Course material

Course:

ENG606 / PHYS 442

Video:

DOE_lesson4_part1_DOEbasics

Concepts (extracted from automatically generated subtitles):

Empirical model. Multiplication of matrix. Important slides of last week. Linear model. Vector of coefficients. First lego block system. Type of model. Matrix of dispersion. Different elements. Last time. Estimate of the coefficient of the model. Ellipsoid of confidence. Most important slide. Second step. Matrix multiplication.

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

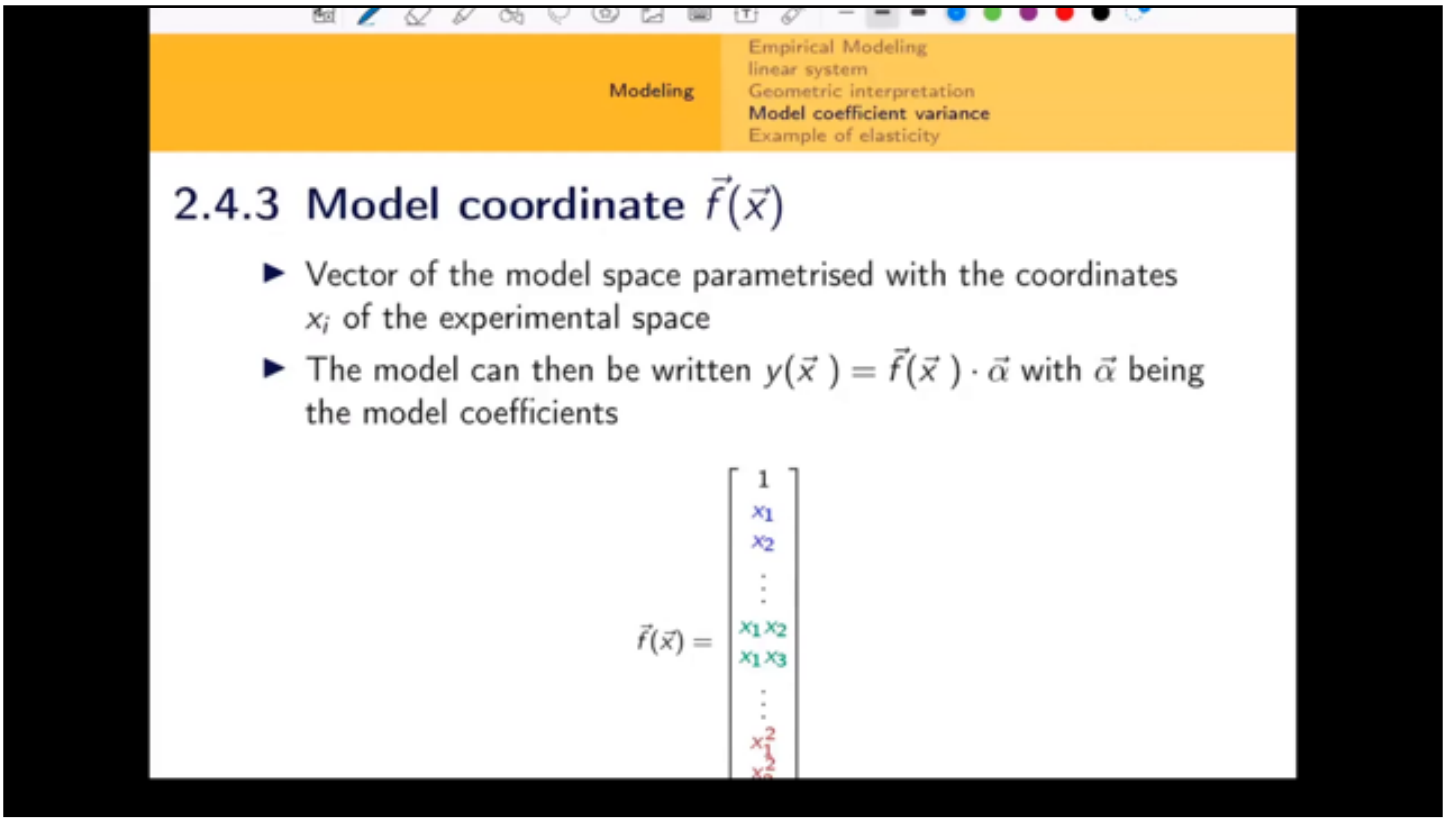
to video

...

notes

summary

0m 0s

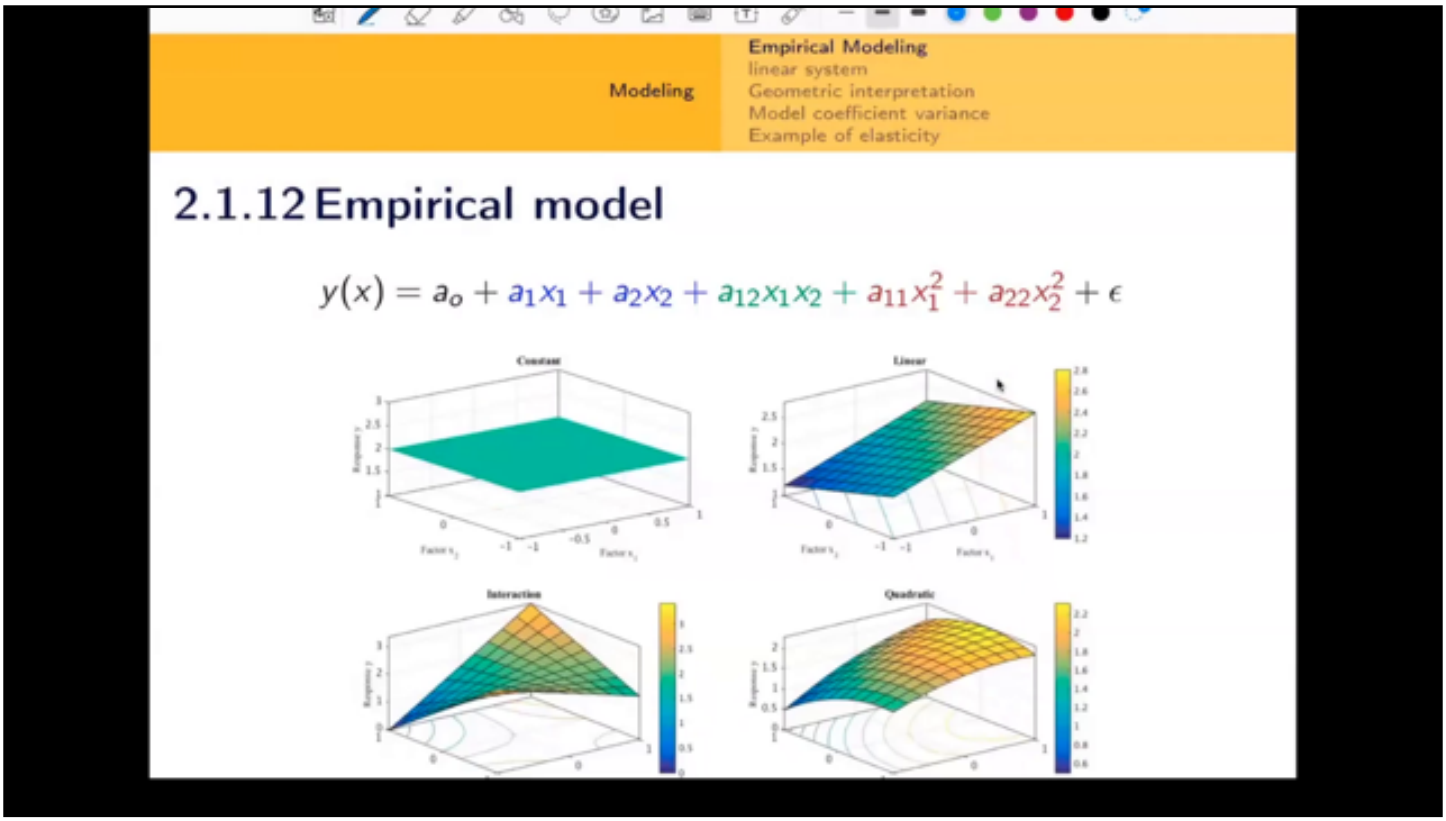


These subtitles have been generated automatically So really happy to see you. I hope you are well.

notes

summary

0m 1s

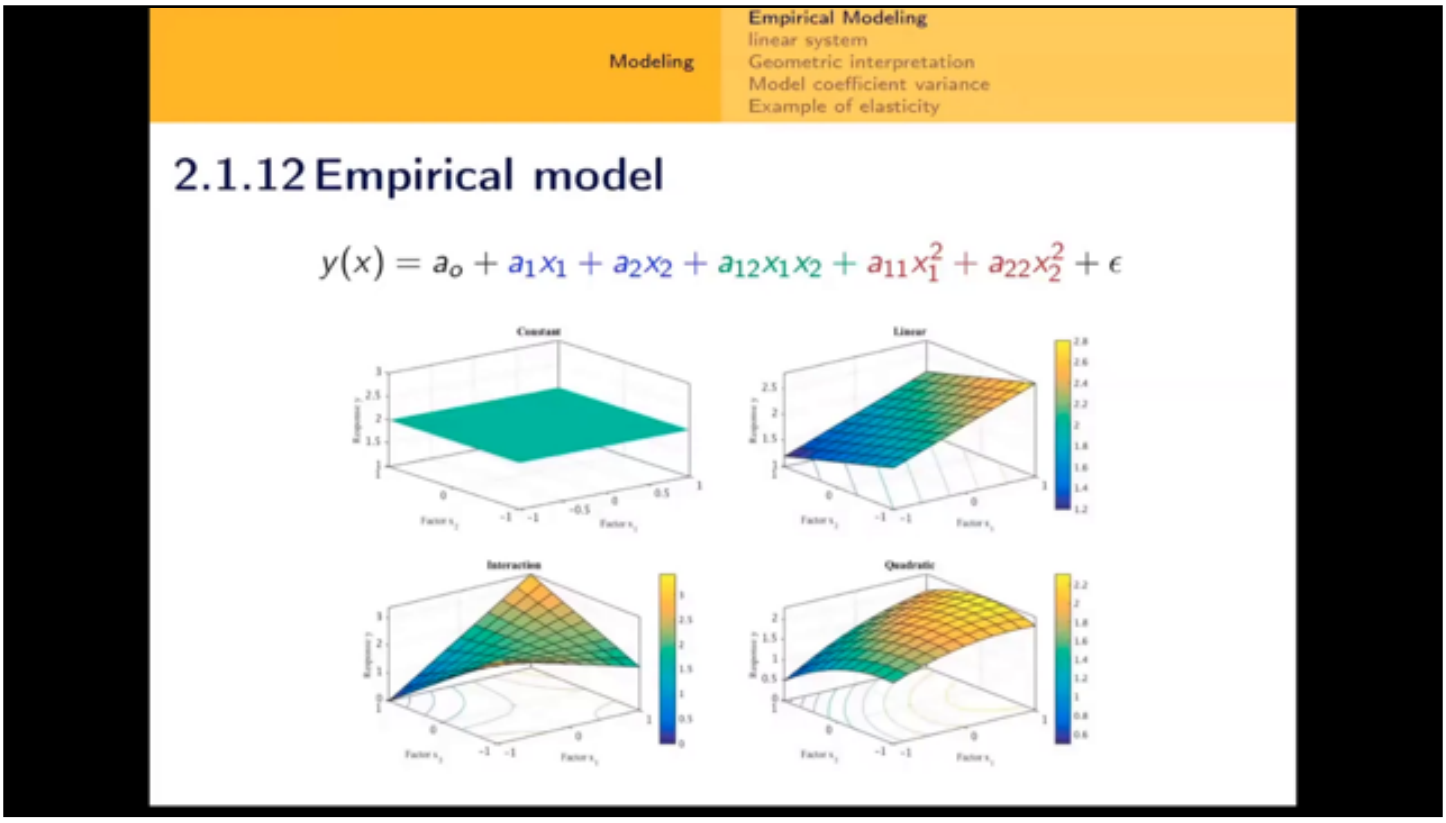


So we will continue today the chapter I call modeling. We will see how to model a situation. And it has mentioned when we know the different elements for modeling, we will look. So we are looking at all the things that we can check before the experiments. The next lesson will be some few things we can check after the experiment. And after that we will enter in discovering a few designs that are very good in very typical situations. But we will have enhanced tools for everything seemed before the experiment and also the very things after the experiment. Sorry if I'm not all the time absolutely clear in that but I try to insist. Because when you are in the experimental situation it's very important to understand what you can get before the experiment and what you can get on the app. Of course sometime I mix and present results and so it's very important to have this clear.

notes

summary

0m 5s



So if you if you remember last time, if we look at a few very important slides of last week, perhaps for me, one of the most important slide is this one about empirical model that doesn't mean that you are obliged to make a model that's all you have a model of your own because you are working on physical biological chemical and you have your model and try to see what type of model you have. But I mentioned that for applying what I propose you need what we call linear model that means models that can be represented with the multiplication of matrix plus a vector of coefficients is what we call the linear model to be differentiate from linear response model. I complain. And so when you have no idea about what could be the model, then this is a first Lego block system that you can start to say okay, nothing happened. It's flat. You have difference in your measurement because you have errors in your measurement etc. In fact, nothing happened. The factors that you have chosen have no influence on your phenomena. The second step is to say okay yes my factors have influence and the first hypothesis which is logical to say they have linear influence direct influence so you have an hyperplane, which is defined by the slope related to each of the factor. After you can you start to sing okay but eventually the effect of one factor depend on the level of another factor. So you have instructions and instructions is all the world we can have very sophisticated type of interaction. But usually we start by the first level of interaction which means that the product of the two factor, which make an interaction, and you see that you start with a page after you have to make a slope and after you start to bias a little

