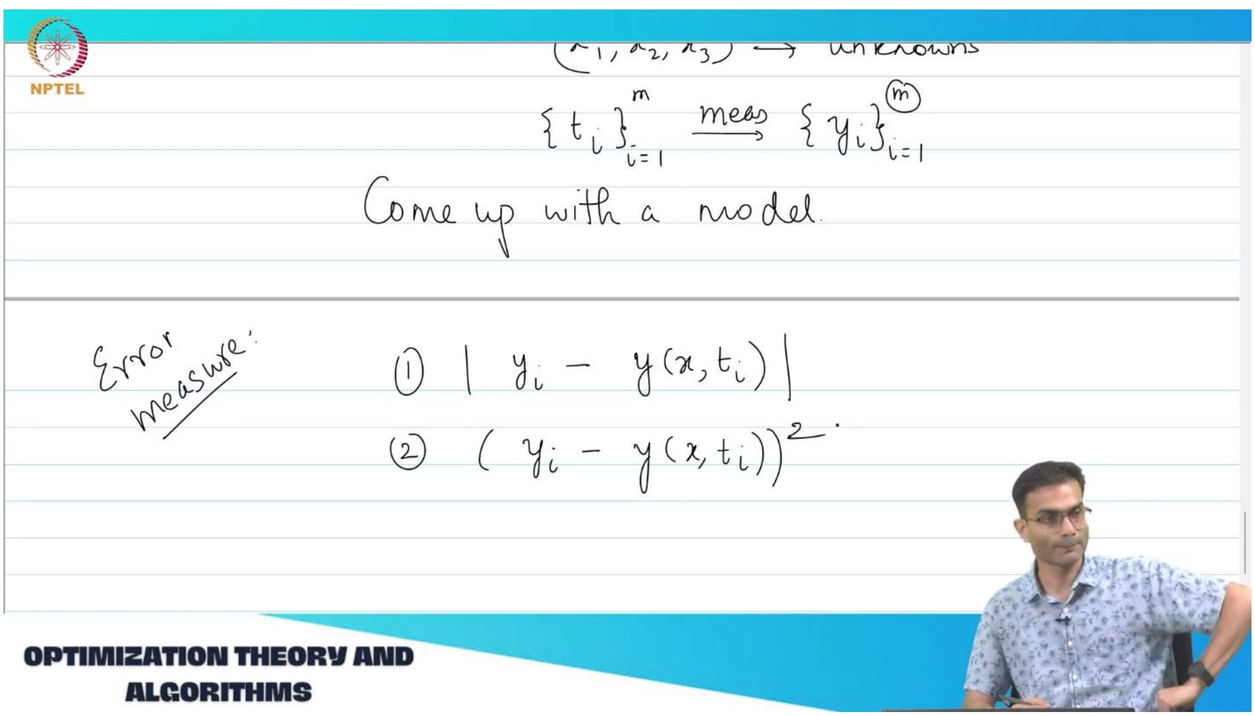
Course Name: Optimization Theory and Algorithms
Professor Name: Dr. Uday K. Khankhoje
Department Name: Electrical Engineering
Institute Name: Indian Institute of Technology Madras
Week - 08
Lecture - 55

Least squares problems

So, now we are going to switch gears a little bit. Instead of introducing yet another method, we will finally take a kind of an application where some of these methods would get used and the application that I will pick is something which is most commonly found in engineering. These are called least squares problems, ok. So, for those of you reading the book, this is chapter 10 of Nocedal and Wright. So, let us start this discussion with an example, a very, you can say real life kind of an example. So let us say that there is some, for example, a nuclear reactor in Japan, which is melted down, right? And you know, you go with a Geiger counter and you want to measure what is the kind of radiation coming from there and you want to model it. So, you can see that you know if I can get a model of it, of the radiation coming from a, you know, destroyed reactor that will be very useful because if that model is accurate you can predict until what time is it going to be unsafe, until what distance is it unsafe and all those kind of things.



