Course material

Course:

ENG606 / PHYS 442

Video:

DOE_lesson3_part2_DOEbasics

Concepts (extracted from automatically generated subtitles):

Matrix of experiments. Variance covariance matrix. Prime of transpose. Model matrix. Name variance covariance matrix. Vector of coefficients. Confidence interval. Value of the first factor. Second of the coefficient. Different models. Variance inflation factor. Dispersion matrix. E prime. Ideal model x. Errors of measurement.

to video sequence search
(within ENG606 / PHYS 442.)

to video

...

notes

summary

0m 0s

2.2.1 Matrix of experiments

$N_{exp} \times N_{fact}$ matrix whose components are the levels x_{ij} of the factors for each experiment :

$$E = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1N_{fact}} \\ x_{21} & x_{22} & \cdots & \vdots \\ \vdots & & & \vdots \\ x_{N_{exp}1} & x_{N_{exp}2} & \cdots & x_{N_{exp}N_{fact}} \end{bmatrix}$$

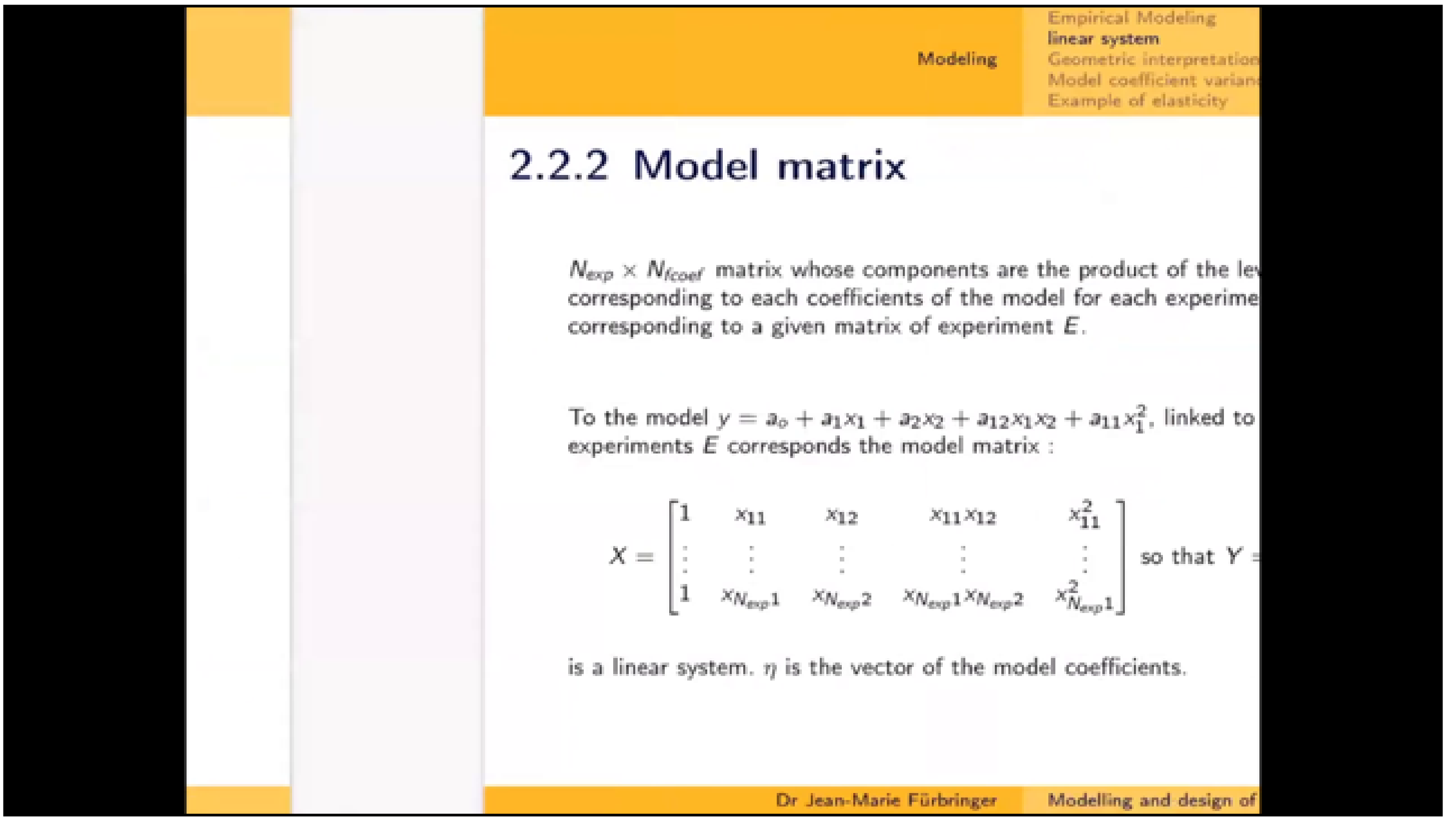
The component can be standardized or not

These subtitles have been generated automatically We'll see very important elements.

notes

summary

0m 1s

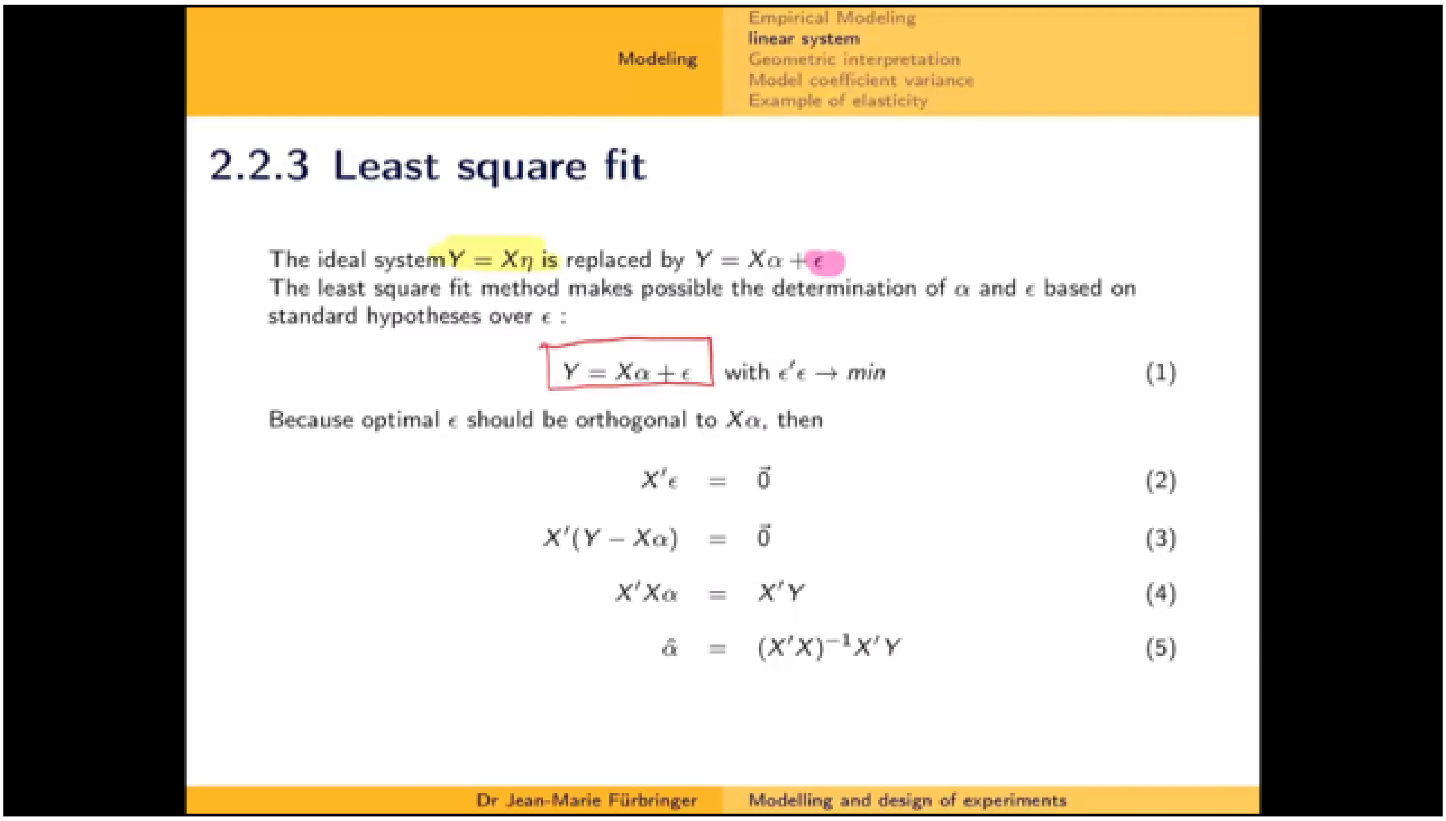


Let's say the first element of the theory. It's called a matrix of experiments. And usually when I'm okay, I call it E for experiments, being E. and we have as many rows as we have experiment. It's your design of experiment. It's what you have to use in the laboratory. It's what you have to give to your lab or on the top. And so because you say that at the first experiment, the value of the first factor is this x. So now I have two indices, x11. So it's the first value of the first factor in the first experiment. We will standardize things in a while. But still we talk about the matrix of experiments. If it's standardized or not, standardized means that you are not standardized, means that you are using the value of the laboratory, meters, velocity, time, temperature, what you really have. And standardized is the possibility of putting those values around an average value and manipulating the standard deviation of them. But it's called the matrix of experiments. And now everything will be built. So in the in the dolls is a small doll that's at the center.