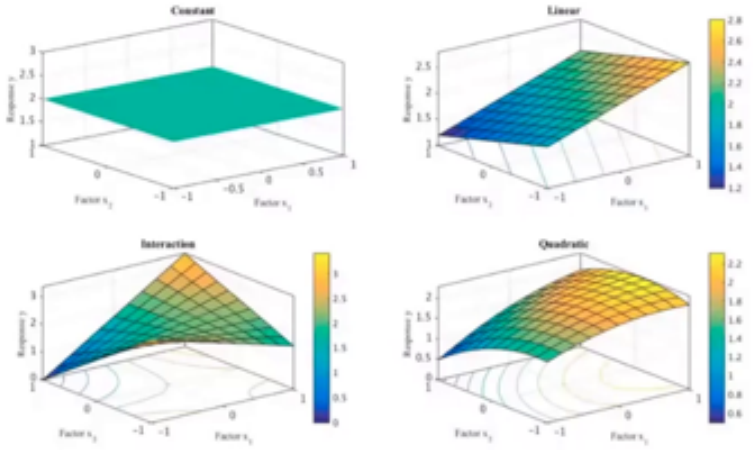
notes

summary

1m 35s

2.1.12 Empirical model

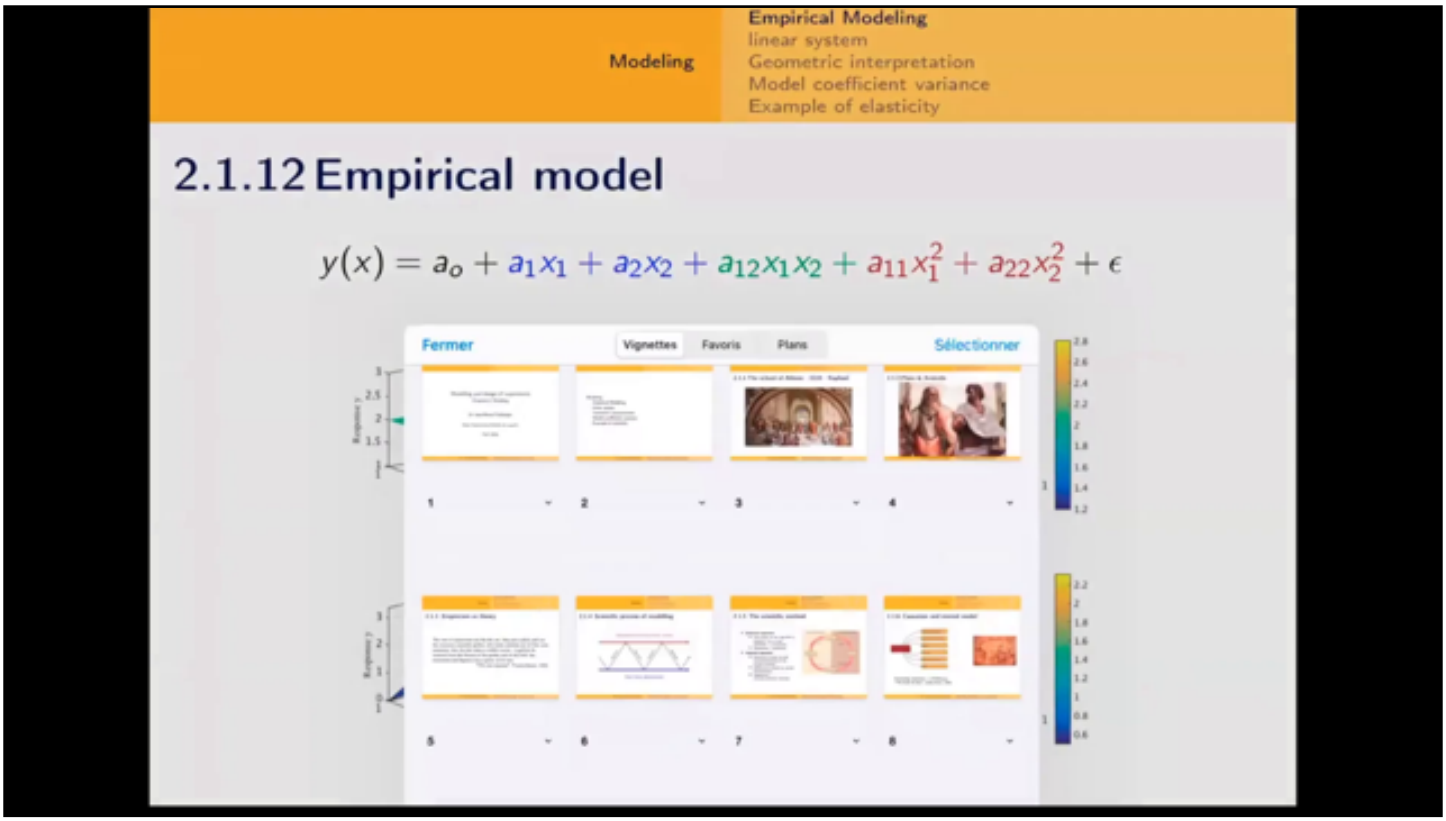
$$y(x) = a_o + a_1 x_1 + a_2 x_2 + a_{12} x_1 x_2 + a_{11} x_1^2 + a_{22} x_2^2 + \epsilon$$



bit the plane. And finally, the force empirical model which is of high interest is having a quadratic model you see so. The second instruction is already a second degree model and that we analytically speaking, but phenomenologically is different from having just it. No, thank you. And so it's different from. I forget my, my, it's different having just influenced from another or having a quadratic which is an influence of your factor of the whole. No, so it's a little bit different. It's the same thing but phenomenologically is different having interaction or having a quadratic model.

notes

summary

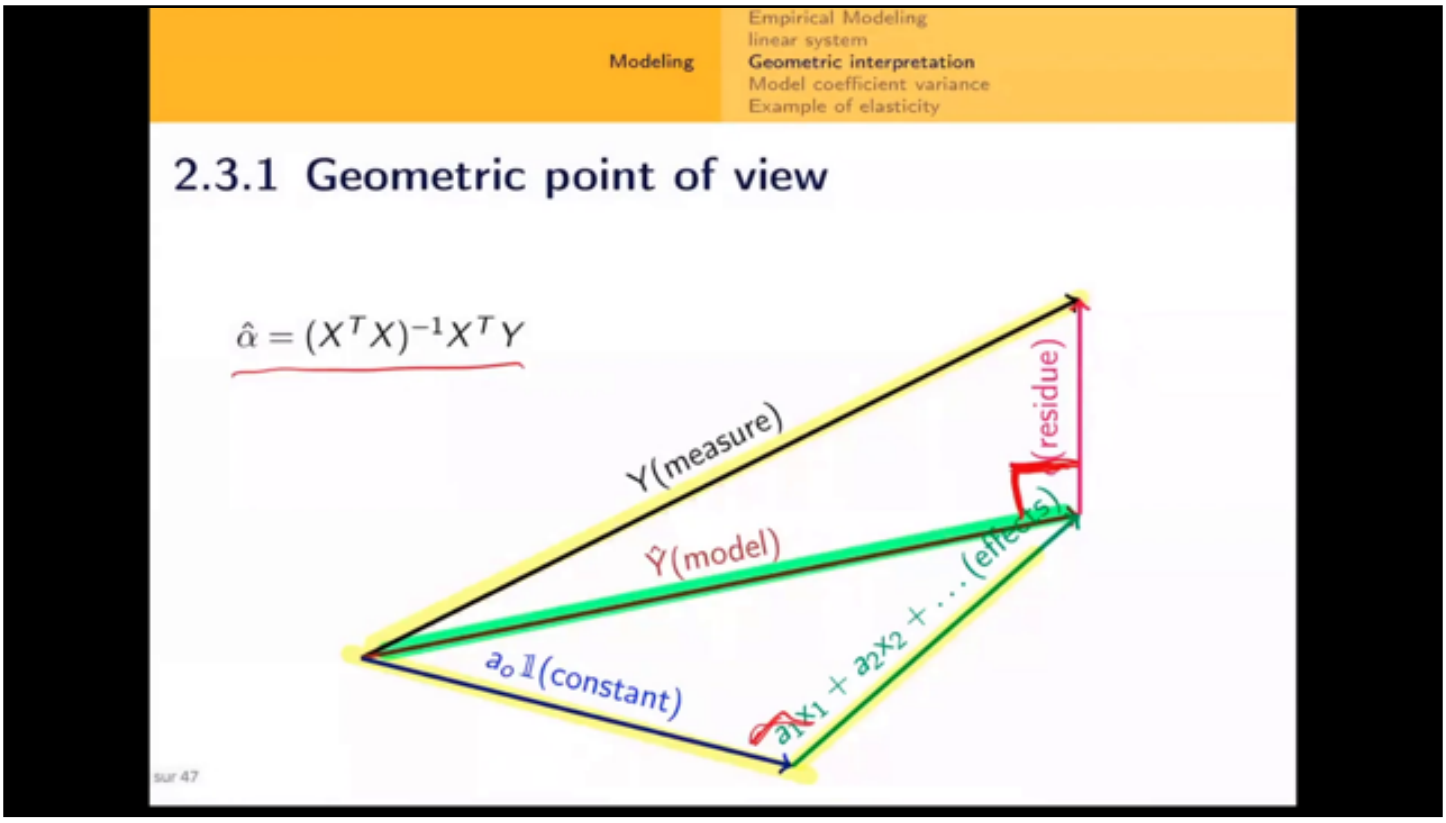


And as I mentioned also, perhaps your reality is not as beautiful as a pure quadratic model, perhaps one side of your heel is different than the other side of the hill, but it's already allow you to make a lot of things. To improve their most sophisticated model but if you are able, if you have no idea about what could be the model it could be a very first try to work and we will with the design, you will see that will present you designs to go around these different basic model, step by step, so possibility or to jump directly to the last one because you already know that you want to optimize some things or you already know that you want a surface response.

notes

summary

5m 13s

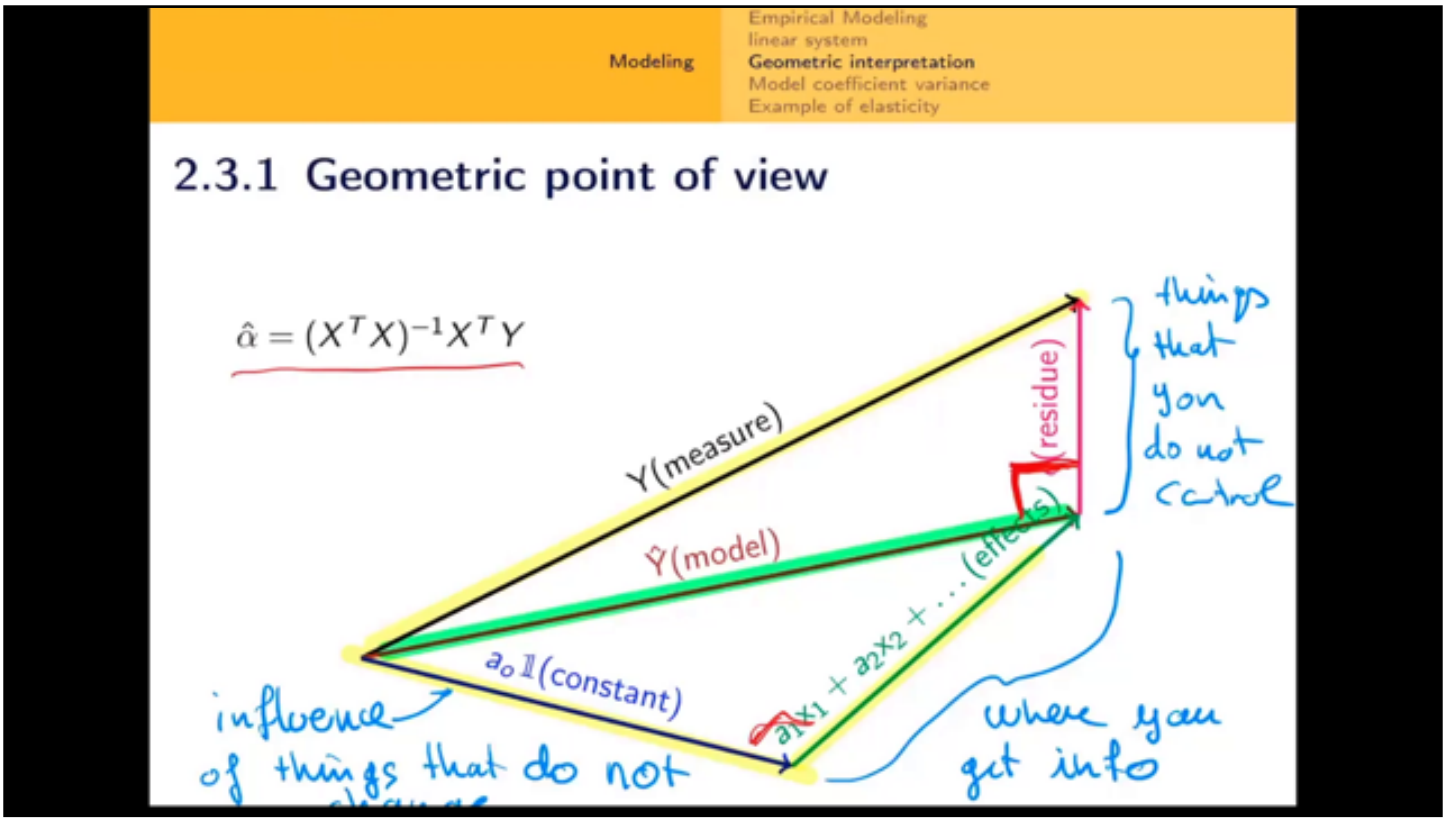


So this was one of the important slides of last week, and another important slide, at least in my point of view, because it helped me a lot to understand what we are doing.