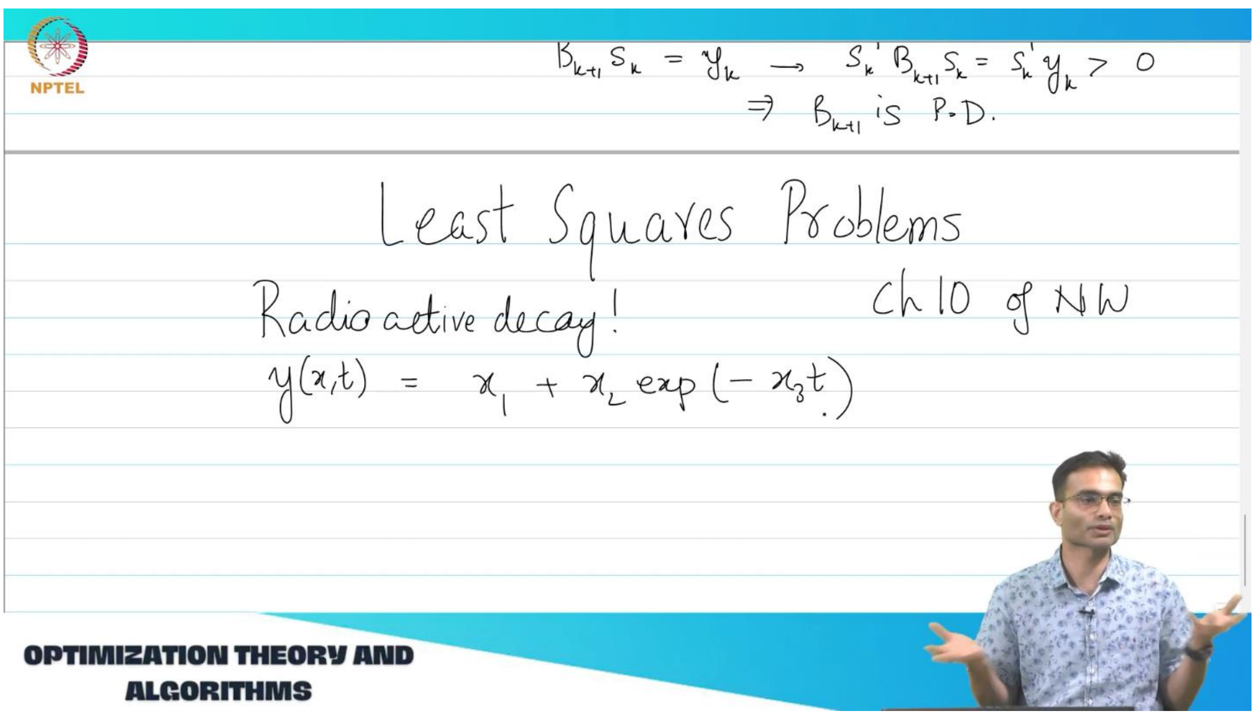
So, these, so this model for example could be some very simple like

$$y(t) = x_1 + x_2 e^{-x_3 t}$$

So here x is not space, x is some parameter that I do not know. So what is it looking like? We know that radiation decays exponentially, right? This, if you remember the very first lecture of this course, we said there is an optimization begins with a modeling phase. And that modeling phase does not come from optimization, that comes from your core expertise whether it is nuclear physics or chemical engineering or whatever, whatever, right? So, the nuclear physicist said that radiation has an exponential decay. Now, when I say exponential decay, immediately without being a domain expert, this is the most general model I can write.