notes

summary

0m 5s

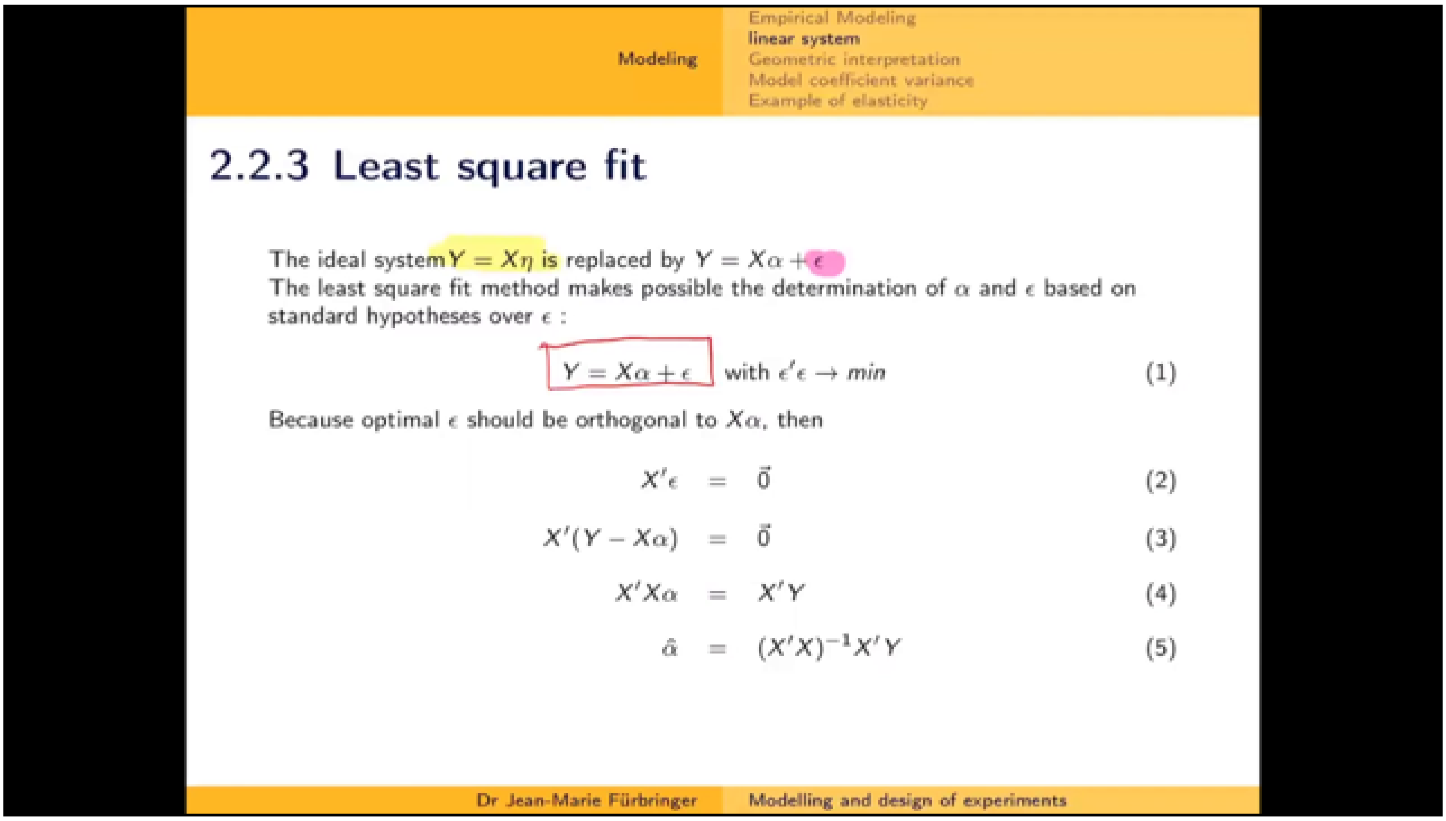


After when we have a model, we have a model matrix. So if we have different models, we have from one model of experiment, we can have different model of matrix. So the model of matrix is built in relation with the model. So if the model have a constant plus main effect plus an interaction effect plus one second of the coefficient, that means that my model matrix will have one column for the constant. So this is related to that. And I put a one because I will not change the constant will be present in all the experiment not changing is because it's a constant is the fact is a constant. I have one linear terms. So I will have one columns with the first factor, the value of the first factor. After I will I have a column for the second main effect. So I will have one column in my model matrix was the first factor after it became interesting. So yes, yeah, at least somebody is listening to me. Thank you. And so now I have an interaction. No, another color. This one. So now I have an interaction factor. So I will have one column. This is a new column in consideration with my my matrix of experiment, which will be each time the products of x1 and x2. And when I have quadratic terms, let's take a little blue. So I have a quadratic terms. So I have one column, which is the square of my, my first factor. So I have one column per coefficient in my model. And I have one role for exposure for the interaction factors. Is there a combination? Does it always have to be multiplication? Yes. We could go higher in more. We could have a cubic contraction with the cube. The square of the first one and the linear part of the second one. After

notes

summary

2m 36s



you can go, but it's with multiplication within this more. So I can represent this. So my answer, why will be equal by the, this model of matrix. Multiply by the coefficient. We can put the coefficients. A zero, a one, a two, a one, two, a one, one. I could put them in evidence and build a linear. Linear system. And in this case, as is the ideal linear systems, I use the, the Greek letter. You will see the reality will not be like that because when you are making measurement, we are making some errors. So we have to add this problem, but now is the perfect system, the ideal system. It's a linear system. So it's why we say that all those models are linear. So I'm not talking all the time correctly, but we should differentiate the linear response systems. That means that all our factor has linear responses and we do not consider interactions. Do not confront with a linear model, which could be anything that you can represent with some matrix of the model, multiplied by a vector of coefficients. So now, if we would like to estimate those coefficients. We use the least square feet. It exists all the time of regression, but we use least square feet and all the theory is made for that. So, as I said, the, the ideal system will be why that multiply a matrix of the model that multiply a vector of coefficients. In the reality, as we have errors of measurement, the real system we are confronted with is this one. It's, it's better because if we do not have this possibility of accepting error, you can have a system that doesn't work or have a rank which is not sufficient for being sold. So we really need that. We have this system. And we want to solve it. We want to solve it

notes

summary

2.2.3 Least square fit

The ideal system $Y = X\eta$ is replaced by $Y = X\alpha + \epsilon$
The least square fit method makes possible the determination of α and ϵ based on standard hypotheses over ϵ :

$$\boxed{Y = X\alpha + \epsilon} \quad \text{with } \epsilon'\epsilon \rightarrow min \qquad (1)$$

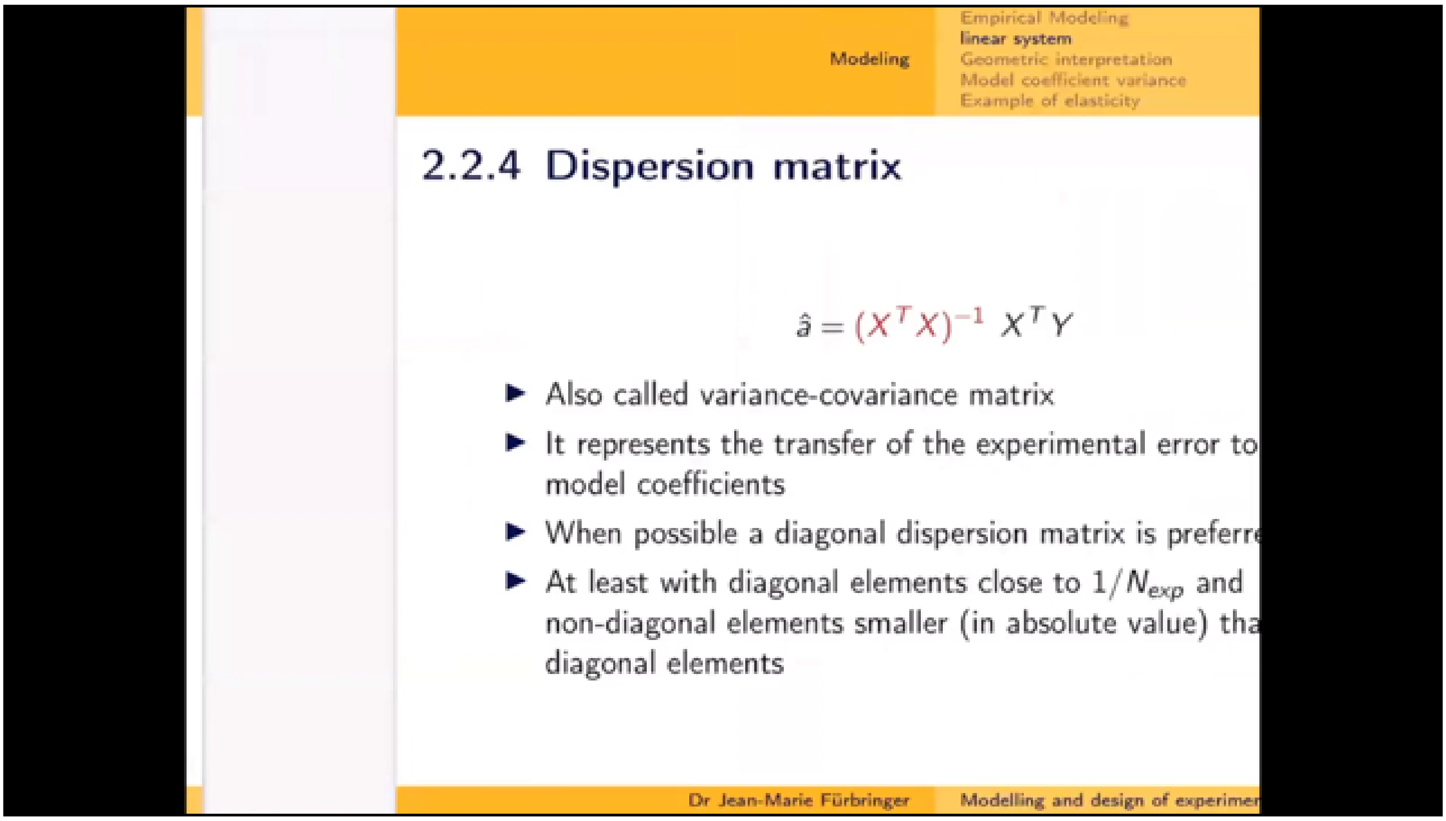
Because optimal ϵ should be orthogonal to $X\alpha$, then

$$\begin{aligned} X'\epsilon &= \vec{0} && (2) \\ X'(Y - X\alpha) &= \vec{0} && (3) \\ X'X\alpha &= X'Y && (4) \\ \hat{\alpha} &= (X'X)^{-1}X'Y && (5) \end{aligned}$$

for the minimum error. So the way of doing it, it said that we want to minimize e prime, it's transpose the prime of transpose transpose. So e transpose by e, this would be the sum of the square of the error. So we use e epsilon is a vector. So if you multiply e transpose by e, that means that to make the sum of the square of the terms of. The idea is that the model. Multiply by the error must be orthogonal one to the other. So you can write x prime by equal the vector zero.

notes

summary

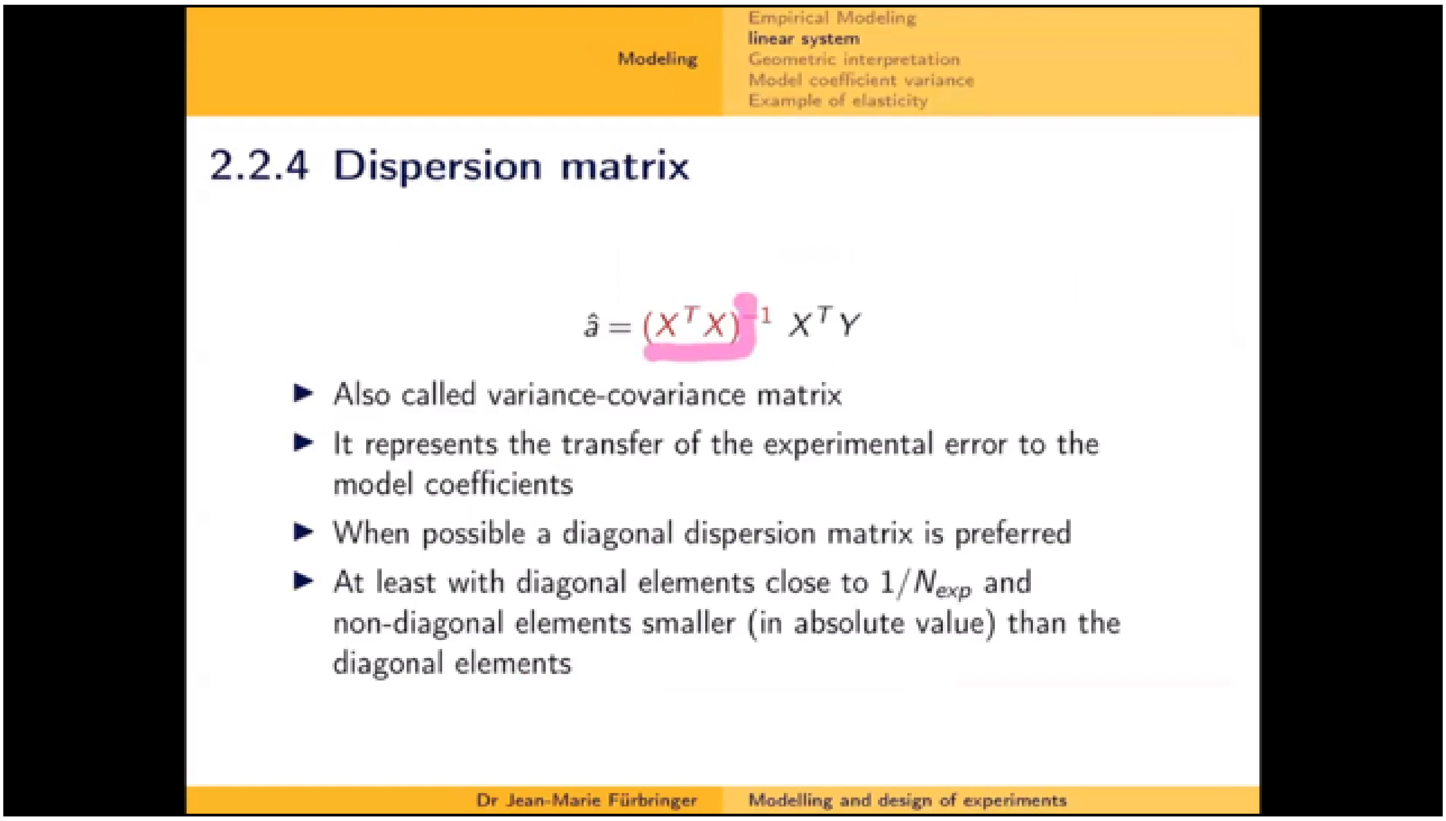


That's epsilon is the error of the model. So it's the difference between why the answers of the experiment, minus the ideal model x multiplied by alpha, we use alpha in this case we forget the effort. So, in that x supplies x prime is x transpose multiplied by x, multiplied by alpha must be equal by X prime, multiply by why so we can have now an estimate of the coefficients. So the hat means that it's a estimate. We make the difference between what we should have ideally and what we will get, accepting some errors will be. And this is a formula you will never forget it for the rest of your life. The estimator of your coefficients. So alpha 0, alpha 1, alpha 2. Yes. X transpose X minus 1 multiplied by X transpose multiplied by Y. So this is your coefficient of your model. The X is your experiment. You decide where you have making your experiment and the model you would like to fit. And Y is your experiment. And so from now, never forget this is the estimator of the Linsquirt mass. X transpose X minus 1 multiplied by X transpose multiplied by Y. Very, very simple calculation. You can even make it very easily in an Excel. It's really simple. And this is the Linsquirt. Matricially explained.