notes

summary

5m 49s

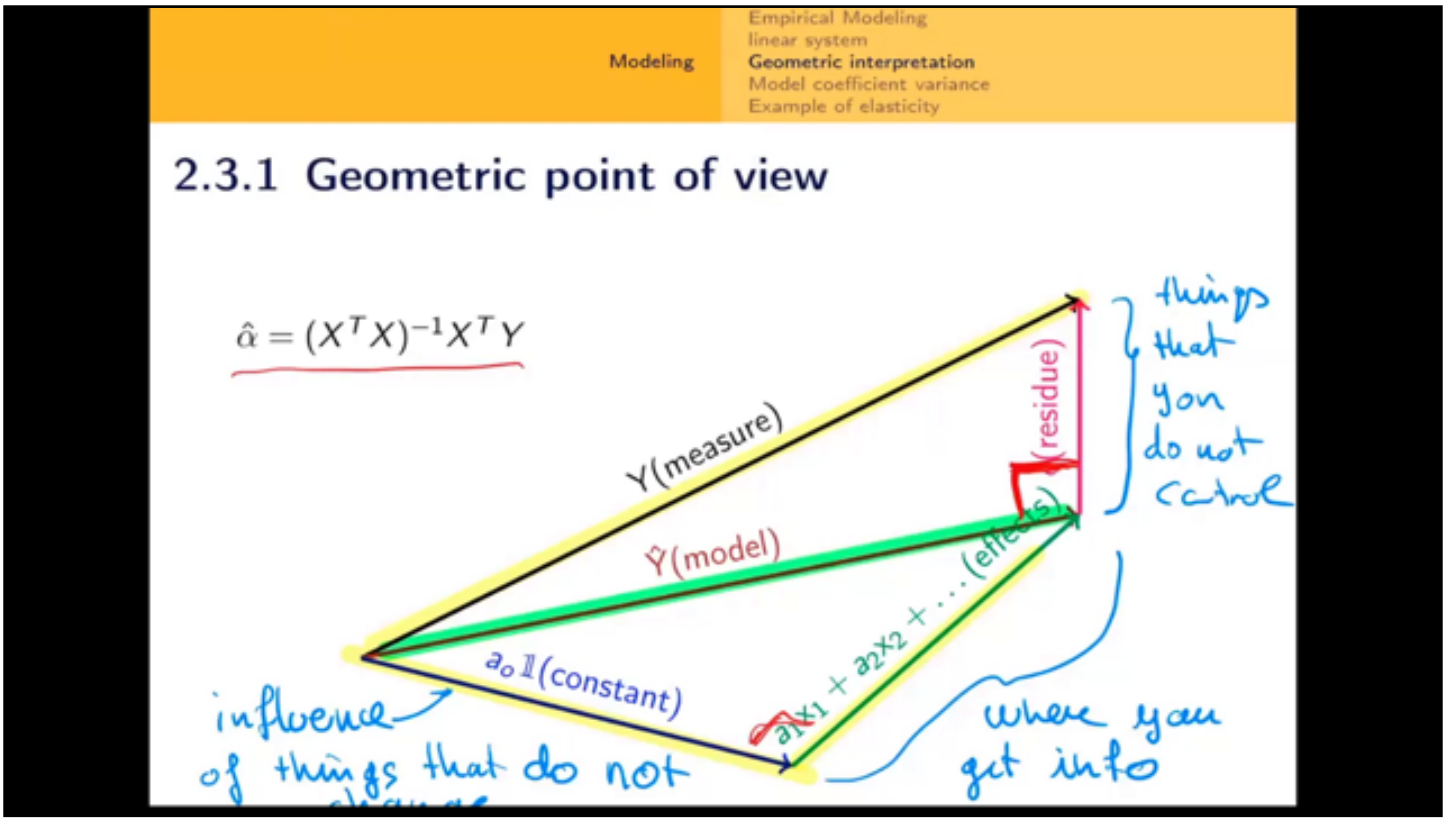


And it was very important for me to make the parallel between the geometry between the statistics understanding what what we do. So in fact, when we are calculating the estimate of the coefficient of the model, geometrically we are making a projection. Because we have one of the first the matrix of dispersion is a question of metrics, what bring most of the information on your different factor is the X transpose why, and it's a matrix multiplication so you understand that the matrix multiplication is just an automatic series of scalar product and the scalar product is making a projection. So in fact, when you are modeling the situation when you are fitting the model, what we're doing, modeling a data. In fact, we are projecting our vector, this vector has as many coefficients, coordinates, has you have Muslim and 100 Muslim and you have a vector of 100 dimension, and you project it on a subspace corresponding to the dimension of the space is correspond to the number of factors that you have. And let's say coefficient, you can have instruction, you can have a second degree. So this represent what you have in three dimension where you have a vector y representing the measurement, and you are projecting the constant I tell you is a fix the direction of the vector is fixed is a vector going from the lowest point in your domain to the highest points in your, in your domain and you just have the length so the degree of freedom is one and after the, the effects you have one vector per per coefficient in your in your model, and after you have your residue, and perhaps something I didn't insist sufficiently. This is very important to understand something that, in fact, the constant corresponds to things. Well, but that was not the right. So this are, let's say, influence of, I

notes

summary

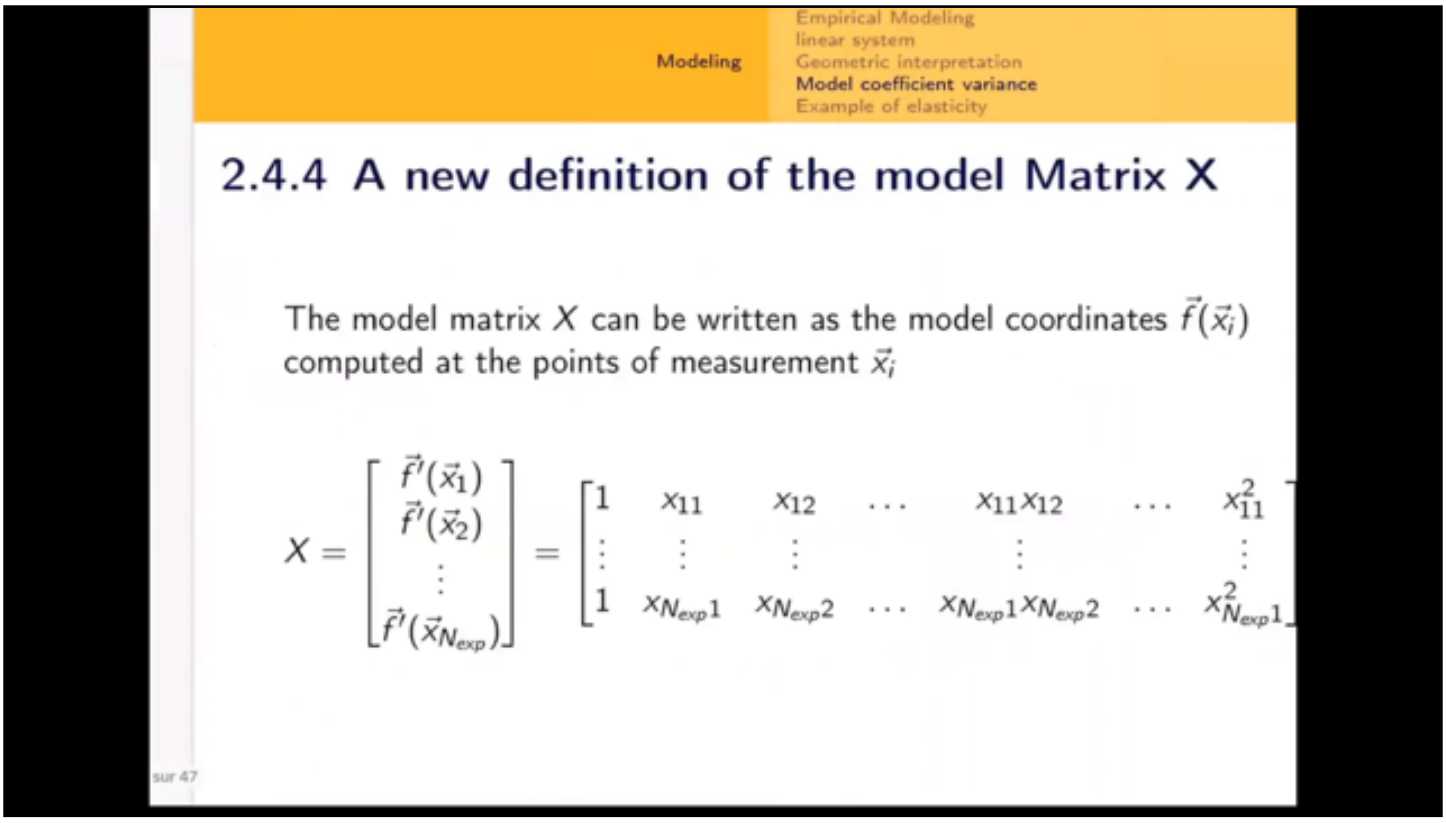
5m 52s



say things for having a very generic words that do not change. So you have to have a lot of experiments. You don't explain so is a part of the reality that you just put a part and say okay, it's not my business this time and I let it apart it doesn't change and you have some consequence on your response but you don't manage it. You get information on this thing that escape your control. There are things that have changed during the experiment, but you don't know what it's random it's the origin of your errors. Other things that you do not control. modeling is all the time the separation of your information in three parts. What doesn't change and you don't manage it, you let it apart, it's okay. And saying that change, you don't manage it, but they change your experiment and what is the result of your work, of your experiment, because you have very factors and they have an influence. So you understand why for a situation you can have different models. It will depend what you manage in the project with a master's student. If we do not manage the temperature of the room or the temperature of the device, then it will be in the constant. If it doesn't change during the experiment, if we are in a room with the temperature changing or you are not managing the temperature, but the temperature of the device is changing during your experiment, it will be in the random part. So those are probably the two slides that really are the most important. The rest is also important, but say it's really explaining the last steps.

notes

summary

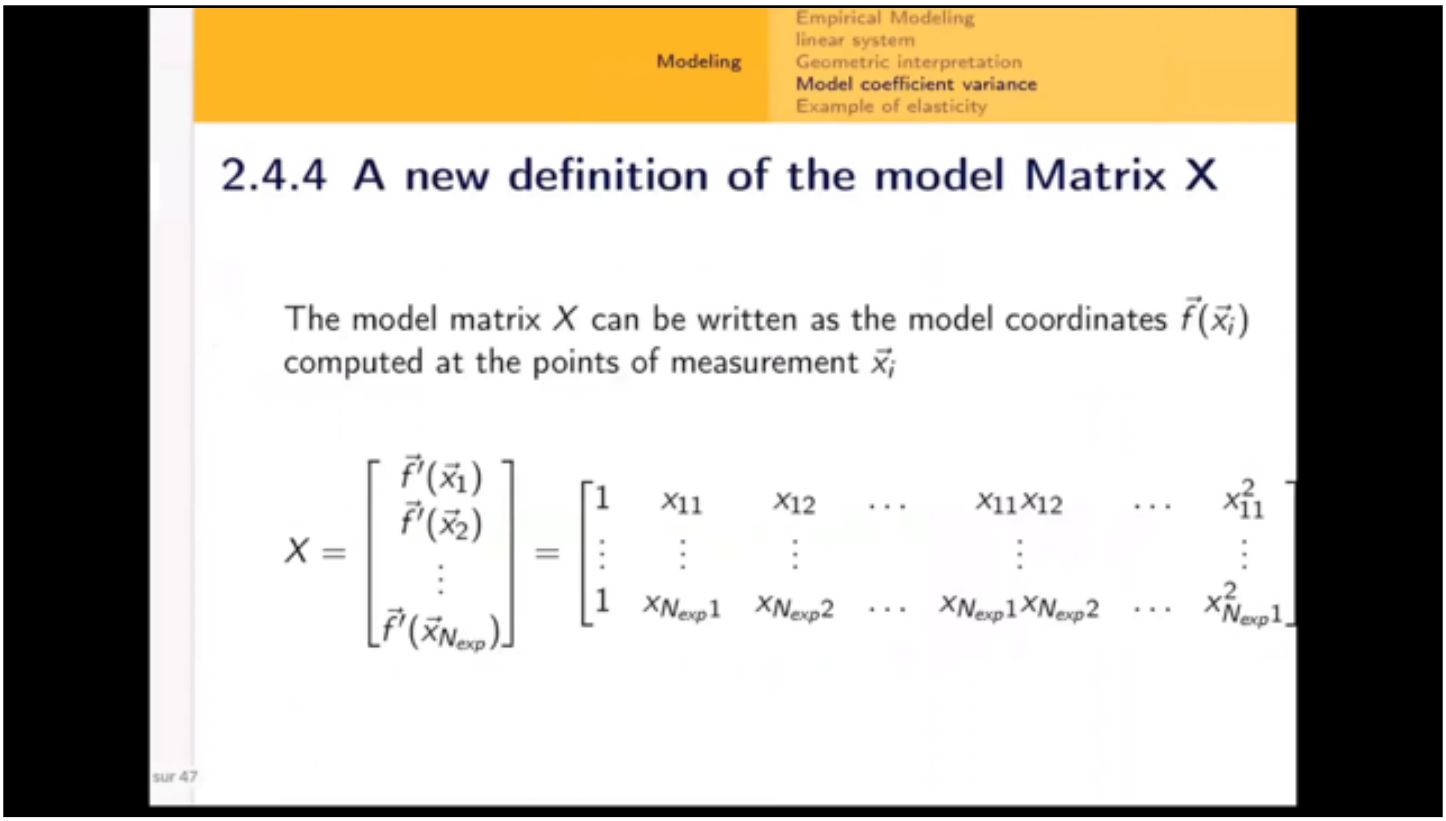


So one thing which is just easy, just easy, it's easy. The writing of a few things is what we call a model coordinates. And so it's in the form of a vector. The dimension is the number of coefficient of your model. So it's a model coordinate, so it's related to the model. It's not related to your movement, it's related to the model that you have chosen. Linear, linear interactions, quadratic, cubic, anything. And it works with linear models because we would like to represent this. We want to separate the coefficient, put in evidence the coefficient. So if you have an exponential within the exponential, you have product of factors you cannot represent with this type of model. So each element corresponds to one part of the model. So the first one is corresponding to the constant. So when you have put in evidence a constant, what stays one. And after you have the linear as a part of your model, so you have put in evidence a one, a two, a three, etc. So what stays is x one, x two, x three, the coordinates of your points. And after you could have any terms that are as that you have in your model. Here I put in green the instruction part of your model. And after the quadratic, the cubic and any part, it could be even producing an exponential part of your model. If you have exponential of x or exponential of x one or x two. But as I said, if you have exponential of x one by x two, it could be complicated to represent it linearly. In design of experiment, we are all the time interested to separate what you decide before the experiment, what you get after the experiment, what is an hypothesis, what in an estimate. So it's really, it's very nice because like that, this is your hypothesis

notes

summary

11m 51s



of the model and you can have different apps. So be careful, we call it x. Sometimes I put a prime, it's not a derivative, it's a transpose, you can really define it vertically, sometimes we use it horizontally. You will see a few examples when we use it, but it's just an instrument. It's nothing, nothing special.