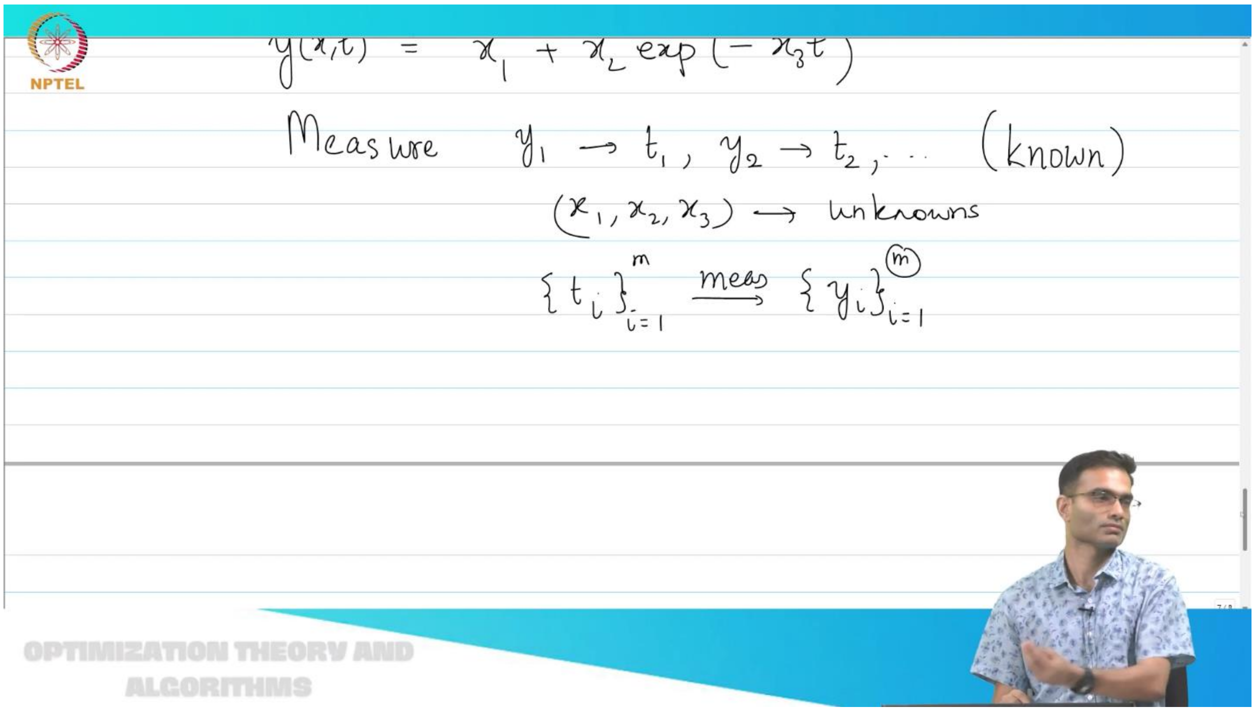
There is some offset at time t, I mean there is some offset, I call that x_1 and there is an exponential decay which has some weight x_2 and the rate at which this decay, that is yet another variable, that is x_3. This is about as general as I can write. Now, what is going to be as someone who knows engineer optimization, what would you want to do? We will go with your counter. You will make several measurements, right. So, you will measure let us say y_1 at what time t_1, then you will measure the radiation y_2 at time t_2, ok and so on.

So, these are the knowns of the problem. Now, as I am making these measurements, what are the unknowns? x_1, x_2, x_3, do they change with time? Hopefully yes. So, the more general way of writing this is that at times t_i, $i = 1$ to m measure y_i, so I get m measurements. Now, having gotten these measurements what is our job? Give me the model that explains the measurements in time, right. So, come up with a model. Now supposing I do come up with a model, what would be a sensible way to test how good the model is? Like what is the measure of goodness of the model? No, I am not asking how to solve the problem. I am asking you having gotten a model, let us say you got a model, how do you, how would you assess whether this is good or not? Does the model prediction match the measurement, it will not.



So how do I then quantify, I mean in real life problem it will not match, why? Experimental noise, measurement noise is always going to be there, you will never get exact match, that is obvious, right. So now that it does not match, how do I quantify this error? I could distance, exactly. I need some measure of distance between the two, right. So, for example, I could do, so error measure, I could do for example, I could compute this for all j's. Is there any other thing I could do? The hint is the title of the section.

I could square it, I could look at this measure, I could look at $y_i - y\big(x(t_i)\big)$. Any advantages or disadvantages? First function mod, the modulus function, is it differentiable? There is a problem at the origin, right? So if I start doing calculus, I may end up with a problem at the origin. The second one, it is nicely differentiable as long as y is differentiable, right? So that is one practical reason why I will choose this, okay? So this is to test the, how good my model is. Now how do I come up with the model itself? I have my measurements right and now in the language of optimization how would I write it?