notes

summary

10m 1s

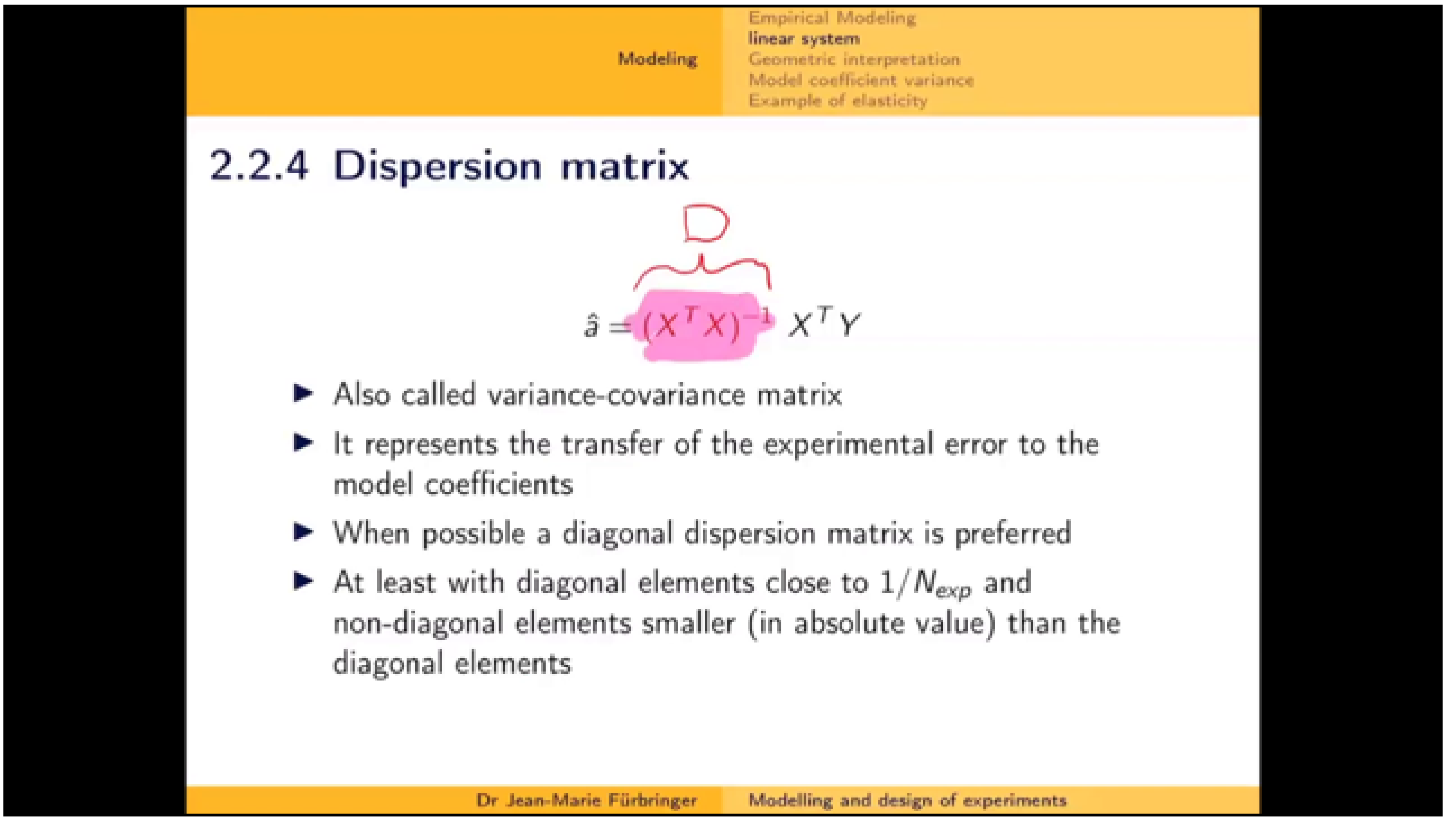


After, we can say something about the X transpose X minus 1. We call this matrix, which is in red in this slide.

notes

summary

12m 28s

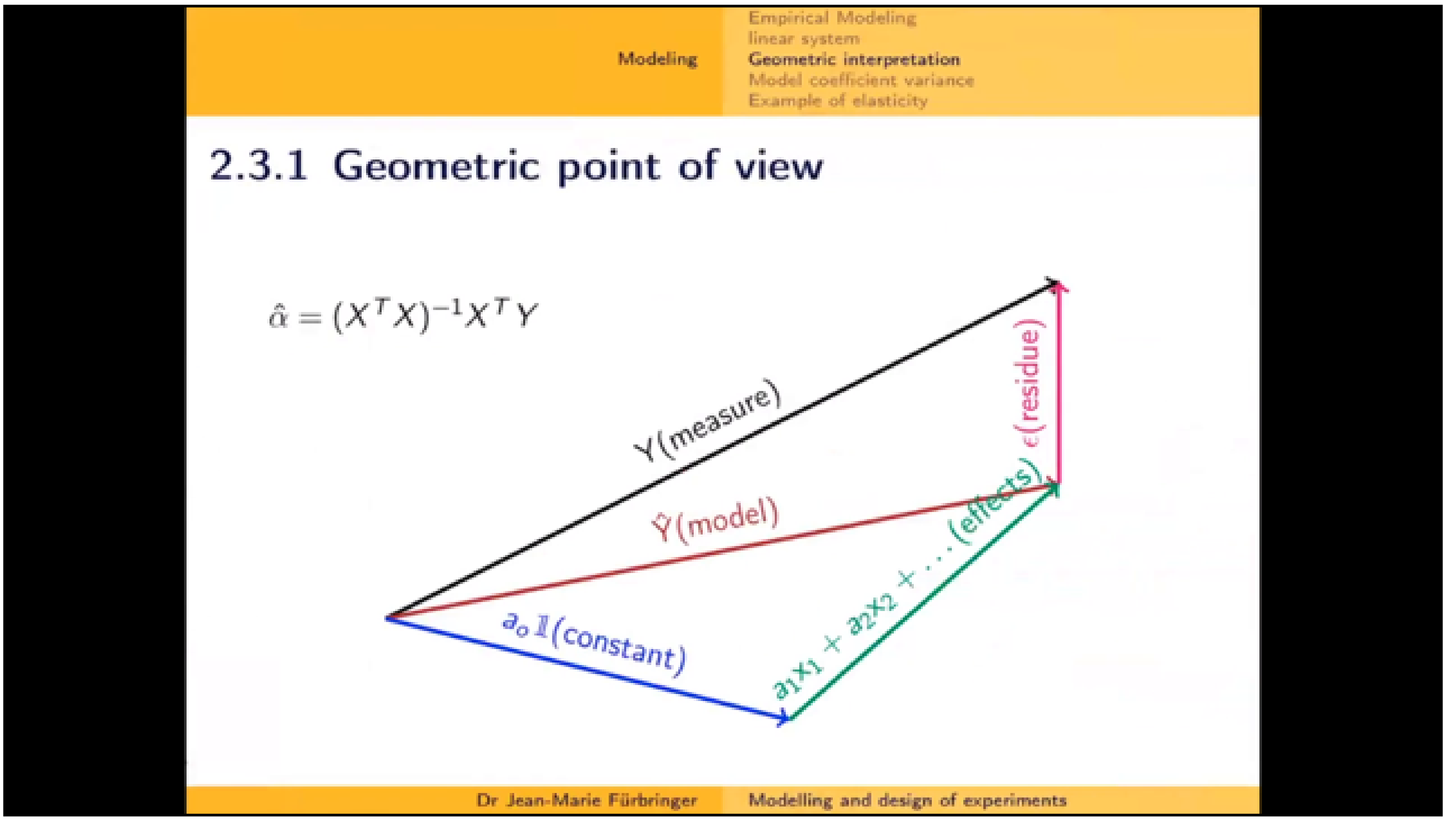


It's called the dispersion matrix. Understand this word, dispersion. So it's really the matrix, which is dispersing the information. It's the matrix which explains how uncertainty is traveling through your process of estimation. So it's very important. And optimizing the design is, in fact, optimizing this matrix. Sometime I give it the name of D for dispersion. So this would be the matrix D for dispersion. Analyzing a design is checking this matrix. If it's diagonal, you are very happy. That means that you do not have covariance between your factors. That means that the information is flowing well from the experiments to your model. When you have interactions, it could be high. Not interactions, covariance, it could be high or not. And after, you would like to have small value, small eigenvalue, because it will make that your uncertainty will diminish when you are estimating and not increasing what could happen. We have good estimator. So this matrix is also called the variance covariance matrix. Be careful. They are under the name variance covariance matrix. It exists two types of matrix. It represents the transfer of the experimental error to the model coefficient. When possible, the diagonal dispersion matrix is preferred, at least with diagonal element close to 1 divided by the numbers of experiments and non-diagonal smaller than the diagonal element. So this is the best one. Diagonal 1 divided by the numbers of experiments, this is perfect. Sometimes cannot. So first diagonal would be of first quality.

notes

summary

12m 47s

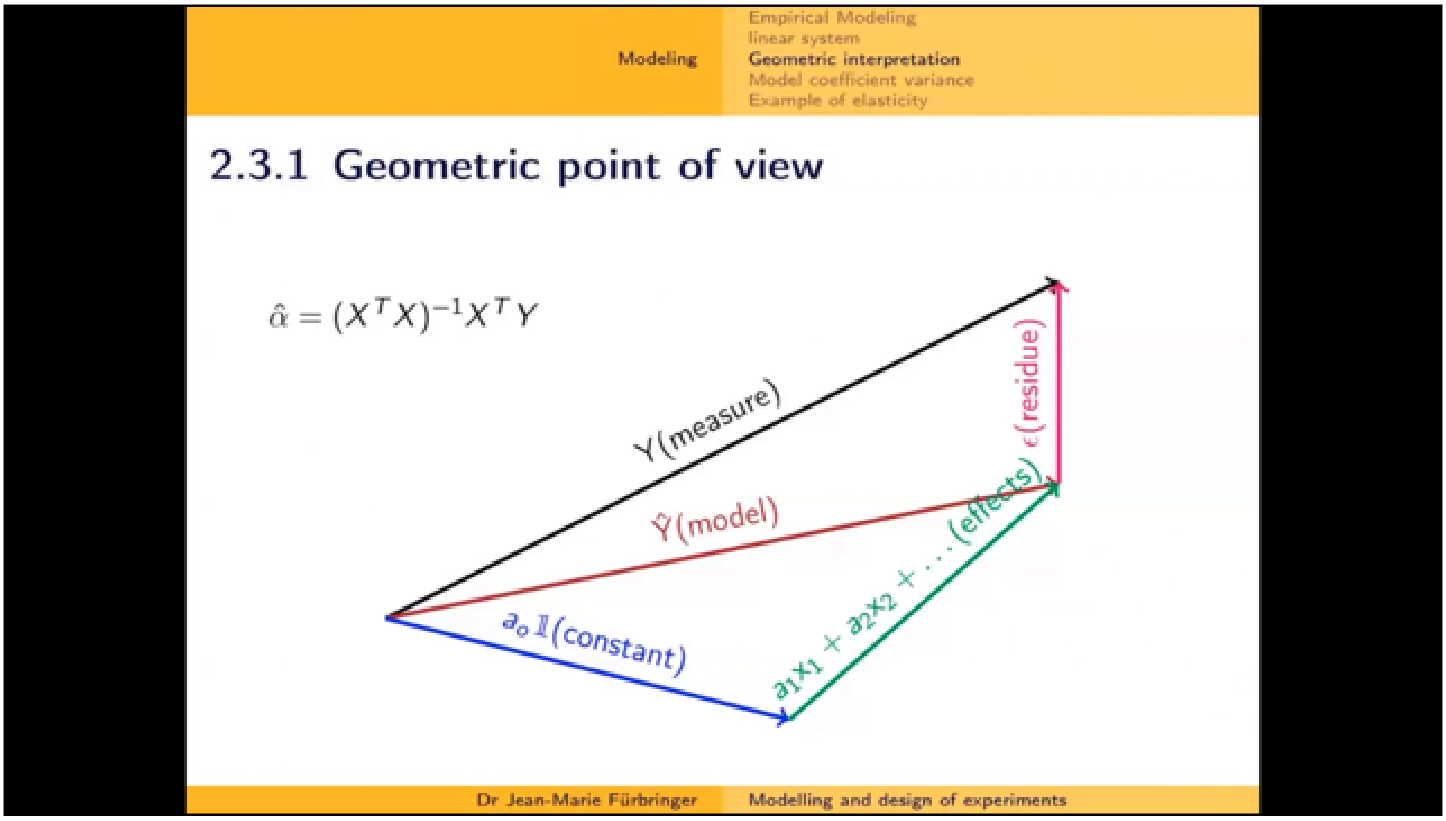


And after having interaction covariance, not so important. Dispersion matrix. So you see, we already see three elements. Matrix of the model, no, matrix of experiment, matrix of the model, and no, dispersion matrix. And you see, it's all come one within the other.