notes

summary

2.4.3 Model coordinate $\vec{f}(\vec{x})$

- Vector of the model space parametrised with the coordinates x_i of the experimental space
- The model can then be written $y(\vec{x}) = \vec{f}(\vec{x}) \cdot \vec{\alpha}$ with $\vec{\alpha}$ being the model coefficients

$$\vec{f}(\vec{x}) = \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ \vdots \\ x_1 x_2 \\ x_1 x_3 \\ \vdots \\ x_1^2 \\ x_2^2 \\ x_3^2 \end{bmatrix}$$

So a first use we can do of this model coordinate is to rewrite my model matrix. So my model matrix, which was the matrix with one column per coefficient, one line per experiment, in fact, is the vector horizontally, the model vector for the coordinate where you make experiments.

notes

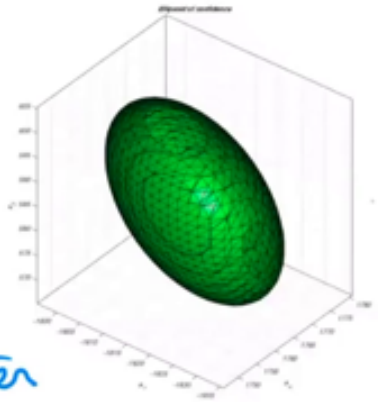
summary

14m 31s

2.4.5 Ellipsoid of confidence

- If the coefficients α of a model $g(x_i, \alpha_i)$ are considered as Normal random function : $\hat{\alpha} \sim N\left(\alpha, (X^T X)^{-1}\sigma^2\right)$,
- Then, if η represents the true model, it can be written $\eta - \hat{Y} = X(\alpha - \hat{\alpha})$,
- And there is a probability $1 - \beta$ for finding the true value of the coefficients α within the ellipsoid defined by $(\alpha - \hat{\alpha})^T (X^T X)(\alpha - \hat{\alpha})/ps^2 = F_\beta(p, \nu)$
- If the regressors are not orthogonal, the confidence intervals of the coefficients are not independent

number parameter



So when you define the model is for anywhere in your domain, so this model coordinates, you can instantiate it for any place in your domain. And if you just take the collection of this model, the model coordinates for your measurement points and you classify them horizontally, it's give you the model matrix and you can write then your model matrix like that. Just writing. Okay, another concept just to let you know we will not use it a lot, but it perhaps helps you to visualize things. So if we are in the space of your coefficients, imagine that we have a linear model, two factors, one constant. So we have three coefficients in our model and we would like to understand what is the uncertainty that we have on the coefficients. So when you make the estimates of the alpha, that means that you have three value, one value for A0, one value for A1, one value for A2. So it's a point in one, three dimension space. It is a center of this pillar. So what we want to represent it, and sorry it's a really statistical way of writing things, but just please understand the spirit of it. It's not so, so, so, so, so, so, so, so, so. So we would like to say that the uncertainty that we have on a fit could be represented as a space around our estimated coordinate of the three A0, A1, A3. And this ellipsoid of confidence would be, okay, so what is the space where we could have the values of A1, A0, A1, A2 around the estimated point with 90% of confidence, etc. So is the generalization of the confidence interval for multi-dimensional fit, where you have to fit an A0, A1, A2, and after you can have a project. So what we say here is that we say, okay, the estimate of our coefficient,

notes

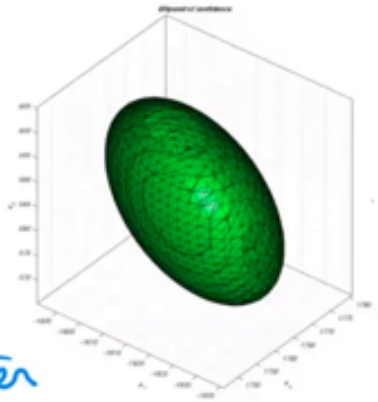
summary

15m 6s

2.4.5 Ellipsoid of confidence

- If the coefficients α of a model $g(x_i, \alpha_i)$ are considered as Normal random function : $\hat{\alpha} \sim N\left(\alpha, (X^T X)^{-1}\sigma^2\right)$,
- Then, if η represents the true model, it can be written $\eta - \hat{Y} = X(\alpha - \hat{\alpha})$,
- And there is a probability $1 - \beta$ for finding the true value of the coefficients α within the ellipsoid defined by $(\alpha - \hat{\alpha})^T (X^T X)(\alpha - \hat{\alpha})/ps^2 = F_\beta(p, \nu)$
- If the regressors are not orthogonal, the confidence intervals of the coefficients are not independent

number parameter



our coefficient estimated alpha hat for saying that it's the estimate. We say, okay, it's following a normal probability distribution with the value of alpha that we have as an average. We won't say, we pretend that we fit quite correctly our word. And after as a variance, we can calculate the variance with the matrix of dispersion multiplied by the experimental variance. We just define what is the random nature of our coefficient. So the hypothesis, we said, okay, the true value of our model minus what we have estimate because we are trying to estimate the uncertainty is the product of the matrix of the model multiplied by the difference between the real value of alpha and the estimated value of alpha. So if we would like to calculate the area where we have a given probability, we can define that the value of the normal, this is the integration of the normal distribution. So it's a very standard statistical function. So it's integration of the normal distribution, which has... Excuse me. Yes. Just one question because we don't see the blackboard. So is it possible maybe to... So what you don't see? We don't see the board, the blackboard. Yeah, yeah, yeah. So when, okay, I say, sorry, I have to present things on the screen. You're right. Thank you, Andri. Oh yeah, no worries. Yeah. So, okay, I was saying here that we have the difference between the real value of our model and the estimated value of the model. And we say that this difference will be the matrix of the model multiplied by the difference between the real coefficient and the estimated coefficient. And after using the primitive function of the normal distribution, so we can integrate the normal distribution for a given probability, we are able to calculate the surface around our point. So again, it's just for explaining things. So don't, if

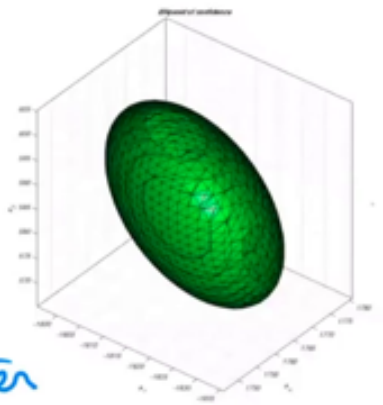
notes

summary

2.4.5 Ellipsoid of confidence

- If the coefficients α of a model $g(x_i, \alpha_i)$ are considered as Normal random function : $\hat{\alpha} \sim N\left(\alpha, (X^T X)^{-1}\sigma^2\right)$,
- Then, if η represents the true model, it can be written $\eta - \hat{Y} = X(\alpha - \hat{\alpha})$,
- And there is a probability $1 - \beta$ for finding the true value of the coefficients α within the ellipsoid defined by $(\alpha - \hat{\alpha})^T (X^T X)(\alpha - \hat{\alpha})/ps^2 = F_\beta(p, \nu)$
- If the regressors are not orthogonal, the confidence intervals of the coefficients are not independent

number parameter

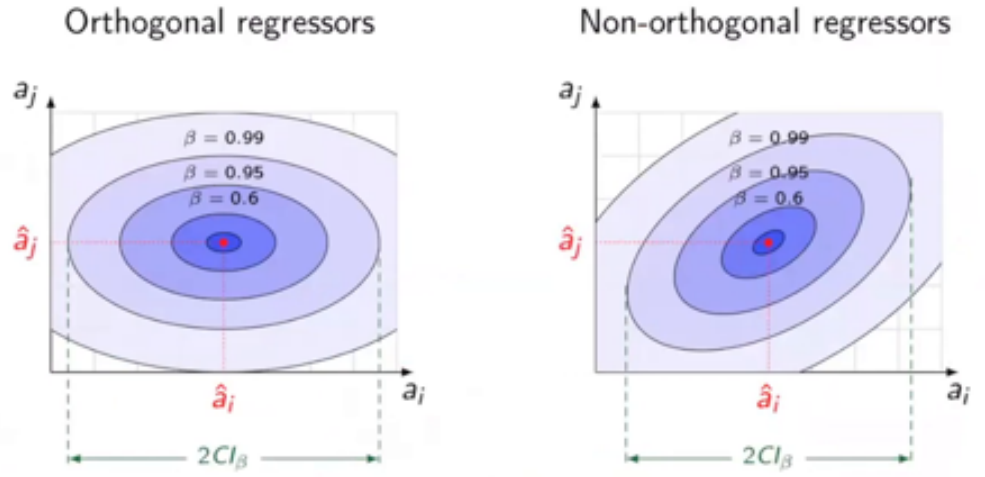


it's not very complicated, but there are points that you don't understand very well, just understand that that around one estimate, which is multidimensional, more than one coefficient, we don't have a confidence interval. We have a confidence ellipsoid. Now probably P is the number of parameters. And what I said, it's not the formula. I just wanted you to understand where it comes from.

notes

summary

2.4.6 Interdependent confidence intervals

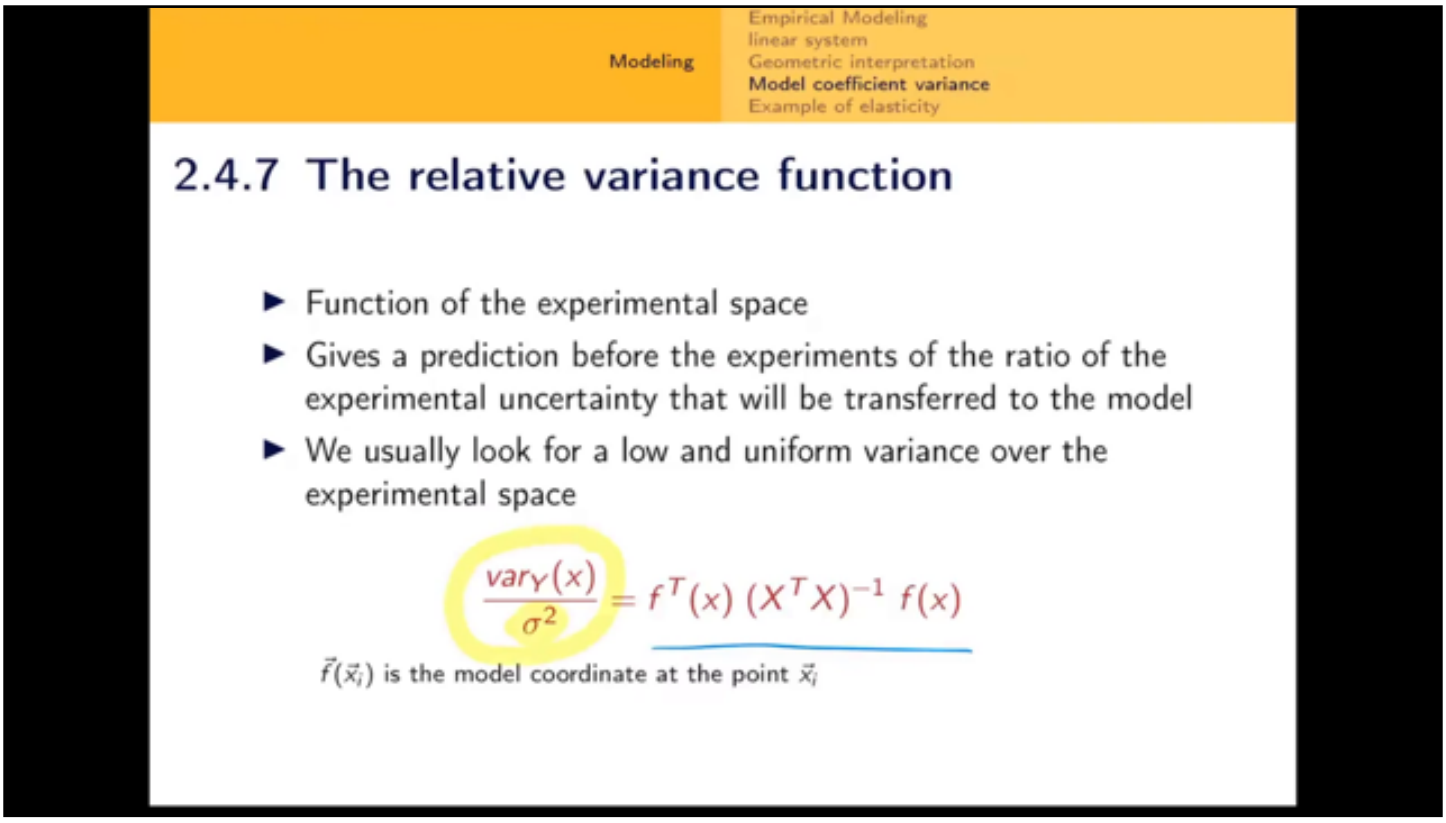


Why we say when we have a confidence interval, so if we are in multidimensional, we have a confidence ellipsoid. And now what is important is the orientation of this ellipsoid. So I wanted to make things a little bit more clear.