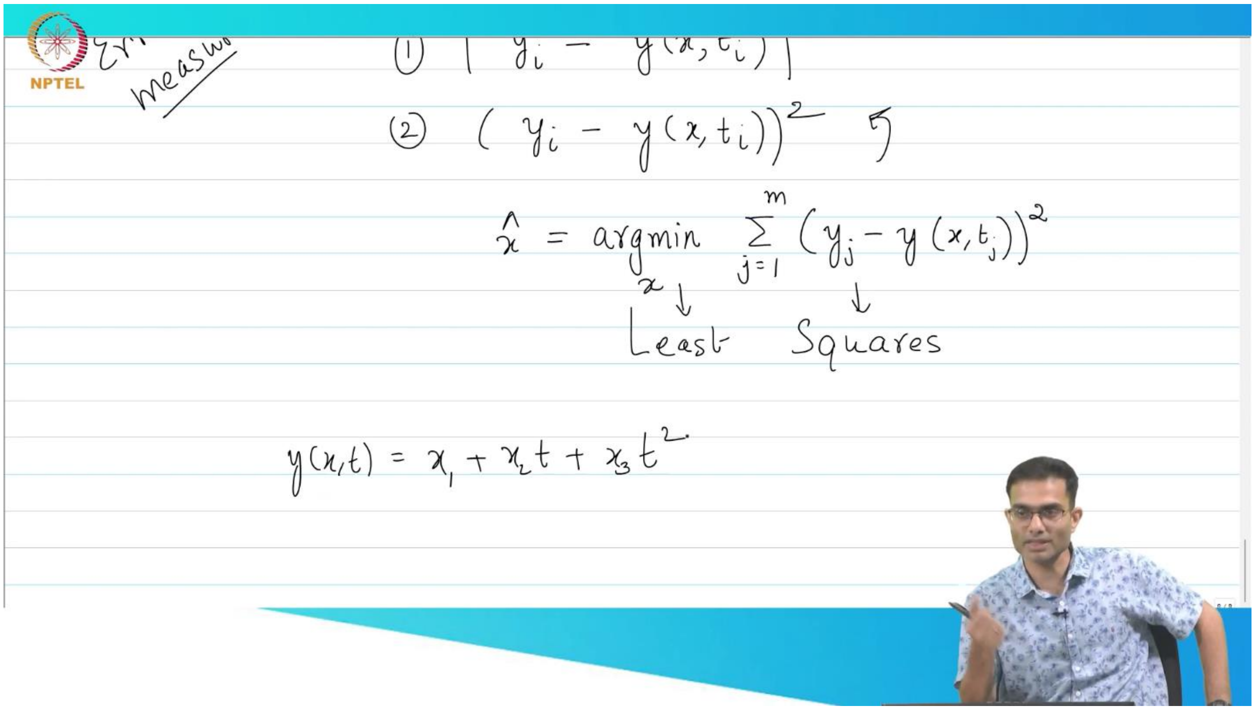
All you have are the measurements and we said that square of errors is a good way of measuring something. So language of optimization I would want to what to these sum of squares? minimize them, right? So, I would say that for example,

$$\hat{x} = \underset{x}{\operatorname{argmin}} \sum_{j=1}^{m} \Big(y_j - y(x, t_j)\Big)^2$$

So, if I find that x, so let us read this in plain English. Argument means that x which is the variable of optimization which, so it minimizes the argument of this right, that is my $\hat{x}$. So I want that x which does this minimization. If I get this, I am done. This is the $\hat{x}$ which explains the data very well.

Why does it explain the data very well? Because when I insert that $\hat{x}$, this sum of squares is minimized, which is what I want. So you see in a very, very natural way we have ended up with trying to minimize a sum of squares. So that is why these problems are called least square minimization problems. And this is virtually I would say 90 percent of engineering is least squares optimization. Whether or not you people realize it formally or not that is what they are doing.

So, this is hence the word comes least from here and squares from here. Now there is another small distinction that is made over here in least squares problem. Like we looked at conjugate gradient method. There were two flavors of conjugate gradient. What were they? linear and non-linear.



Similarly, they are going to be two flavors of least squares problem, linear and non-linear. This real life problem which I have written over here, radioactive decay from $x_1 + x_2 e^{-x_3 t}$. Is this linear in the variable? Which is the problematic term? x_3. x_3 which is sitting in the exponential is not linear with respect to the y with respect to the variables of the problem, right. So, this is an example of a non-linear problem, ok.

So, this is a non-linear, sorry, non-linear least squares. This is a non-linear least squares problem. This is of course going to be a bit more difficult to solve than a linear least squares problem. On the other hand, let us say that my model was something like this. This could for example come if I am trying to model, you know, distance covered by a ball that is dropped.

Maybe it should be something like this, right, $u + at + \frac{1}{2} bt^2$. So, it will be a quadratic. So, is this, is this linear in t? It is not linear in t, but we do not care about t because t is going to be my measurement, t is going to be given to me at time 1.1 second, I measured this, this thing. So in terms of the variables x what is it? It is linear.

It is x_1, x_2, x_3 they are appearing by themselves. They are not multiplied by each other, they are not in log or exponential or whatever. So this is a linear least squares problem. Right, so this is as far as the introduction to at least the motivation of least squares problems course, okay. In the next class we will look at having, now we have done the formulation of this, right.

Once we have done the formulation, obviously the next step that we will do will be calculus on this, so that we can start our journey of minimization, coming up with a solution and then solving it. We will get two different cases for the linear and the non-linear and we will look at what happens over there, ok. So, those of you who are currently solving a research problem for your thesis, for your UGRC, for whatever right, go back and try to look you with high probability you will come across a least squares problem somewhere in your research, ok. Alright, so I shall see you folks next week.