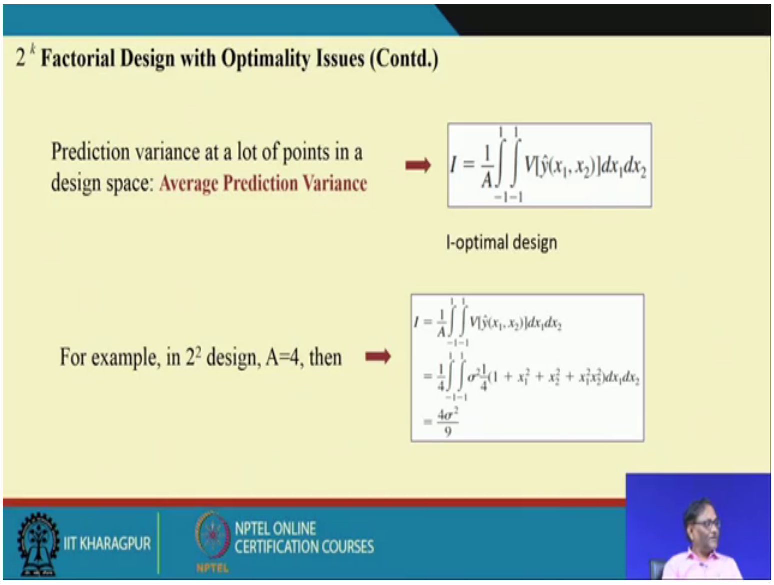
(Refer Slide Time: 22:43).



So, we you want a design that minimize, minimize the maximum of predicted, maximum variance, maximum of variance of predicted value. So, minimize that many maximum variance for the predicted response will occur at the corner points and you want a design you have to minimize this one, these design is known as G-optimal design. And 2 to the power k design, factorial design is G-optimal with first order and model first order model and with interaction and model with first order model with interaction are G-optimal.

So, this is the second issue, second optimality issue here. Now you read out the this point third point, here the last but one point, the smallest possible value of the maximum prediction variance over the design space is p sigma square by n, where p is the number of model parameters and n is the number of runs in the design. This is what is the smallest possible value of this maximum prediction variance, 2 to the power k designs are G-optimal designs for fitting the first order model and first order model with interactions.

(Refer Slide Time: 24:36)



Suppose, we have lot of points so, you want to see that, what will be the predicted variance for lot of points. You are not basically predicting at one point. You have lot of several points, so that means, what will happen; the you will find out different kind of variance at different predicted variance at different points. So, under such situation what happen, you will basically find out the average prediction variance.

So, we were talking about Average prediction. So, this average prediction variance is denoted by, you just first find out the total area. Here we have the, what is the design space x 1 minus 1 to plus 1 and x 2 minus 1 to plus 1. So, within this design space, what is my variance of this one that predicted one, x 1 x 2. So, you integrate this, d x 1 and d x 2.

This is basically total you are getting. Now once you divide it by the area, 1 by a. So, then you are getting the average prediction variance, this is the formula. So now, let us think a 2 to the power 2 factorial design case, what this is. So, this is minus 1 sorry, let me write minus 1 plus 1 and then minus 1 into plus 1, this is 2. Similarly, minus 1 plus 1, what is the length; this is 2. So, what is the area 2 into 2, 4.

So, I will write these then, 1 by a, here minus 1 to plus 1 and then minus 1 to plus 1 and then what is this value? You have already seen what is the predicted variance of this? I think variance of this is you see, we have written variance of this equal to sigma square by 4 into this. So, you write down this one as sigma square by 4 and this is 1 by 4, I am writing, a equal to 4, 1 by 4 sigma square by 4 into 1 plus x 1 square plus x 2 square plus x 1 square into x 2 square, d x 1 and d x 2. So, first you do it with reference to d x 1 then, put the values and x 2 put the values, then the result and value will be 4 sigma square by 9.

So, in 2 to the power 2 design, average prediction variance is 4 sigma square by 9. Now what do you want? We want also, this to be minimized the minimization of average prediction variance. So, when you say lot of points to, in the design space, you are predicting the values at lot of in the design space lot of values. So, this is known as I-optimal design, I-optimal design.

So, these are the few things you learnt that, your design is when you develop regression surface, that time how you use regression equation that is you have learnt earlier. Today, we have repeated this thing, then we say that when you are estimating some of the parameters, the variance of those parameters to be minimum and that was that is in mathematic known as D-optimal design. So, we have seen that 2 to the power k design is D-optimal for the first order and first order model with interactions and then we have we have shown you that G-optimality.

So, G-optimal design is one which minimizes the maximum prediction variance and we have seen that our, 2 to the power k design is G-optimal and finally, I-optimal design. And here, in the case of 2 to the power of k design, the I-optimal value is just 4 sigma square by a this is I-optimal also. So, this is little bit advance topic for that optimality issue, that optimality issue will be dealing further in detail in response surface when we discuss response surface methodology, that time this optimality again come and we will discuss to some extent there also ok.

Thank you very much, hope that you have enjoyed this class also.