notes

summary

21m 1s

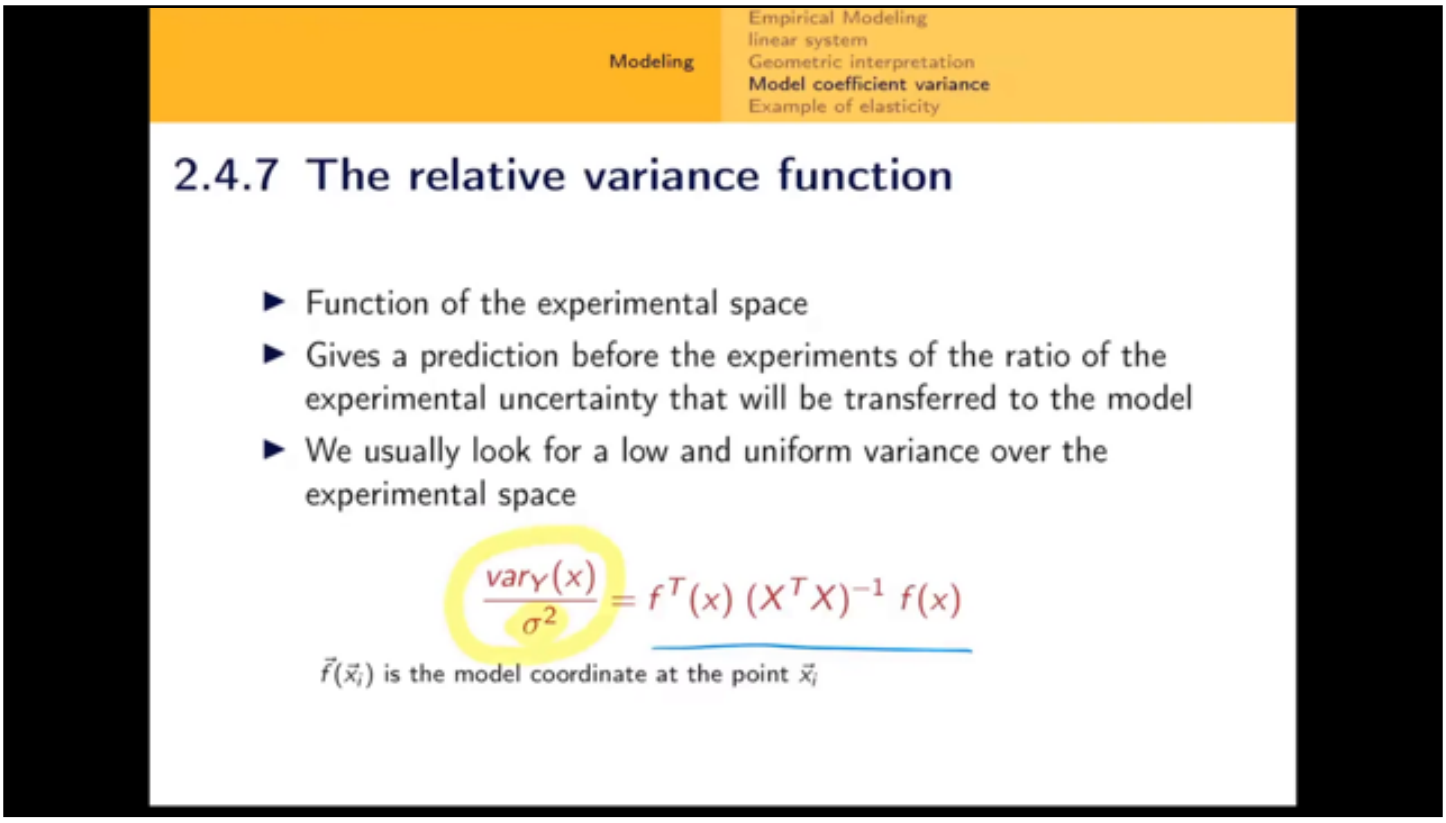


So let's make the rational in two dimension. So you have very two different situations. The first one you have what we call orthogonal regressors. That means that you have your matrix of dispersion, which is diagonal. And that means that you do not have covariance between your vectors. It's again my example, if you want to evaluate the influence of humidity and temperature on the solar cell, and you make measurements where you have quite all the time, hot and dry climates or meteo and cold and wet meteo, it will not be easy to separate the influence of temperature in front of humidity. So that means that you have covariance between these two estimates. So it will be more this situation. So when you are well and you have orthogonal regressors, we will have time to come back also in this concept of orthogonal regressors. You are very happy because your ellipsoid is oriented with your axis. So the confidence interval individually of your coefficients are clear. It doesn't depend from another coefficient. When you do not have this orthogonality, that means your ellipsoid at an angle in relation with your main axis are your factors. And so that increase numerically the confidence interval and it's make a mess. So it's why we insist a lot for having orthogonal estimators. We could not all the time. Typically, we can never have orthogonal estimators when you have a second degree. You all the time has a covariance between your first degree coefficients and your second degree coefficients. There are no charts. It doesn't exist. Okay. So the point of these two slides is every time as you have choice, try to have orthogonal estimators. It's it have a lot of interest, but one of the interest is minimizing the confidence interval on your result. So it's also explained a little bit why in this scale experiment, the experiment with

notes

summary

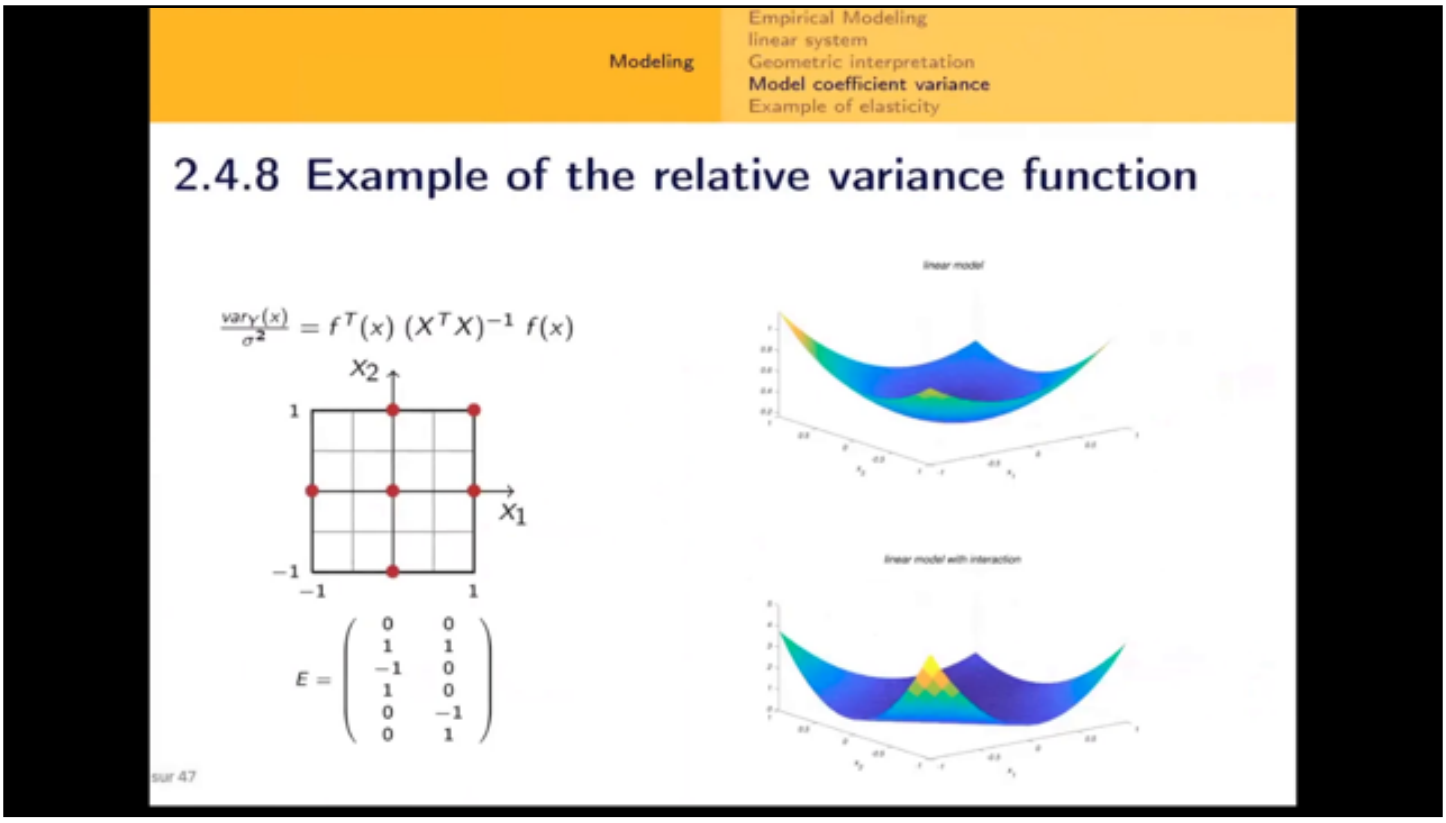
21m 17s



the special matrix at the end with all the objects where we were using was better than the one when you test each factor one after the other. It was because it was an orthogonal system and the other system was not orthogonal. And the ratio was between two sigma square and one divided by four sigma square. There are a few other things that we can calculate from a situation before experiments. One is the variance of the model. You want to estimate the model over a space and you would like to know if the quality of your model is everywhere acceptable or if they are placed in your model in your space, they are placed in your space where the estimate is worse than in other. And sorry, it's all the time worse in some parts than the other. You know that extrapolation is a very bad idea that every time that you extrapolate data, you increase exponentially your confidence interval. So you can calculate before the experiment because you see in my in my formula, at least at right, you only have X value. So it's only your decision. It's only the thing that you decide I want to make my measurement here and not here, etc. So this is known absolutely before the experiment. So I say this, we call variance function. So it's a function of your domain that tell you what will be the accuracy. But before the experiment, you cannot know what will be the accuracy of your measurements. So it's why you have on the on the left part, you have the sigma square. You don't know it before the experiment.

notes

summary

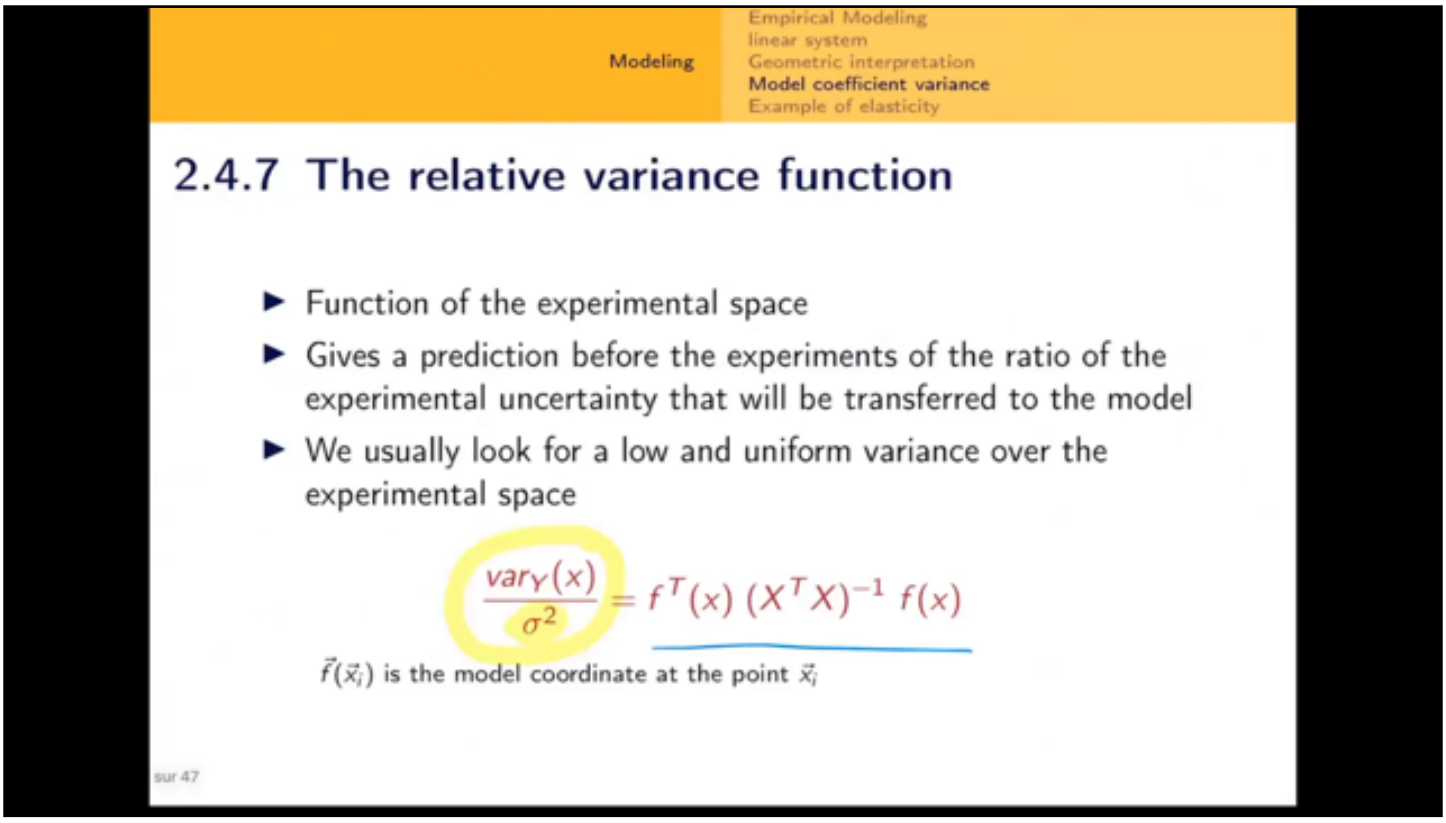


But what we said here, what you will get, you will have the variance of your model divided by the the experimental variance. So this, you know it before the experiment. Don't be trapped because you have sigma square. You don't know sigma square, but you can have only the ratio of the variance with the experimental variance. And after this value must be multiplied by your experimental variance for getting something real, the real variance in your model. And this is quite important because like that, you know where your model is okay and where your model is not okay. Probably you have a client, another process could be really a client, could be your colleagues, that need your data. You can say, okay, if you would like to have a 10% accuracy on my prediction in my model, then given a level of measurements or sigma square quality of your measurement, I can predict you if you put data measurement here, here, here, you have this type of quality and perhaps we have to double to triple the numbers of experiments.