notes

summary

15m 1s

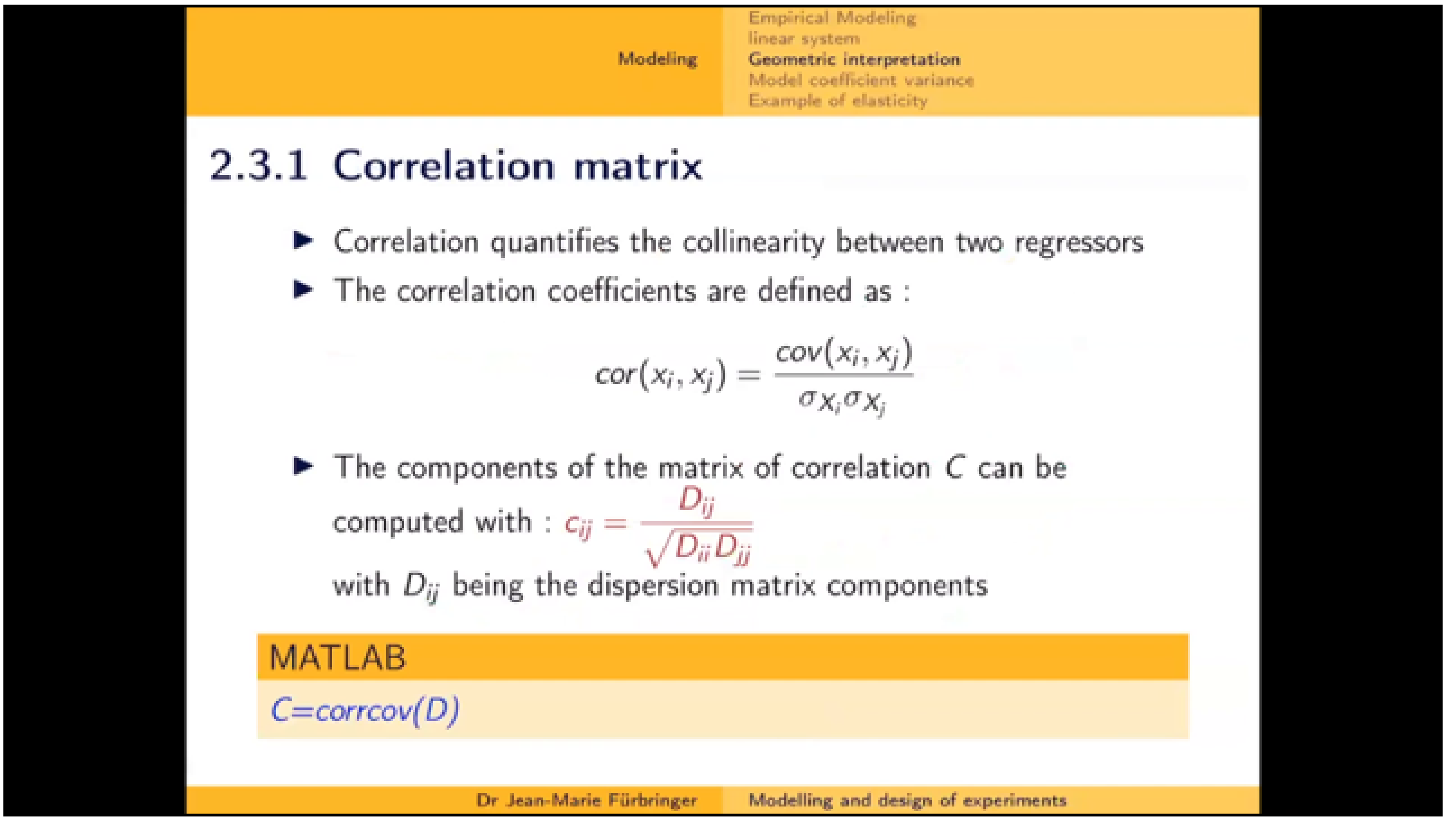


So now let's understand what we are doing, because what is a matrix multiplication? I don't know if it's for you, it's clear. Geometrically, a matrix multiplication, is that what you are projecting the vector representing by one of the matrix, the second one, on the vector of the other. So they are projection, because projection is the product of two vectors.

notes

summary

15m 19s

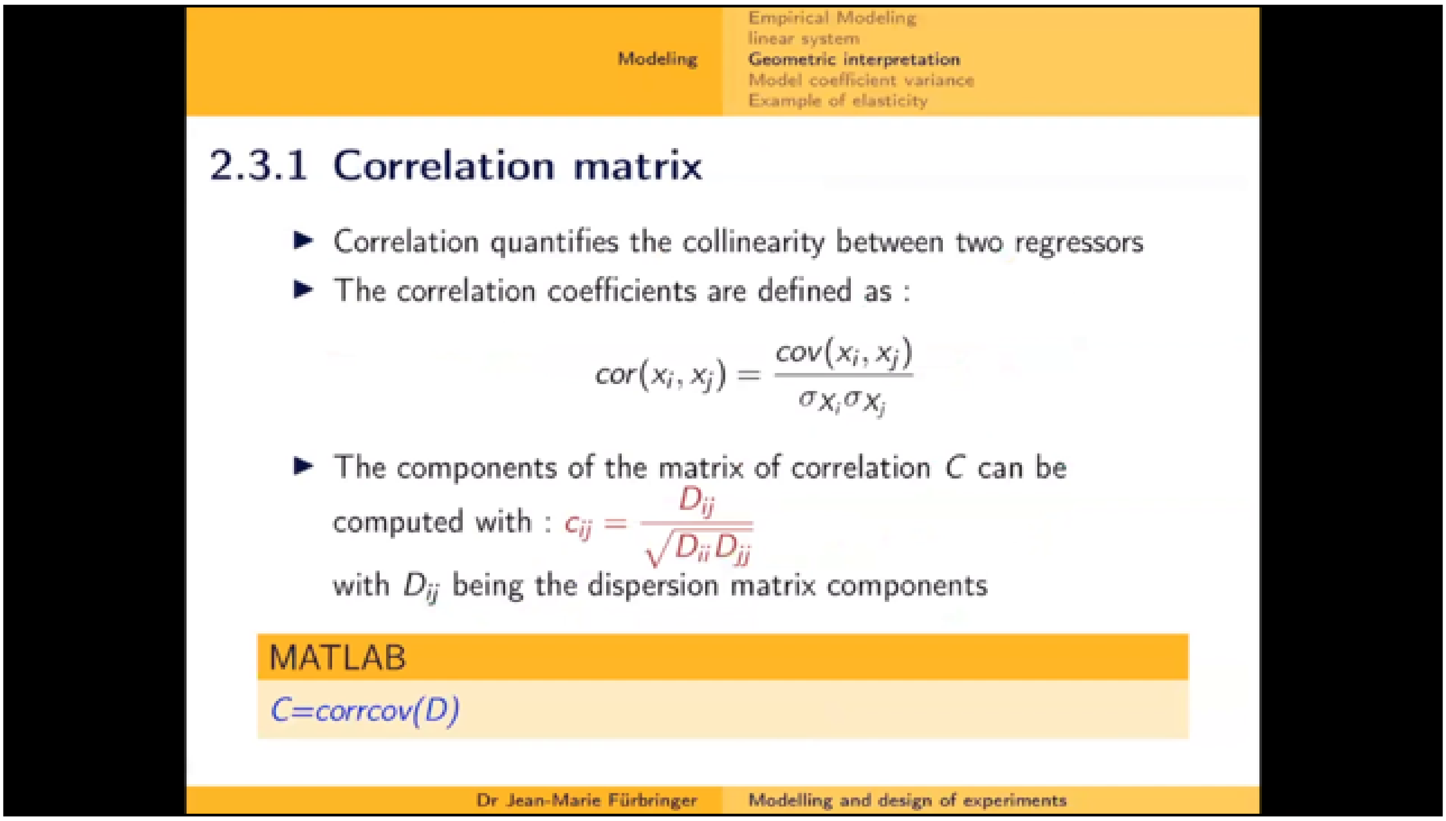


The matrix show might be, that is the name of the argument, the scalar multiplication. So the scalar multiplication of two vectors is in fact the projection of one on the other. And so a matrix multiplication is the series of scalar multiplication. So in fact, when you are multiplying two matrix between them, you are making a projection. So look at this slide. It's represented, so it's in three dimension. So you have your y represents my measurement. And my measurement, I wanted to, from my measurement, I would like to make a model. And for making a model, I'm calculating a constant. And a constant is a vector, which direction is clear, because it's a vector, which each component is the same. It's the value of the average. So it's a vector, which is crossing your space from the low corner to the high other corner. So its direction is clear. And after, depending on where you have making your experiments, you are orientating, you are projecting your black vector on different greens. So I just draw one green vector for all the effect. In fact, I have one for each of the effects that I can do. And your model is the sum of all your small vectors. So it's now here in the red vector. This is a model. And what you have, by definition, you have a right angle here. This is what is the modeling of a situation. You have a vector, not in three dimension. In fact, it could be in 10, 100, 1,000 dimension. Depending the numbers of data that you have. And you are projecting this vector on subspaces. One is very clear is the average, which is this vector crossing your space from the low corner to the high corner. And after, depending where you have making your measurements, so the vectors on which you project correspond to the columns