notes

summary

26m 13s

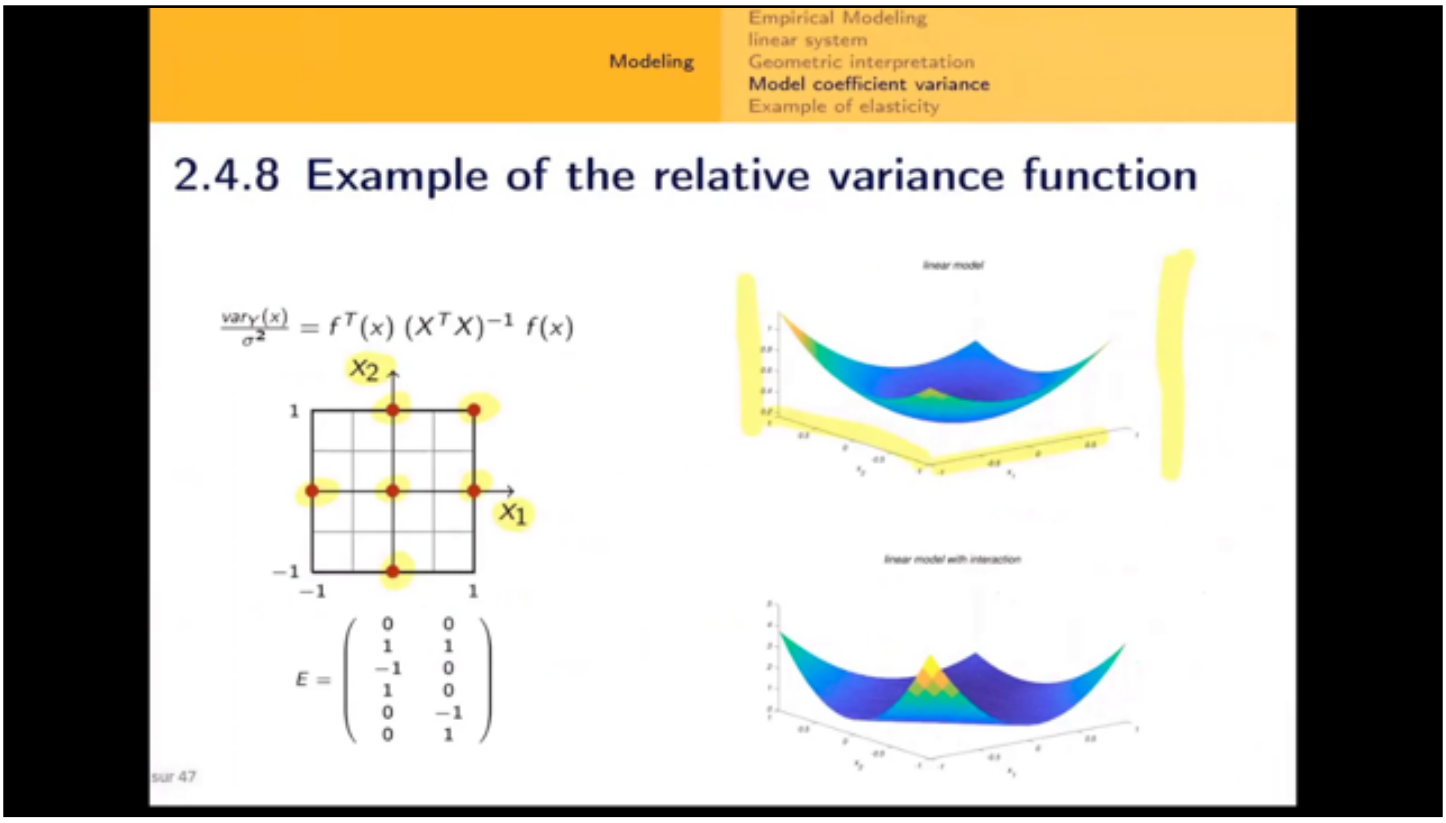


So this is an example for illustrating that. Imagine that you are interested by an experiment over a domain of two factors. So you have x1 and x2 and we are making six measurements. We are making measurement at the center. We are making experiment in the center of the edges. And for, I don't know which reason, because we have time, I don't know why, we add an experiment on a corner. So you can see the quality of what you will get will not be the same depending on your model. So it's all times the balance, what you have as data and what type of model then you can estimate. So in the first example here above, we are showing you the variance function over the domain. So the axis here are my domain x1 and x2 and vertically it's the value of this, of

notes

summary

27m 25s

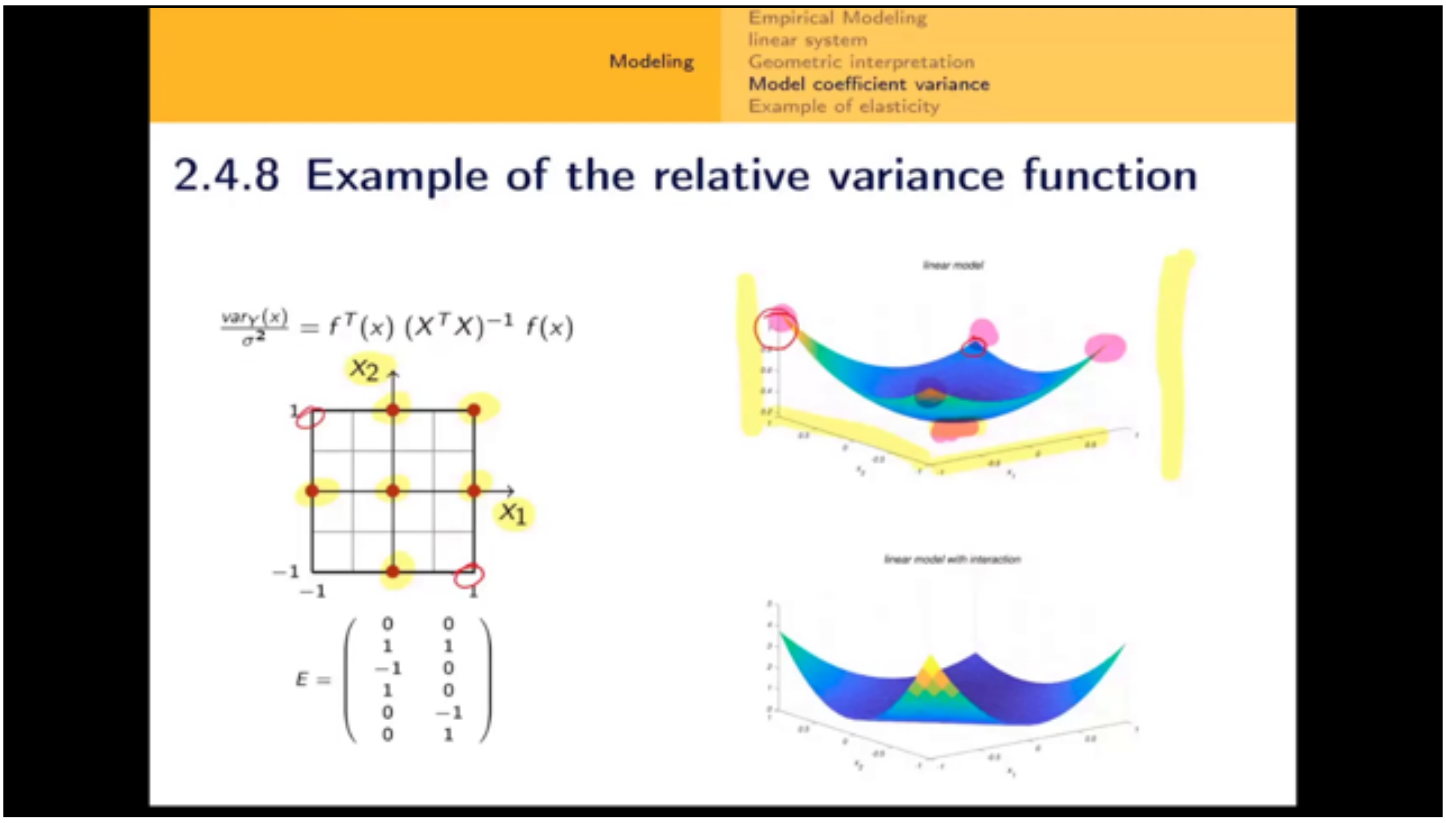


variance divided by sigma, because I don't know the sigma as a sigma square yet. I didn't have made the experiment.

notes

summary

28m 42s



Again, I know that before making the experiment. So it's interesting because here is not sufficient pressure. So I can add some data here because I need to have a good. What we usually look for is a flat surface because we would like to have a predictive quality of our model all over our domain. But nevertheless, the variance is a quadratic function. So it's never flat. It's impossible to have it flat. So the question would be probably what is the value at the center of your domain, which will be probably the place where you have the most information, but more of that you would like to understand what is the value of your model at the quality of your model, the uncertainty of your model at the limit of your model and probably the corner is the worst place. And you can see that in the limit of your domain, that there are a place, this one, where the model is more precise. It seems that the value is something as point point six. So you will have the variance of your measurements that you have to multiply by point six or quite half of your accuracy of your measurement as predictive quality of your model is that place. That's what you want. What is the worst? What is the place where the model has the worst quality? And you see here in this example, so in the corner minus one, one, so this one, one, it seems that it's this corner or this one. The variance coefficient is one. That means that you have the experimental variance. The variance of the model will be sigma square there because it's all time multiplied by the quality of your measurement. I don't know what the client wants. It is, that's the ant also your measurement, imagine that your quality of your measurement is around 1% if you

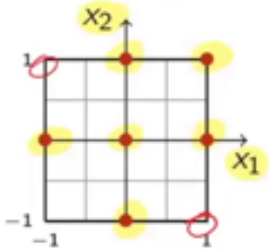
notes

summary

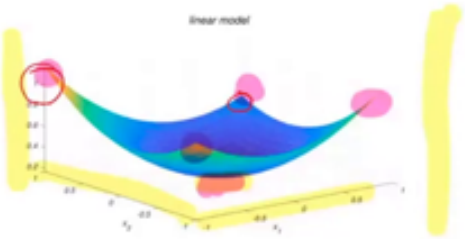
28m 51s

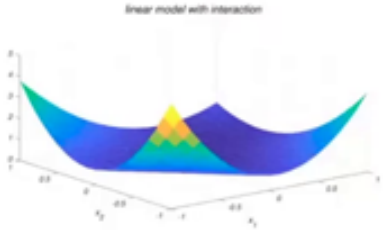
2.4.8 Example of the relative variance function

$$\frac{var_Y(x)}{\sigma^2} = f^T(x)\,(X^TX)^{-1}\,f(x)$$



$$E = \begin{pmatrix} 0 & 0 \\ 1 & 1 \\ -1 & 0 \\ 1 & 0 \\ 0 & -1 \\ 0 & 1 \end{pmatrix}$$