notes

summary

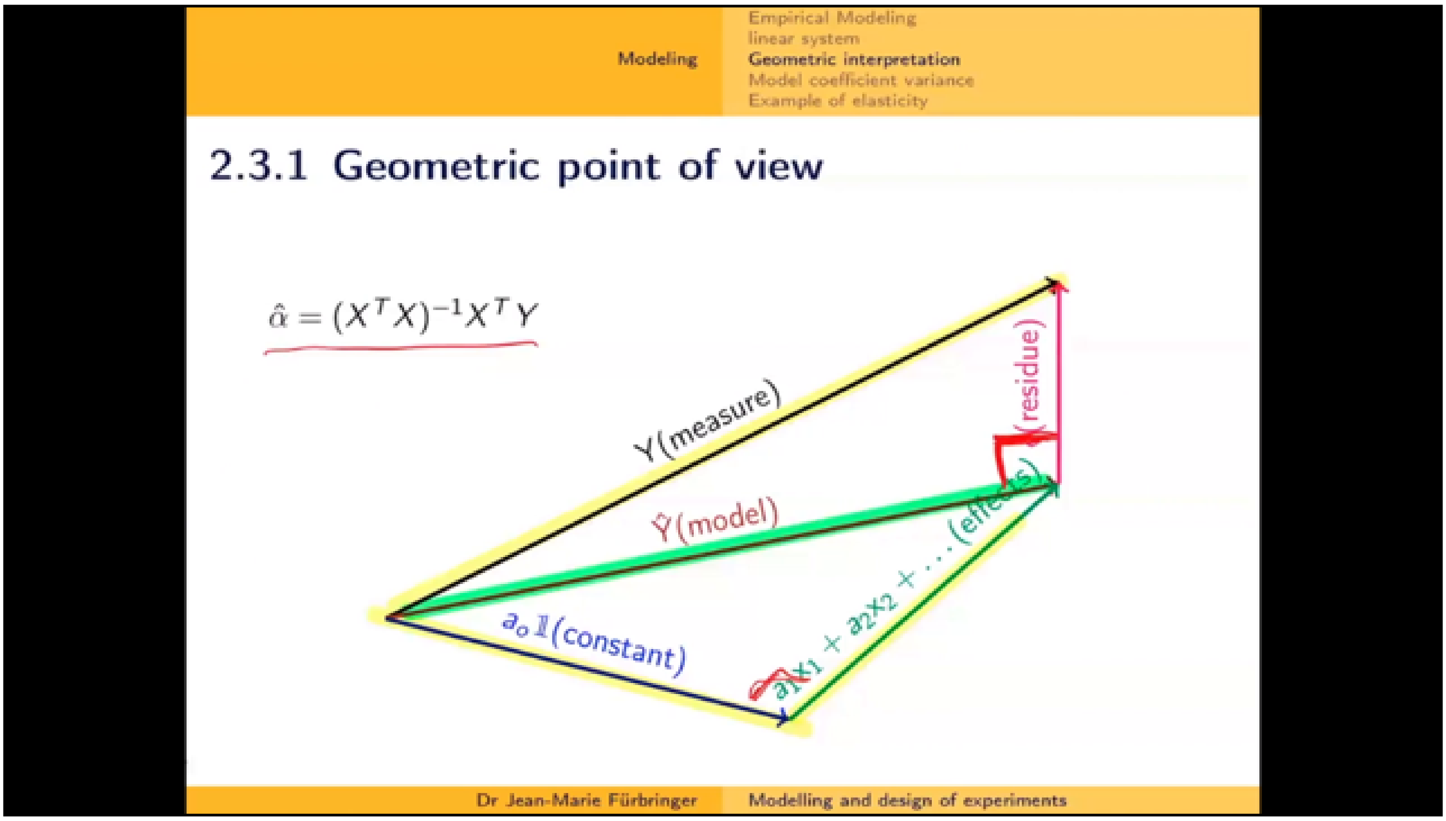
15m 49s



of your matrix of the model. In fact, the matrix of the model correspond to the subspace on which you would like to project your data. So you have the equivalence between this estimator, alpha hat x transpose x minus 1 x transpose y. It's exactly the same thing of my drawing. Taking a vector in the dimension of the space is the numbers of experiments that you have. And you are projecting your vector. So you understand that something which is all the times OK, it is the right angle here. After those angles are right angles or not. If they are right angles, it's make a diagonal matrix. That means that the vector that constitutes the model matrix are orthogonal one to the other. And so it's perfect because you are making a model with vectors at right angle. When it's not, most of the case, well, you have to consider that after. That means that you are sharing information between factor. You will not really guess everything for your model. There are some covariance. This is what I call the geometric point of view. I will come again and again on this slide. And this is one of the questions of the exam.

notes

summary

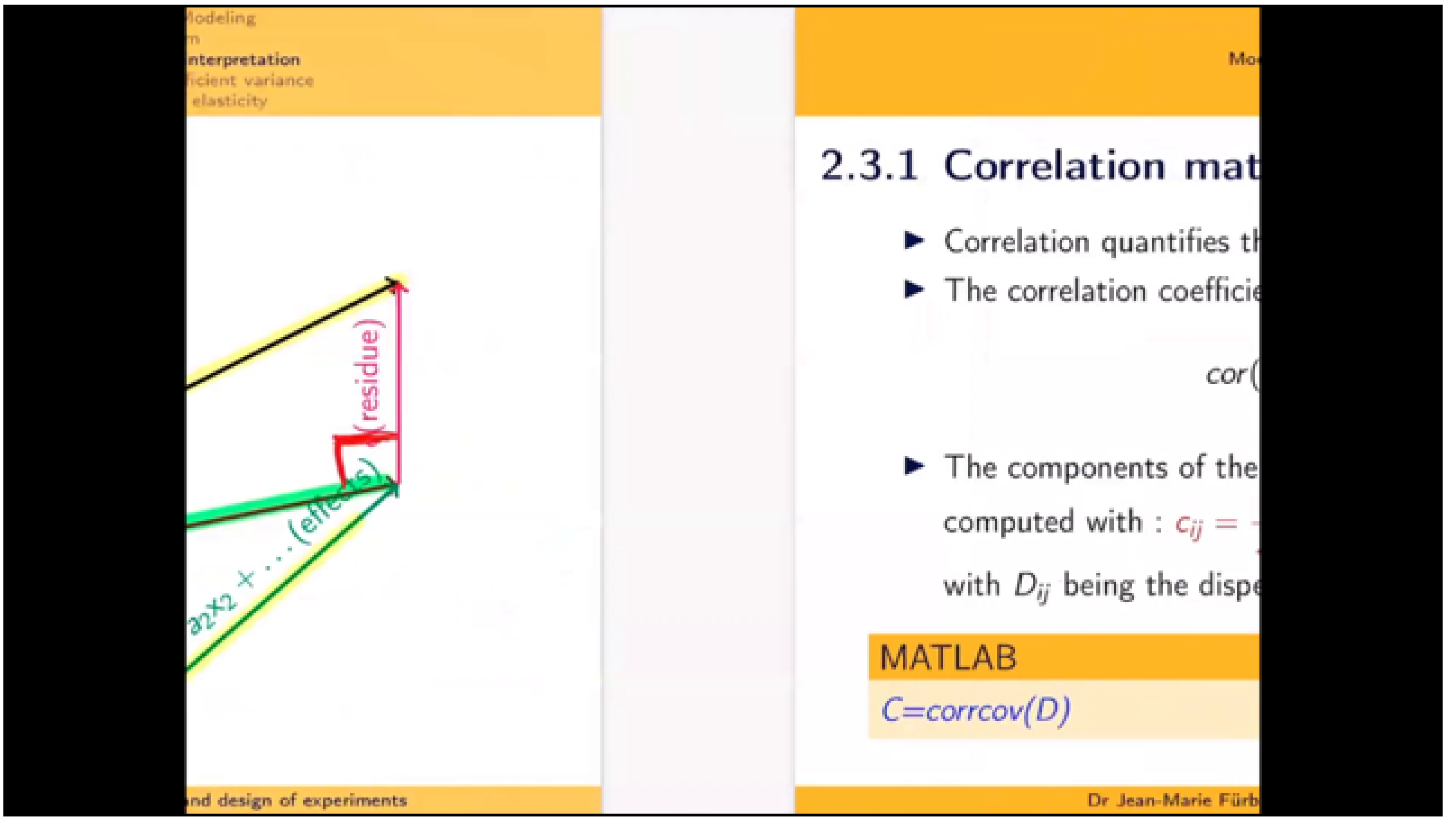


Yes, please.

notes

summary

20m 13s

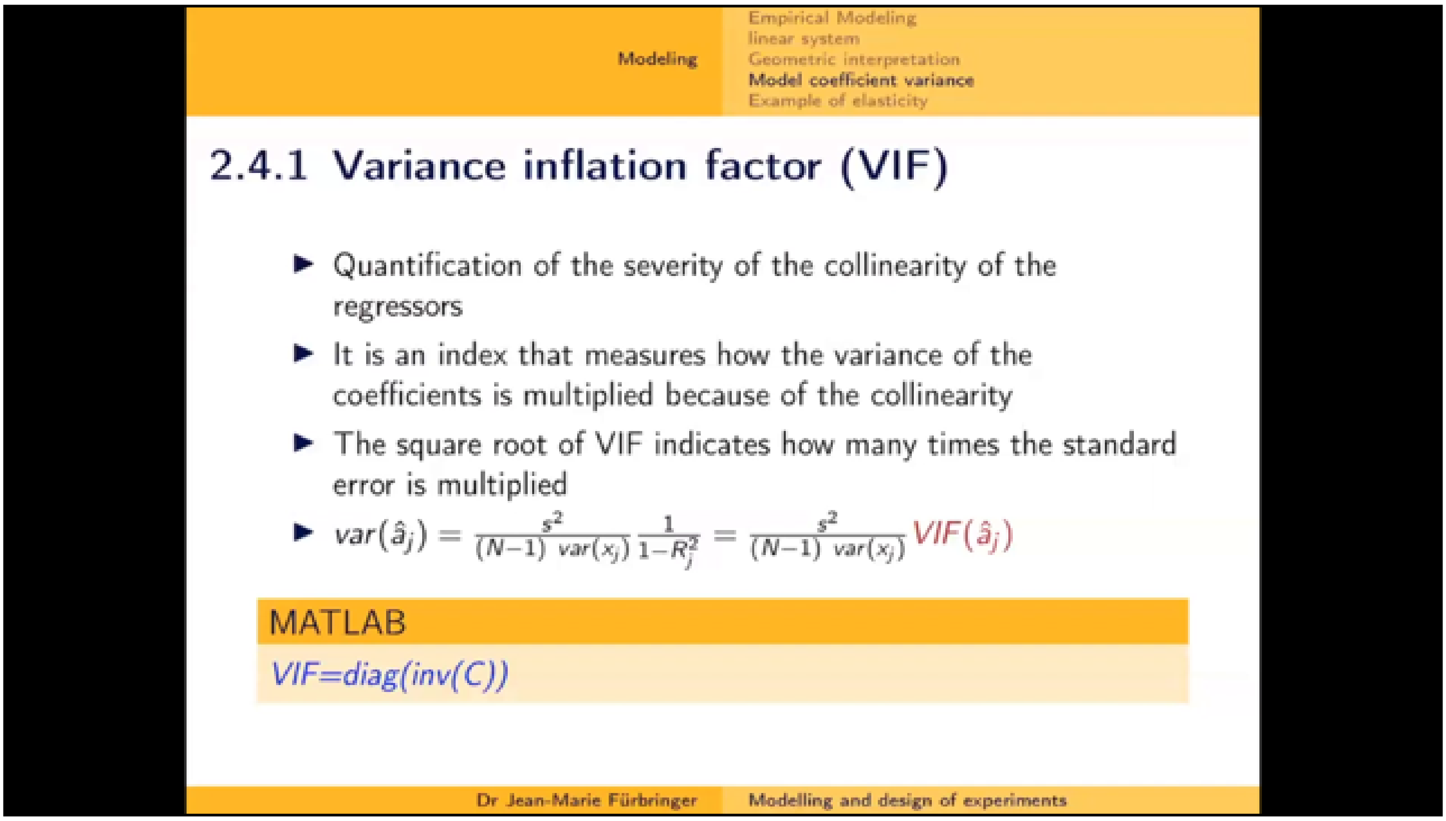


Thank you for this representation. Yes. Is there a ground tree? There are. Ground tree? Like because what are the calculations there with respect We stand from errors in our model. No, it's not the error. So this will be quite the last lesson of this chapter. We will make the difference between pure error and lack of things. So now we are mixing what could be the error of model and the error of the measurement. But what we consider as true, very close to the trueness, the truth, is the average of the measurement. It's not one measurement. But we consider that one series of measurement is really, we do not have a general bias.