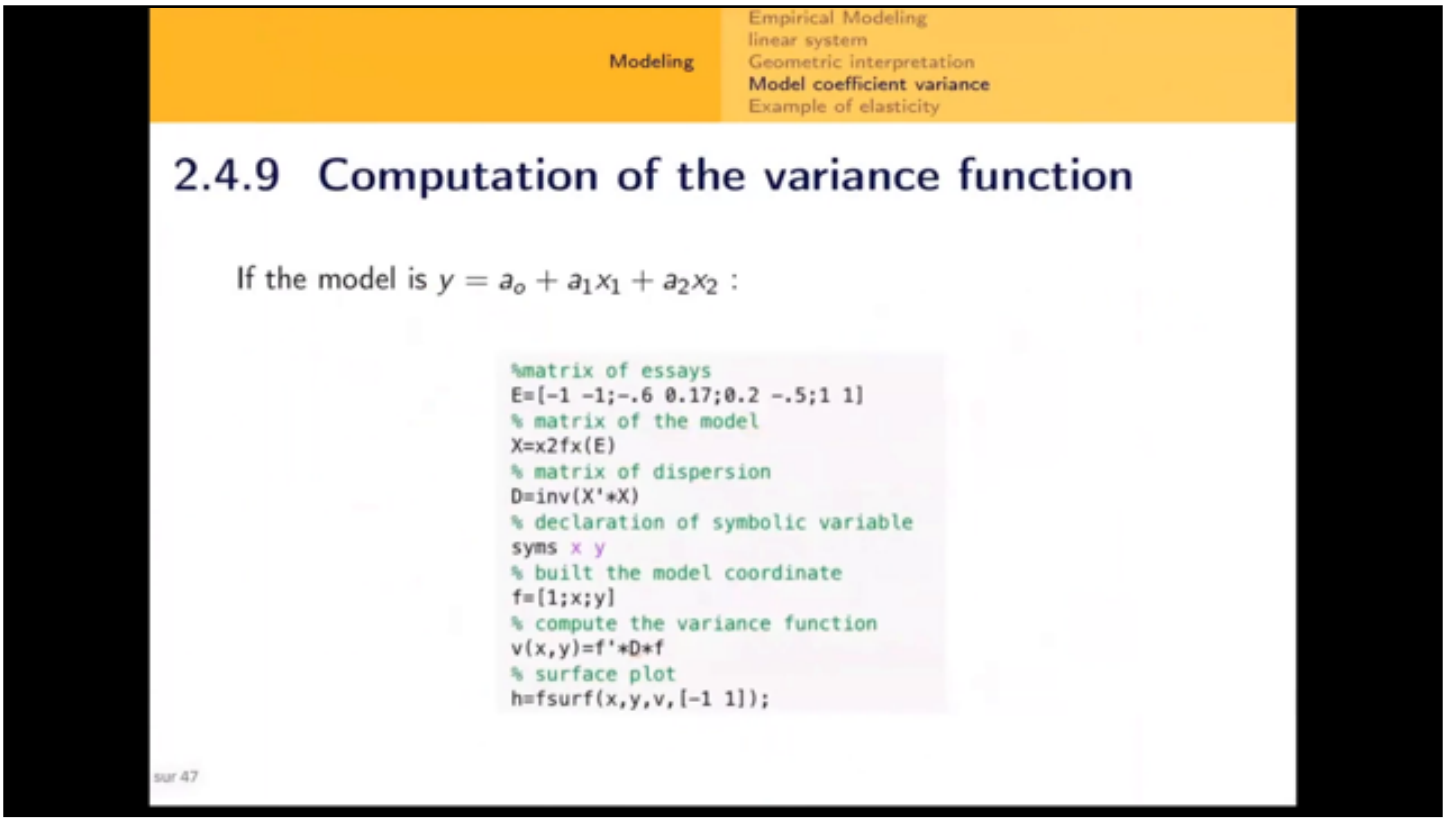




are happy that you can check before the experiment. or you have a very unprecised movement, say, okay, I better double or multiply by or by 100 the numbers of experiment. You know where?

notes

summary

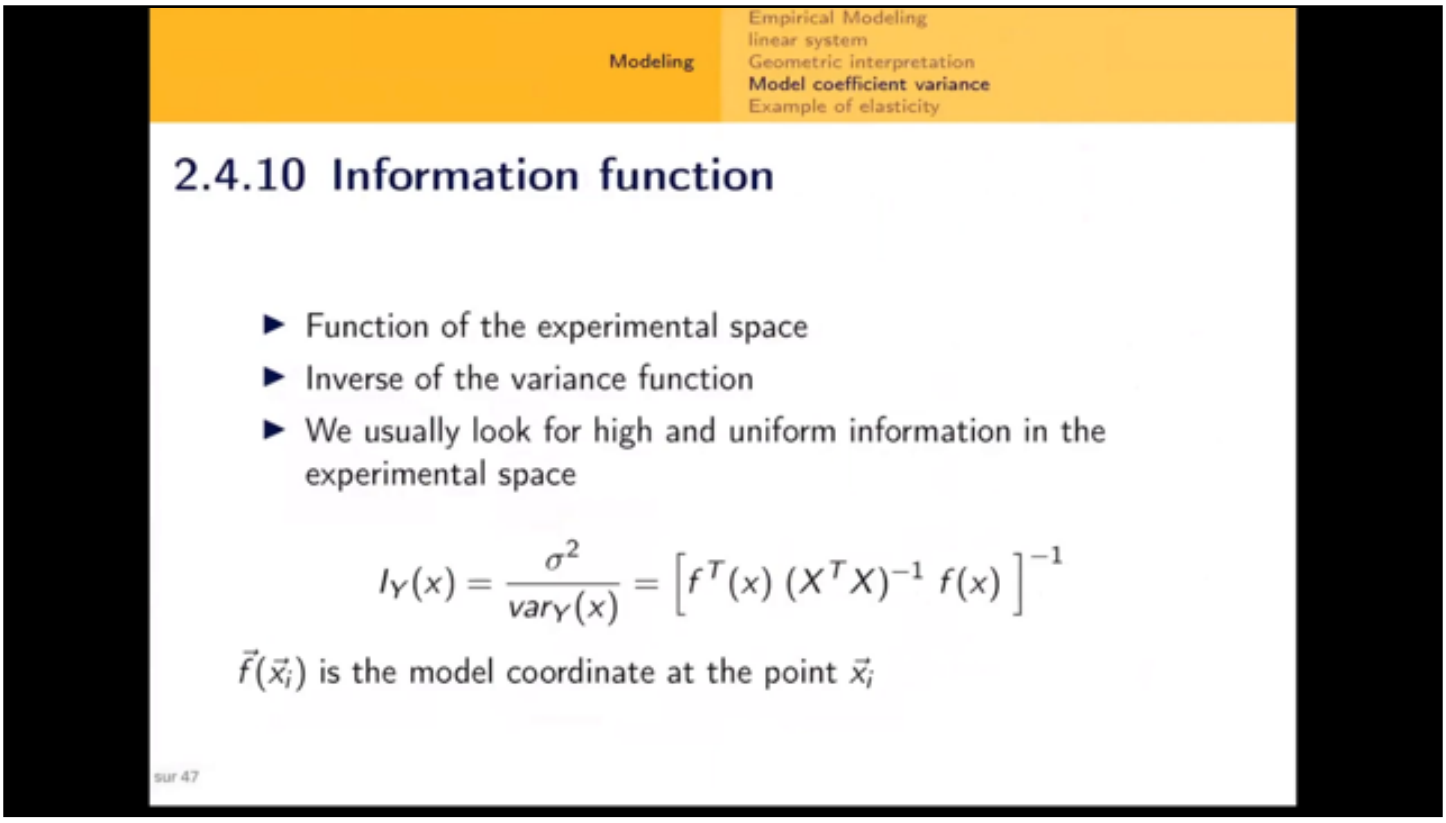


After you also see if you have a client, you have also to tell me where you would like to estimate the model. If you are interested in a linear model with interaction, with the same information, you see first that you have multiplied a lot the extreme value. So now it's around 3.5. So you see that with the same data, depending on the model that you would like to estimate, the predictive quality will not be the same. What is this question of do I consider the linear aspects of the model? So, okay. So this is why it's very interesting to check that. So in most of the cases, quadratic function, and usually we check just the limit because we know the shape. You could have a very complicated model, perhaps you, or if you have data everywhere, you have not applied a regular design and you have data everywhere, you could check where it could be interesting to put an additional measurements. Trying to get a better quality all over your surface and not only in one form. So to remember here, there are things we can calculate before the experiment. And the second thing, the variance function, so the uncertainty on your model depend on the model, depend on the experiment, depend on what you have. That's for different model, for the same data, you have different uncertainty.

notes

summary

31m 37s



Here you have the way to compute it. I try to convince to use the symbolic calculation. So I should present you the calculation without symbolic, so it will be 10 times longer because you have to make a lot of calculation of loops and et cetera. But so if you want to calculate this variance function, it's very easy. So you see, well, first, oops, no, the right, okay. So here I've just have decided the matrix of S.A. So just where I put my points of measurement. After I have calculated the model matrix for a given model, in this case, it was a linear model because I say no, no keywords. So when you have no keywords in F2Fx, it's a linear model. Or it could be interaction quadratic. After I have calculated the dispersion matrix, and after I start with a symbolic calculation. So I define two variables, x and y, as symbolic variable. I have built my model coefficient. So it's this vector with one in the first coefficient for the constant. And after the coordinates x and y, because I have a linear model. And if it would have been a quadratic function, so I have to add more. And look, the calculation is just as you write it to the variance function. I just write what was the equation. F transpose multiplied by the dispersion matrix, multiplied by F. And then I just have to use a function for graphing a surface with a symbolic function. So it's very easy, very practical to use symbolic calculation. If you have to make the calculation, you have to discretize your space, you have to do a lot of other things.

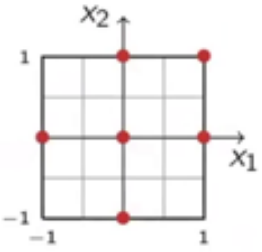
notes

summary

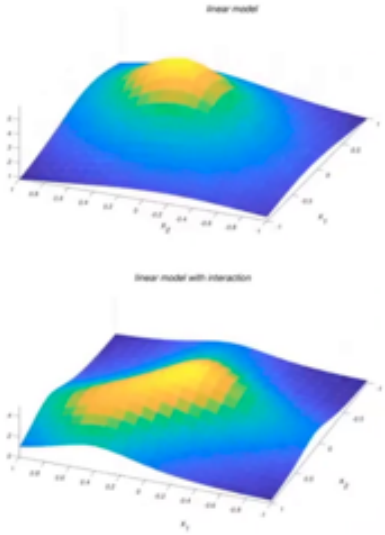
33m 22s

2.4.11 Example of the information function

$$I_Y(x) = (f^T(x)\ (X^TX)^{-1}\ f(x)\ )^{-1}$$



$$E = \begin{pmatrix} 0 & 0 \\ 1 & 1 \\ -1 & 0 \\ 1 & 0 \\ 0 & -1 \\ 0 & 1 \end{pmatrix}$$



sur 47

You have a function which is quite the same. It's just the inverse of the variance. And the inverse of the variance, we call it information. And so it's just a formula inversing the value.

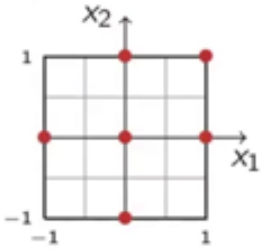
notes

summary

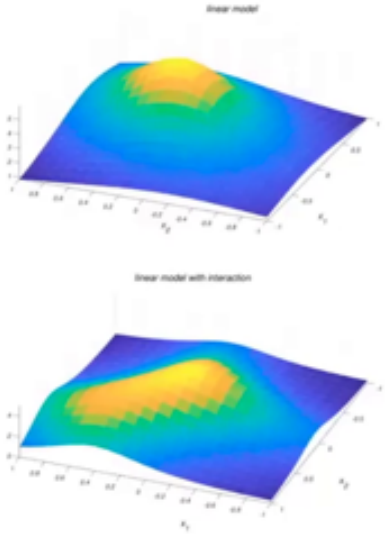
35m 30s

2.4.11 Example of the information function

$$I_Y(x) = (f^T(x)\ (X^T X)^{-1}\ f(x)\ )^{-1}$$



$$E = \begin{pmatrix} 0 & 0 \\ 1 & 1 \\ -1 & 0 \\ 1 & 0 \\ 0 & -1 \\ 0 & 1 \end{pmatrix}$$



And it gives you not where you have uncertainty. It tells you where you have the information. And it's the same information, just one's the reverse of the other. And you understand that where you don't,

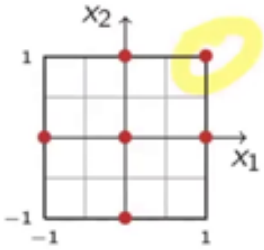
notes

summary

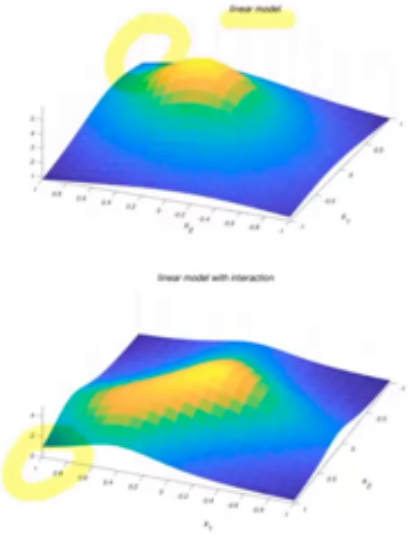
35m 45s

2.4.11 Example of the information function

$$I_Y(x) = (f^T(x)\ (X^TX)^{-1}\ f(x)\ )^{-1}$$



$$E = \begin{pmatrix} 0 & 0 \\ 1 & 1 \\ -1 & 0 \\ 1 & 0 \\ 0 & -1 \\ 0 & 1 \end{pmatrix}$$