notes

summary

20m 16s

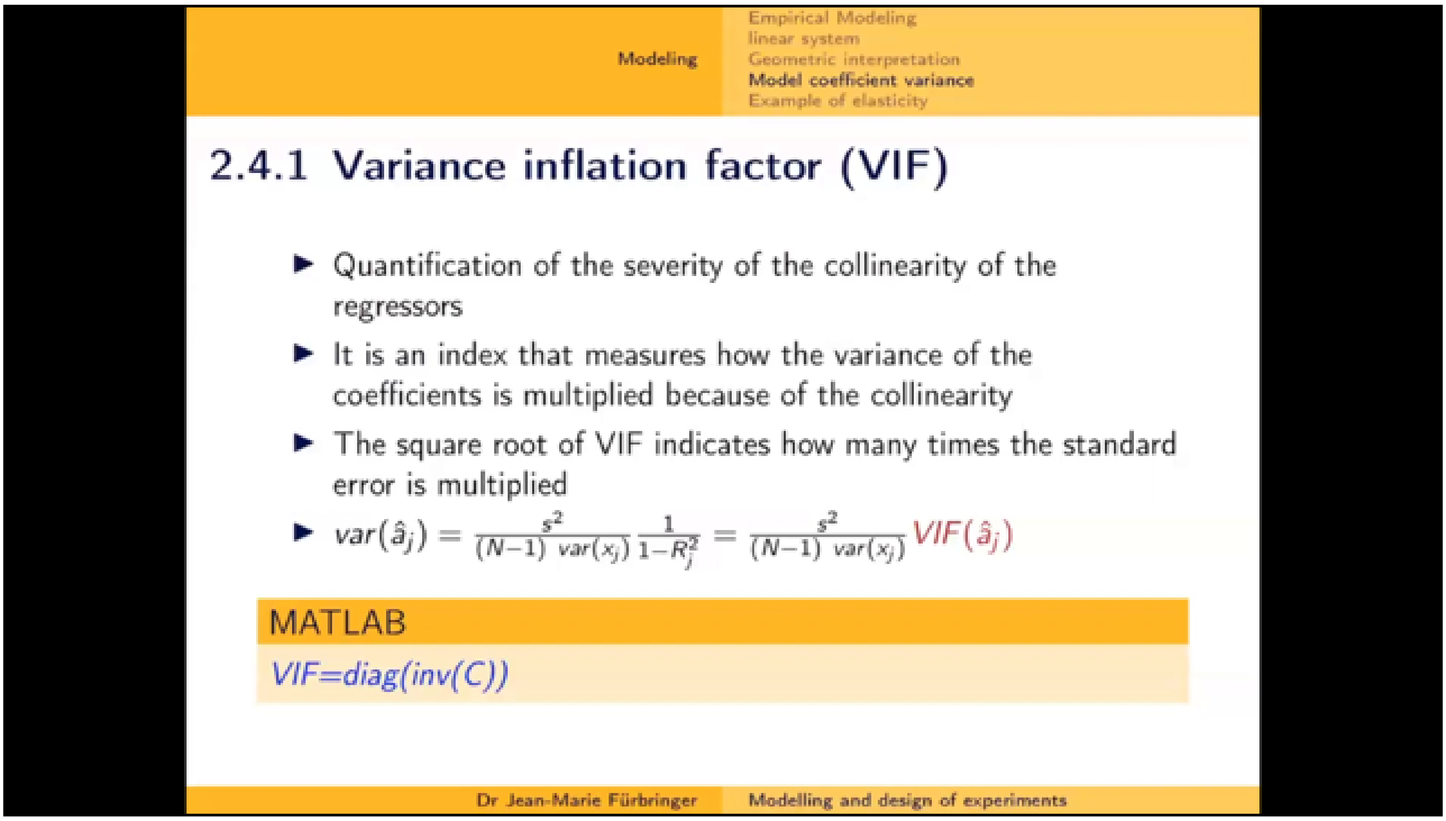


After we can calculate a few other things. So we can calculate a correlation matrix. It's not so interesting by itself. The correlation matrix is just the ratio of the element of the dispersion matrix divided by the corresponding diagonal elements. So it gives you a correlation for like that you can answer. What is the correlation between my factor 1 and my factor 2? But we don't do so much things. It's easy to calculate with MATLAB. Because if not, you have to compute the algorithm. So you have a function already, cork of the dispersion matrix. And it gives you the correlation matrix. But in fact, we calculate this correlation matrix for what's come later is the variance inflation factor. So here you have a statistical definition of what is the variance inflation factor. It's the reverse of 1 minus the square of the effect. So it's for calculating the collinearity. So this slide to understand what is it. But this equation is not important for you. This is just for explaining what is it. Just I appreciate why I think what is the variance inflation factor. What is important is how to calculate it and is what it's represent. So if you want to calculate it, you have to take the diagonal element of the inverse of the correlation matrix. So I tell you, the correlation matrix doesn't interest me in C. But it's interesting for calculating the variance inflation factor. And the variance inflation factor are very interesting. Because I didn't say that here. Sorry. So if it's 1, the variance inflation factor of 1 coefficient is 1, you are very happy. So that means that in fact, you do not have influence of collinearity between your factors. If the value of the variance inflation factor is 1, it's the best possible situation. You cannot be better than that. That means that you do not have

notes

summary

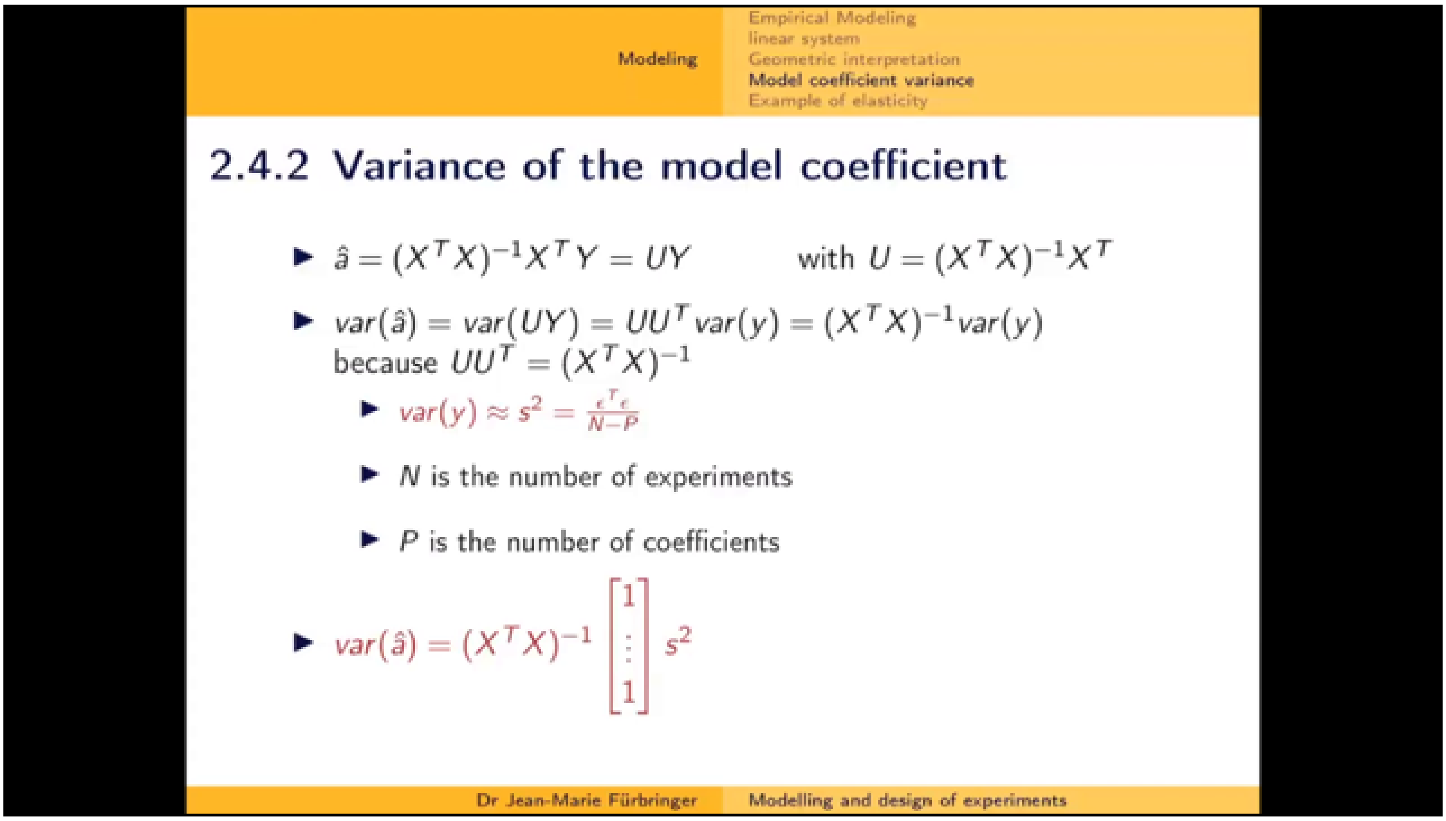
21m 40s



collinearity. What you can do is make more measurement for having the more precise value, but it's not influenced by all the factors. All the coefficient of the factors. When it's below 6, it's quite OK that you understand that you are having collinearity. Because as I mentioned, you are measuring the efficiency of the solar panel when it's dry and hot and when it's cold and wet. And so you cannot differentiate the effect of the temperature and the effect of motion. You have correlation between your factor. You are not really able to understand who is guilty. Shallow columns have difficulty. It has to suspect that it cannot because probably they work together. OK. And when it's above 10, you will have big problems of errors.

notes

summary

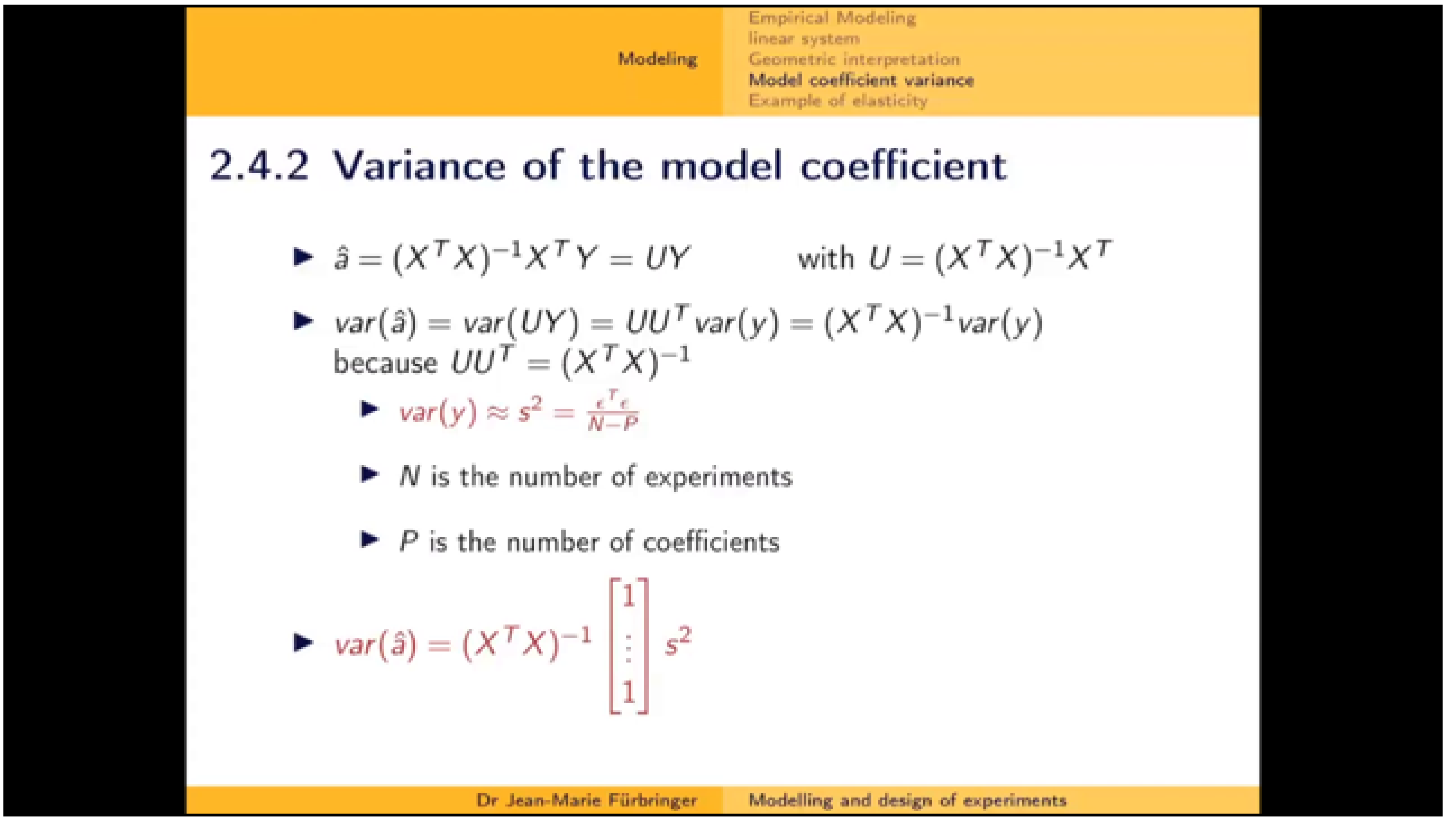


So the error will blur your model. And very probably, you have to renounce making experiments and find more experiment to do for diminishing absolutes. So usually don't try to make experiment with the dispersion matrix. Sorry, the variance inflation factor are higher than 10. It's really bad. 6 is already an error. So question? So we have a singular value for each variable. Yes. Yes. Because it's a diagonal or the inverse of the correlation. So it's. But then, number 6, is there a reason behind why they're bad? More, I say it's empirical evaluation. So 1, very happy. And 2, 3, it's OK. You understand? So when you have 4, you understand that you have to make a very attention to your error in your measurements. So it could be an argument for asking your boss for buying a more sophisticated instrument to be sure that you are really minimizing all source of error. But when it's above of 10, even if you have a very, very, very. I'm not talking about an error of some PPM. Probably you can manage that. Usually the standard measurements are wrong. Some persons know. This is a very standard measurement. If you want to understand. So I think the variance inflation factor indicates you are the culinarity is increasing. And so how many times you are multiplying the variance, experimental variance on your coefficient. So if you take the root of, you understand how many times you are multiplying the standard error. So if you have 9 as a variance inflation factor, that means that you are multiplied by 3 is uncertainty because of the culinarity. No, because this would be a very bad reaction. That means that you change your idea because of the situation. No, the modeling, you choose it because of something. So don't change your model. Change your experience. That means that your experiment are not adequate