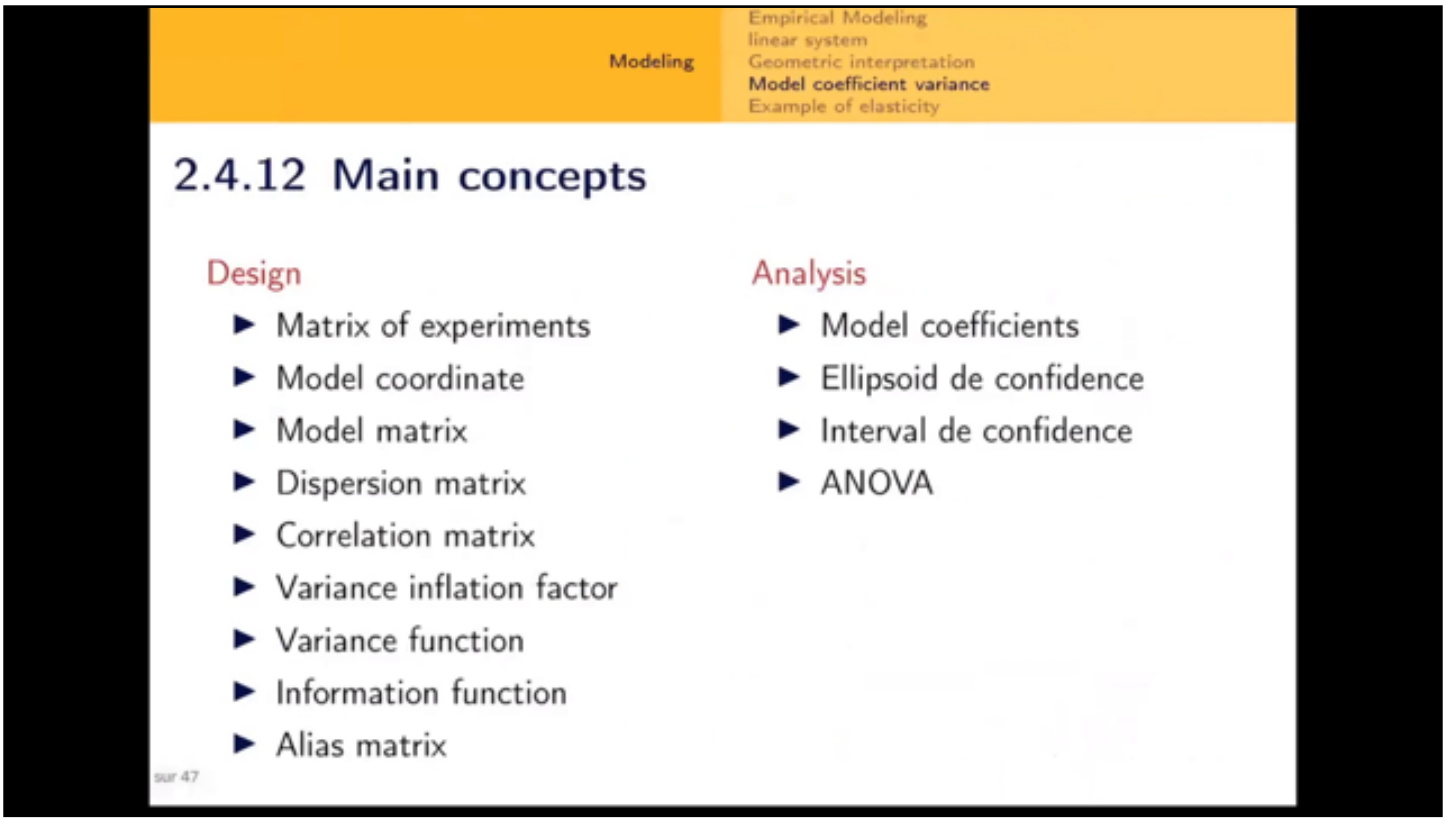


when you have data, you are increasing the information. But not only locally, because you have conjunction of information coming from other points. So for the linear model, you see that you have, I don't know, a hill of information, quite by the center, but a little bit displaced on the corner where I have my additional points. It's not that I do not have a Dirac of information where I have information. I have my information diffusing around. And just be careful because my access are not the same in the two picture. So now this point is here. It's a quite, I didn't pay attention

notes

summary

36m 1s

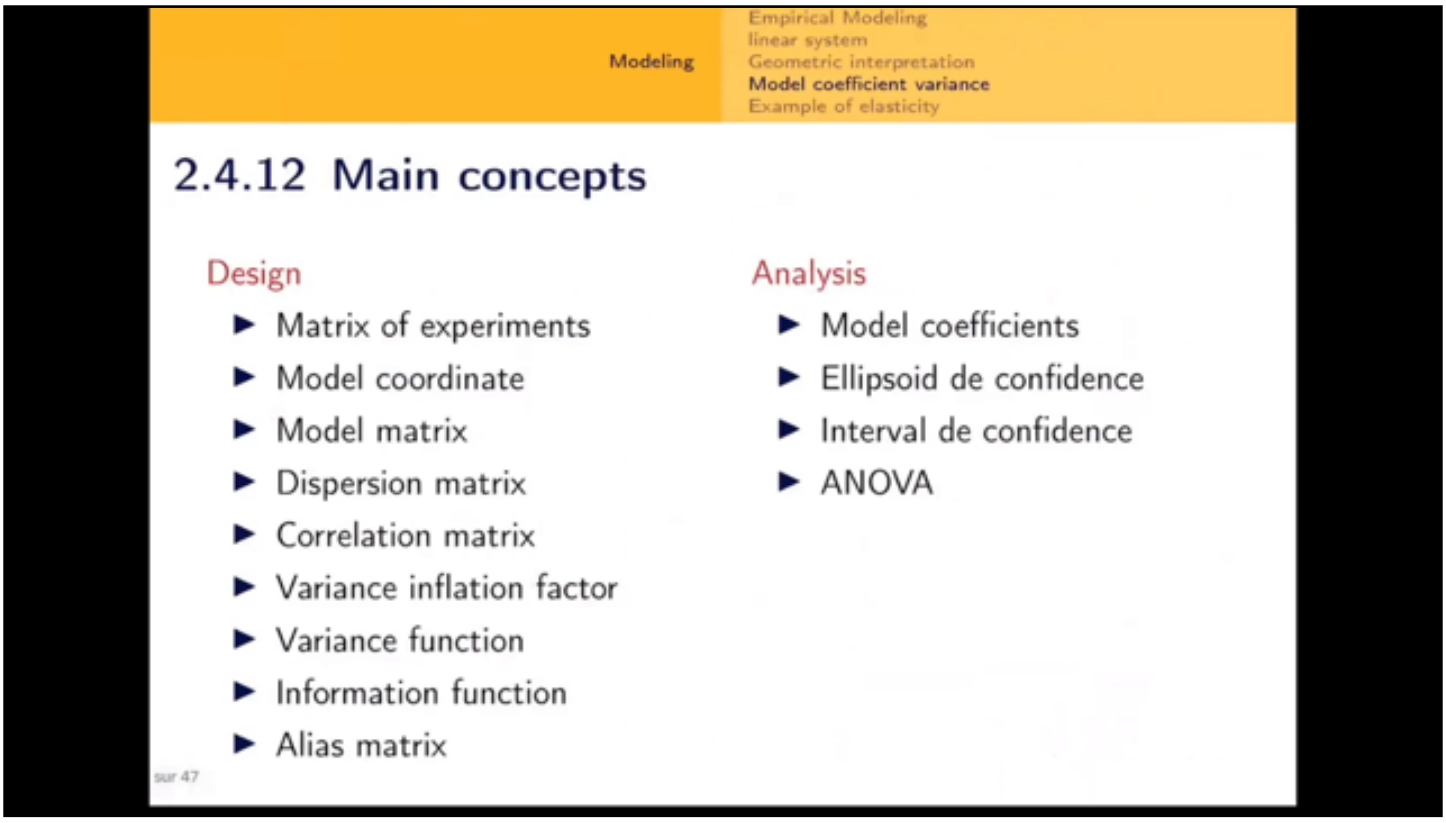


when I make my graphic. If you see the access, the plus and minus axis of one of the axis that they've been inverse. So you see that adding a point in one of the corner has put a lot of information as improve a lot of information for my model with instructions. It's clear because if you just do the first one, you are unable to calculate interaction coefficient. But you understand that you know quite well the interaction in this region from the center to the corner, but you don't know well the interaction information outside of this domain. So this give you a sort of information on where do you have information in your domain. So if you want to improve your model, it's clear that you have probably to put additional points in the blue parts of the space. Very easy at two dimensions. So if you are in five dimensions, okay, but you can understand the principle. Okay, so again with my Matruszka approach, so we see a few of elements. So I would like that out here. Matrix of experiments, all your factors in columns, all your experiments in rows. And model coordinates, you define your model and you are able to calculate some pseudo vectors, just representing the model and easy to manipulate in equation. Model matrix is depending from the matrix of experiment and of the model. So you mix both together, both information, where do you have data? What is the model that interest you? So when you decide an experiment, you decide a model of experiment. The model of matrix belong to the interpretation. You can, the same data, you can interpret it as linear model, as a second degree model, quadratic model, exponential model, any model that you want. After when we have a model matrix, we can calculate the dispersion matrix. This matrix is not X transpose X

notes

summary

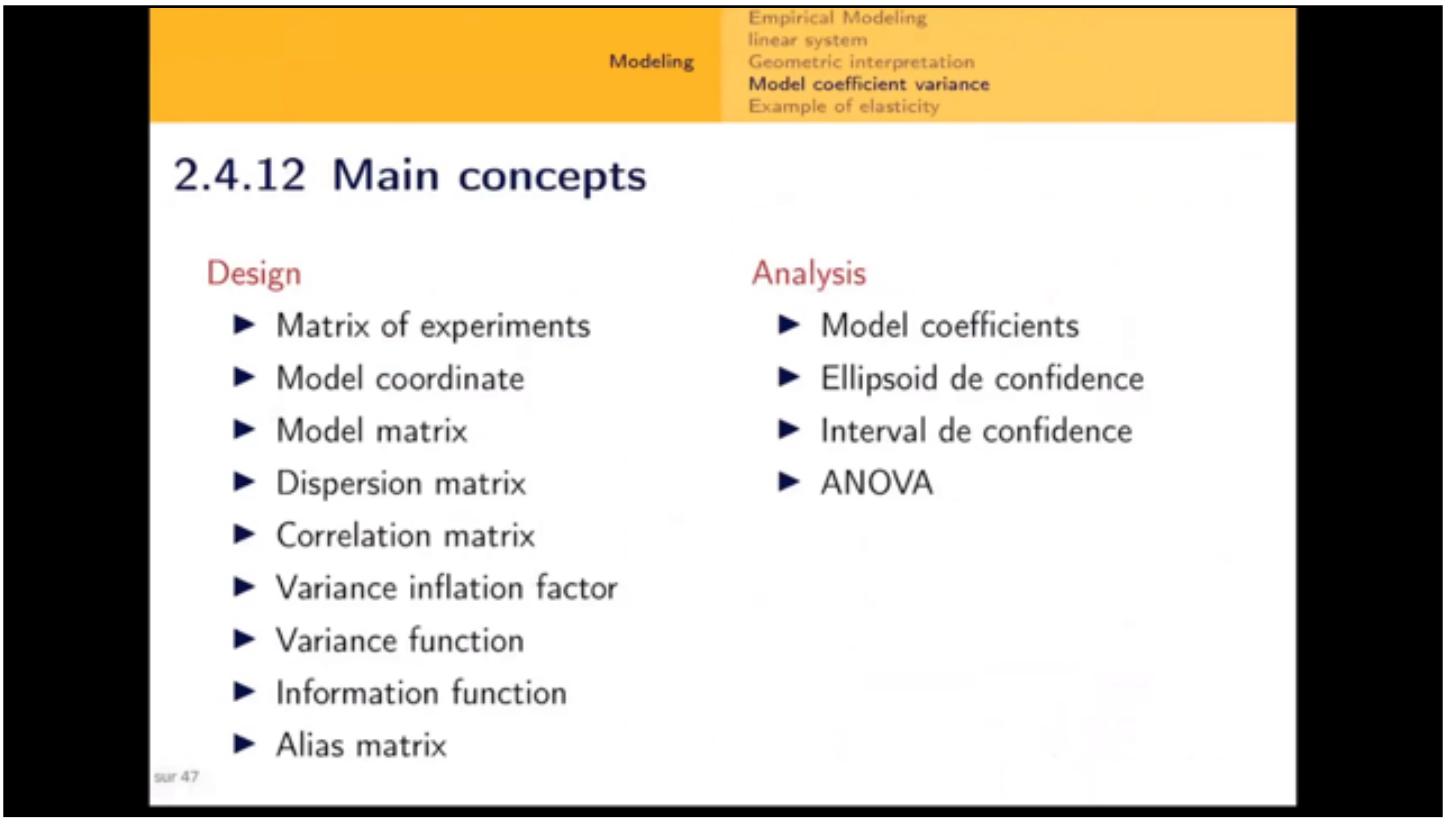
36m 49s



minus one. Both information X transpose X is called the information matrix. So it's all the time between dispersion information, the reverse. So this matrix tell you how you will transfer the uncertainty within your fit. You can calculate the correlation matrix, but mainly the variance inflation factor. So I tell you, so it's just the standardization of the dispersion matrix with the individual variance coefficient. Not interested really the correlation matrix. What interests me is the variance inflation factor. Are the diagonal element of the inverse of the correlation matrix, but what is it is not so much clear, but it's bringing you information on the consequence of collinearity between your factors. So it's bringing you the consequence of collinearity is the opposite of diagonal orthogonality. So if it's collinear, it's not orthogonal. You understand that one is the opposite quality of the opposite quality. After now I present you the variance function and the let's choose this. So the variance inflation factor, the variance function, the dispersion matrix information are elements that could let you guess before making an experiment if you have a good design or not a good design. The first one, honestly, most the most things I check is the dispersion matrix. I see already a lot of problems is diagonal, is not diagonal, the coefficients are over one. That means I'm multiplying the experimental error or I'm diminishing the experimental error. And after the other could be interesting more locally to understand where I have a problem in my domain. We still have to check the alias matrix. You don't know yet. The model coefficients is the alpha. So you know what they are. The ellipsoid of confidence, the interval of confidence I mentioned and there are still we have to talk about another. So we are still two concepts in this Matruska approach to have quite all the element necessary for juggling with

notes

summary



these concepts of modeling.

notes

summary