notes

summary

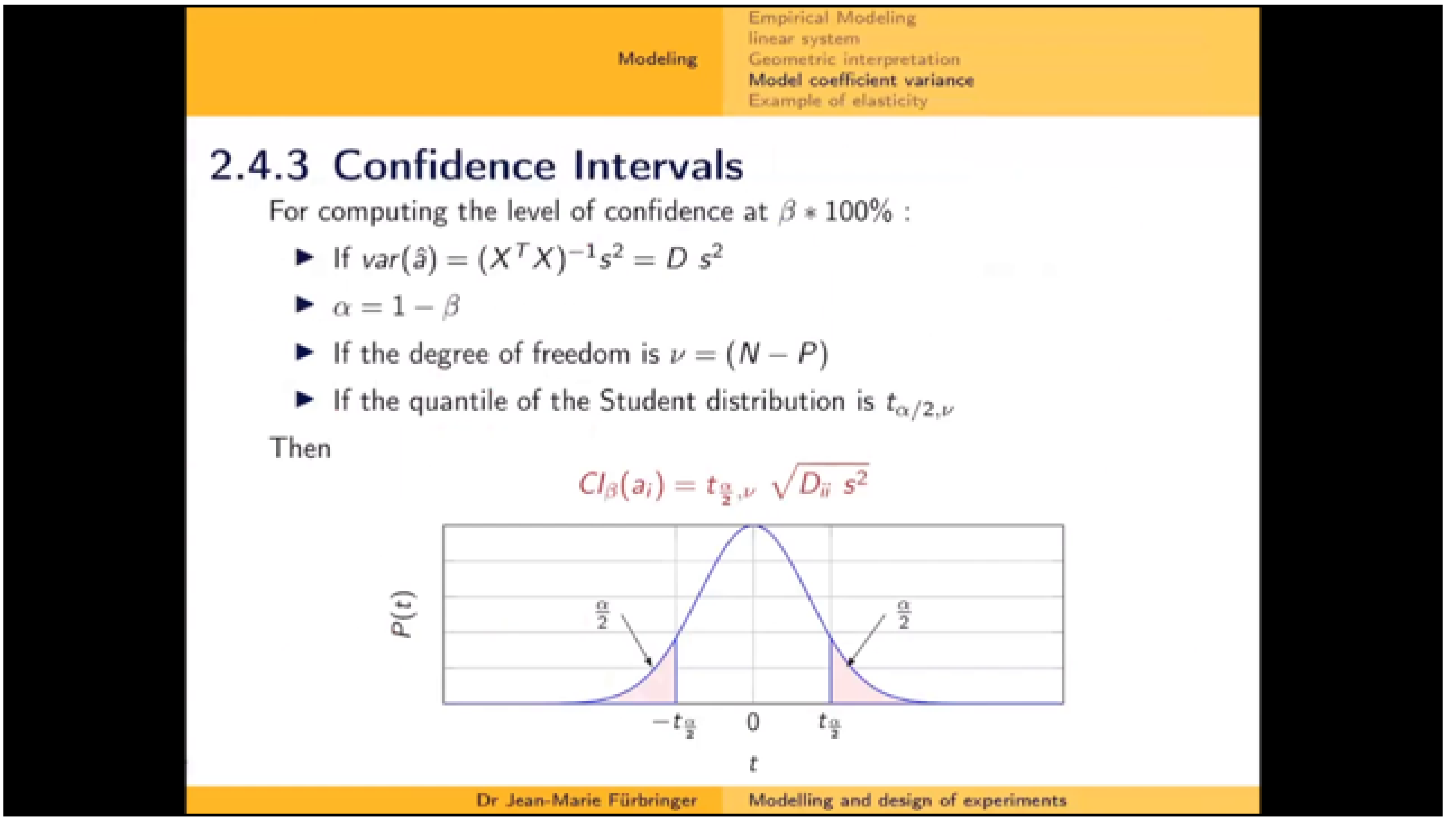
25m 13s



for checking the model. But the model is come of your hypothesis. So finally, if you have to do that, especially if you have to do that, it's a failure. Because you think that it could be a model and you have some physical, chemical, biological idea why it should be like that. And you have to change your model because you cannot make measurements. In fact, it's a problem. Yes, because you will see we'll talk next lesson about ANOVA. In fact, we are sharing some of the information alive with sum of squares. And so the sum of squares should be attributed to some factors. So we call it analysis of variance, so ANOVA analysis of variance. But you have a package of variance who are not able to attribute to a factor. So you have some variation and you don't know where is this variation come from. You know, it's come from two factors that you don't know which one is responsible. Exactly shallow conflict from two suspects and they make your answer so that is enable it. It's him, it's him, it's him, it's him. And it's exactly that situation. You are unable to separate. So one solution is to take this information out and try to use the rest of the information that have been attributed. But it's still a weak situation because you have this variation.

notes

summary

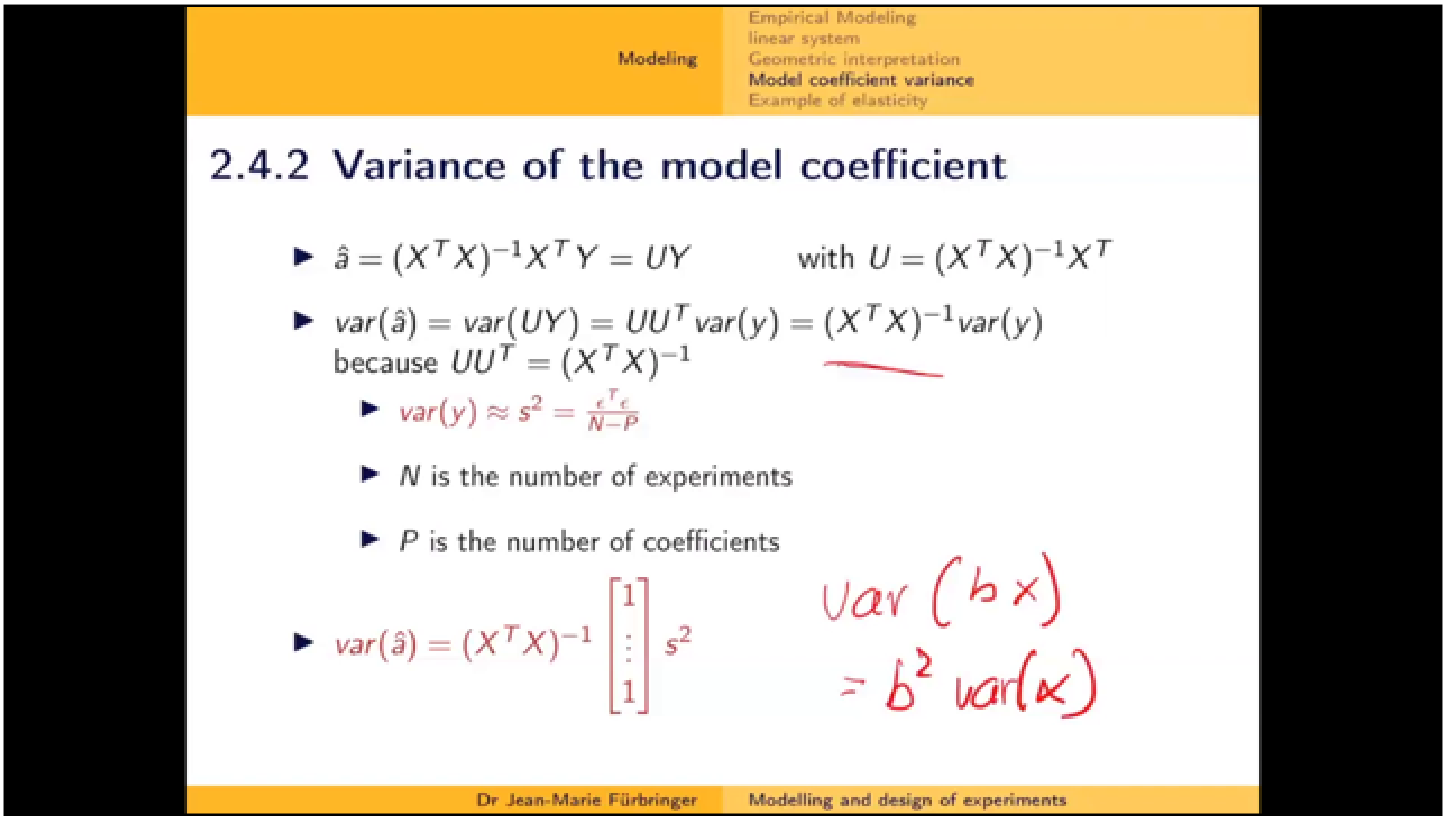


It's come from somewhere or to cut things and say, okay, the first one takes the responsibility.

notes

summary

30m 13s

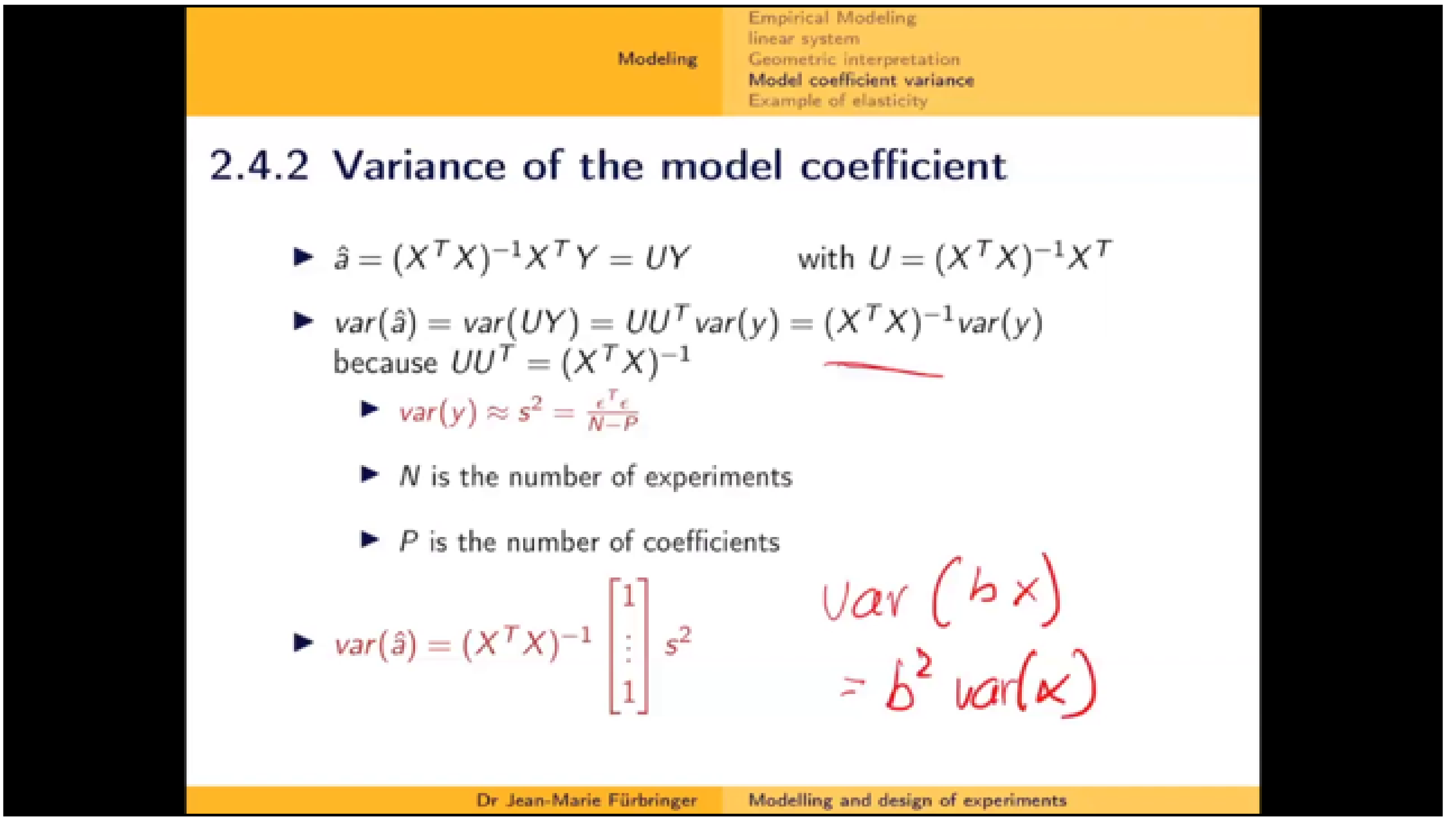


of the coefficients alpha hat. So the way of writing it is X transpose H minus one X transpose Y. I repeat it because I want this formula entering your memory and I can represent it more simply defining all this estimator by the letter U. So just saying, okay, I define U as X transpose X minus one X transpose. So I write U equal Y just for making the calculation more and more easy. So if I would like to calculate the variance of my coefficient, my estimate, if I repeat the fact is the waste statistician indicated is my estimate, what I have estimated. So the variance of my estimator is then the variance of the operator. So the variance of hat hat is the variance of U multiplied by Y. And so in this, when we make least square feet, we consider that the X value do not have errors. It's not true, but it's the way of calculating. So each time you make a linear regression, you accept errors in your response, but you do not consider errors in your, if you want to accept, you need to make bias ends estimation. So this would be another question. But in this course, we apply least square feet to standard method in statistician. We do not accept errors in the X values. So in the variance of U multiplied by Y, Y is a random variable. But U is not a random variable. It's considered as a coefficient. So if you want to remember the rule of the variance, the rule of the variance, if you have a coefficient to a random variable, when you put it in evidence of the variance operator, you have to take it at the square. of variance because the variance is a quadratic function. So we do that except that u is not exactly a constant, is an operator. So

notes

summary

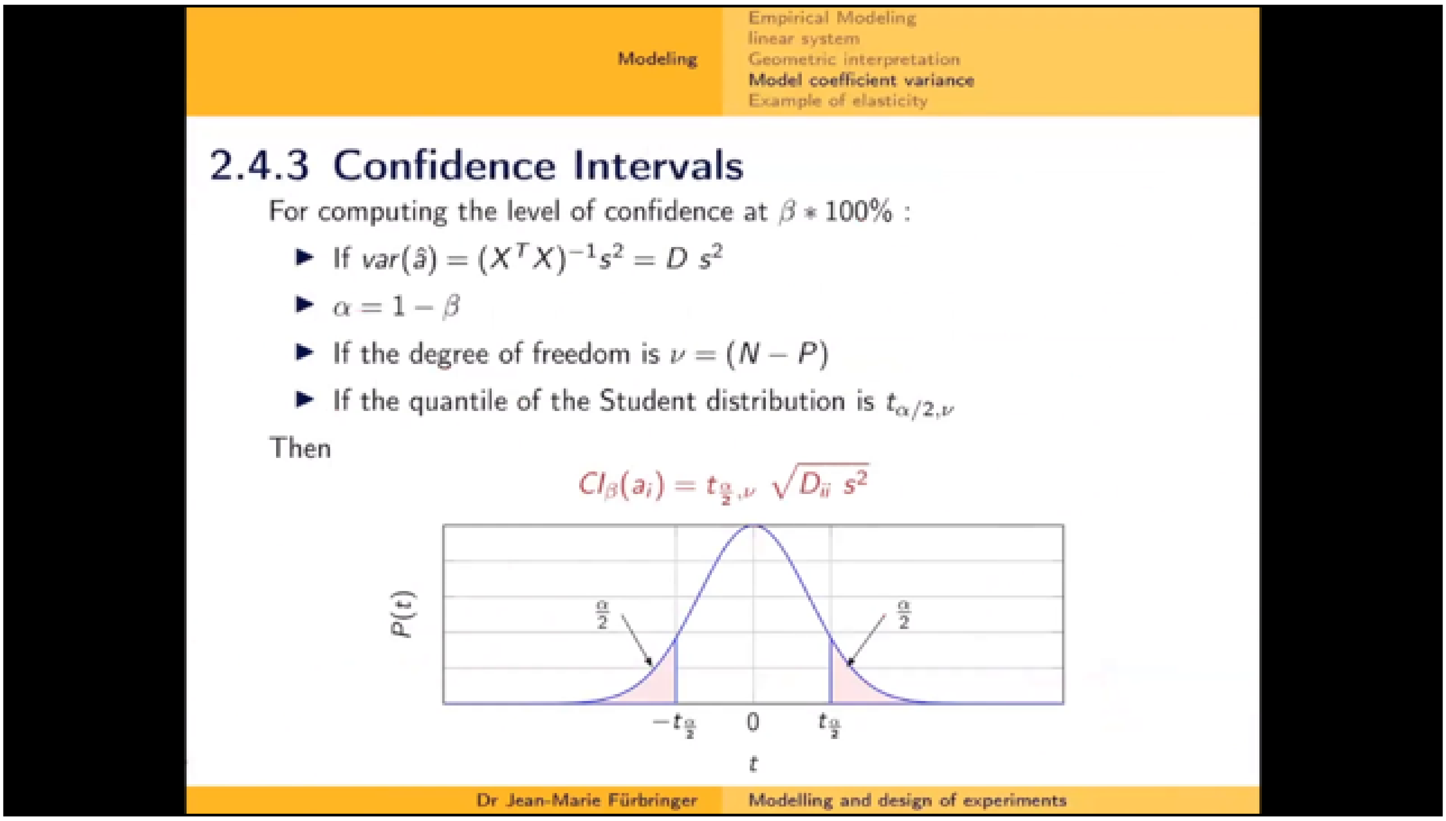
30m 43s



is the matrix. So when you take this square in evidence, so it will be u multiplied by the transpose of u. You can try it. And so if you make u multiplied by the transpose of u corresponding that u is x transpose 8 minus multiplied by x transpose, you finally obtain that it will be the dispersion matrix multiplied by the variance of my measurement. So it's why I mention it that, in fact, the dispersion matrix is representing the dispersion matrix, is representing how the uncertainty is flowing through your estimation process. Because finally, the variance of the coefficient will be the matrix of dispersion multiplied by your measurements. And if you would like to calculate

notes

summary

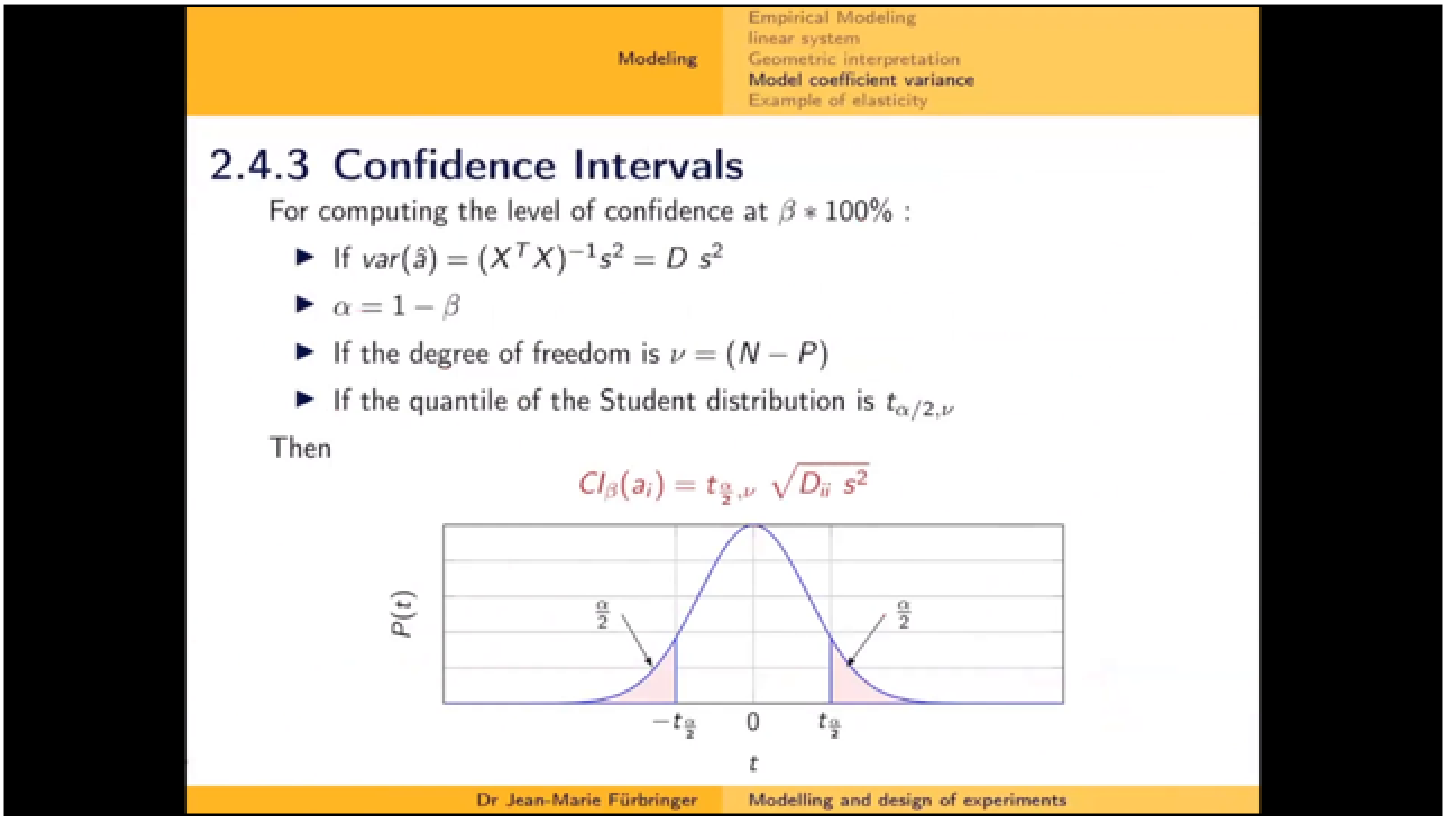


the variance of your measurement, you have different way. You can make several times your measurement and get variance for your measurement. You, perhaps, know your instrument. You have made several. But a way of doing it is what is called the sample variance. You can take the variance of your measurement. We can call it with an s square. It's not a sigma square because it's not the theoretical variance. It's a sample variance. So we use Latin letters. And it's the square of your errors divided by the degree of freedom. So the square of your errors is the residue of your simulation. It's a difference between your model and your measurement. You take the square, so epsilon transpose epsilon makes the sum of the square of your errors. And you divide it by the degree of freedom. And the degree of freedom is a number of measurements, not minus 1, because usually when you calculate just the variance like that, it is the variance of a series. So it's the variance of a model. So it's minus the numbers of parameters in your model. So if you have a linear model of two factors with interaction, p would be 4. A0, A1, A2, A1. And then you have a variance. So now you can calculate the variance of your coefficients. Having your matrix of dispersion multiplied by a vector of 1 for having the algebra respected. And multiplied by the sample variance. And this gives you the variance of your coefficient. And in this situation, no, I don't say more. What about my time? Yes, I will finish these slides. And you will see that in fact, yes, I'm talking. I'm going at the half of my velocity at the previous level. It's OK, no problem. We have time. So no exercise this week. You can try, but usually you better wait next week for making the

notes

summary

35m 13s



exercise. But some of you have videos to check the past videos that you didn't see, et cetera. Work you have, but no exercise this week. So just finishing, I will finish the course with this slide. So if I want to calculate the confidence interval, so I'm able to calculate the variance of my coefficients. The matrix of dispersion multiplied by the sample variance. Algebraically, I should have this vector 1 that I didn't write it. So it's d multiplied by the vector. If I want probability of my confidence interval, let's say 95%. So I have to calculate the beta would be 95%. So I will calculate what's the. So it will be 5% in this case. So the degree of freedom will be the numbers of measurement minus the numbers of parameter. And then using the theorem of the confidence interval but adapting the rule, I can calculate my confidence interval for probability beta, imagine 95% by calculating the, now you know it, I spent some time in the previous lessons, the student coefficient for the probability 5% divided by 2, because I won't have 5% adding both, the both limits. And multiplied what's changed with the previous formula presented the first slide and multiplied by the roots of the diagonal element of the dispersion matrix. So the E is in the dispersion matrix with diagonal element corresponding to the coefficients I am estimating and multiplied by the variance sample bias and taking the hoop. This give you the confidence interval for your coefficient at beta probability.

notes

summary