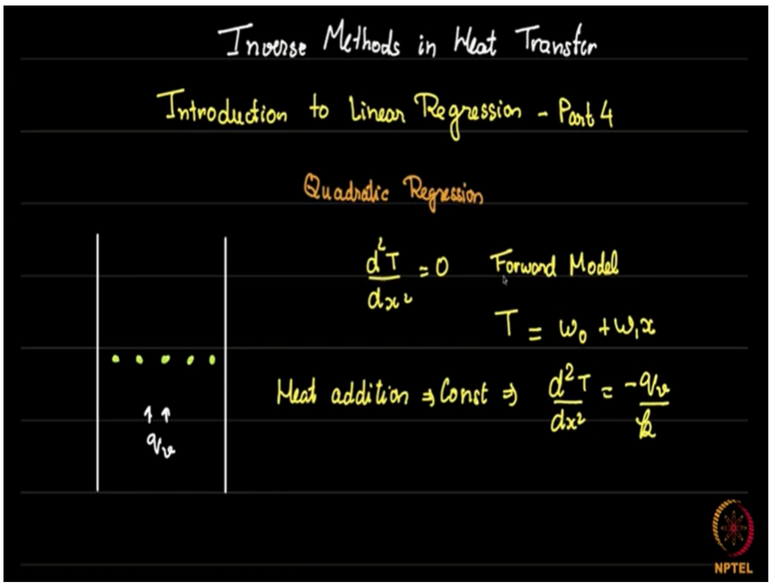
INVERSE METHODS IN HEAT TRANSFER

Prof. Balaji Srinivasan
Department of Mechanical Engineering
Indian Institute Of Technology, Madras

Lecture No # 09
Module No # 02
LINEAR REGRESSION WITH QUADRATIC MODEL

(Refer Slide Time: 00:18)



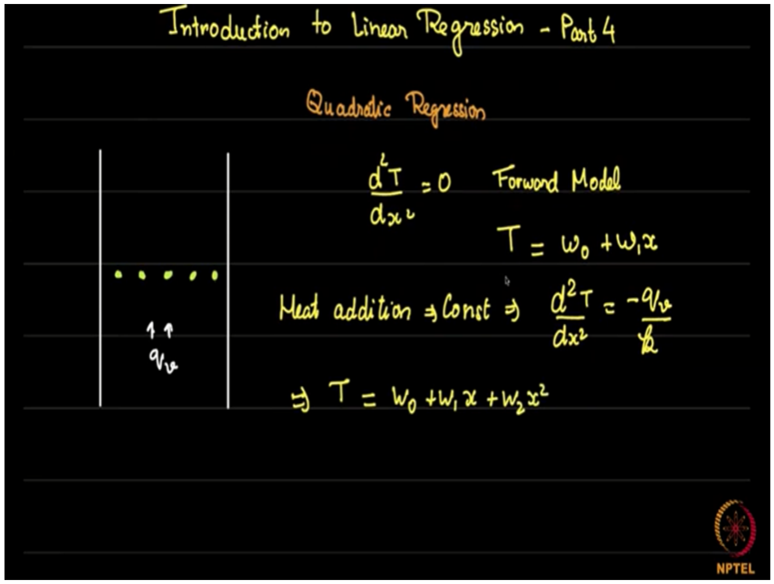
Welcome back. in the last few videos, we saw how to use linear models, so the models that we had been using but because we were dealing with steady state heat conduction without heat addition, we had

$$\frac{d^2T}{dx^2} = 0$$

So, our forward model which was from physics said that the temperature would be of the form of a constant plus some other constant multiplied by x. Now let us say we have heat addition and let us say constant heat addition. So, this would give us something like, is some heat addition.

$$\frac{d^2T}{dx^2} = \frac{-q_v}{k}$$

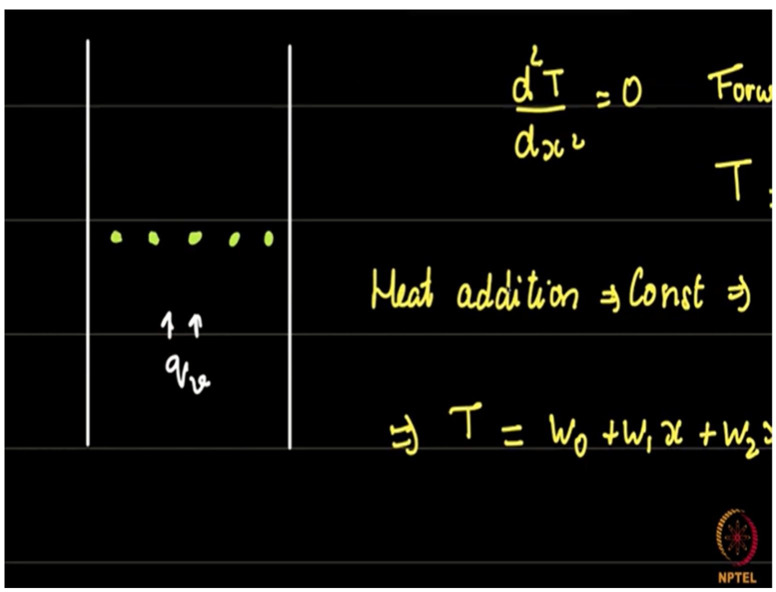
(Refer Slide Time: 01:17)



So, you would have a heat source term on the right-hand side. So, the solution to this invariably is going to be a quadratic so something like,

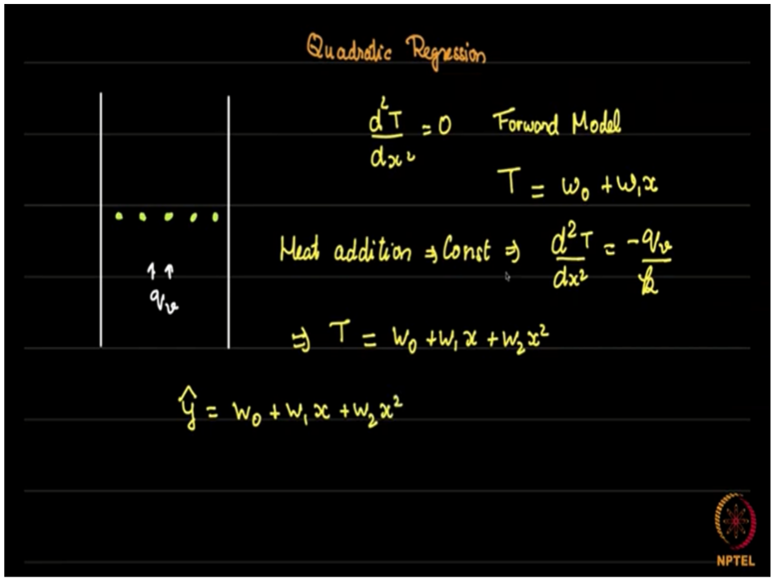
$$T = w_0 + w_1 x + w_2 x^2$$

(Refer Slide Time: 01:37)



So, you can imagine a case very similar to the one that we have done, where you have a slab, some thermocouples somewhere in the middle, but you have heat addition and in fact, we have given one such problem within the assignment for this week. So, if we have something of that sort and we have a similar problem to the one above like, I give you the temperatures and I ask you to find out what this heat addition is or what the conductivity etc. then how would you go ahead and proceed with that.

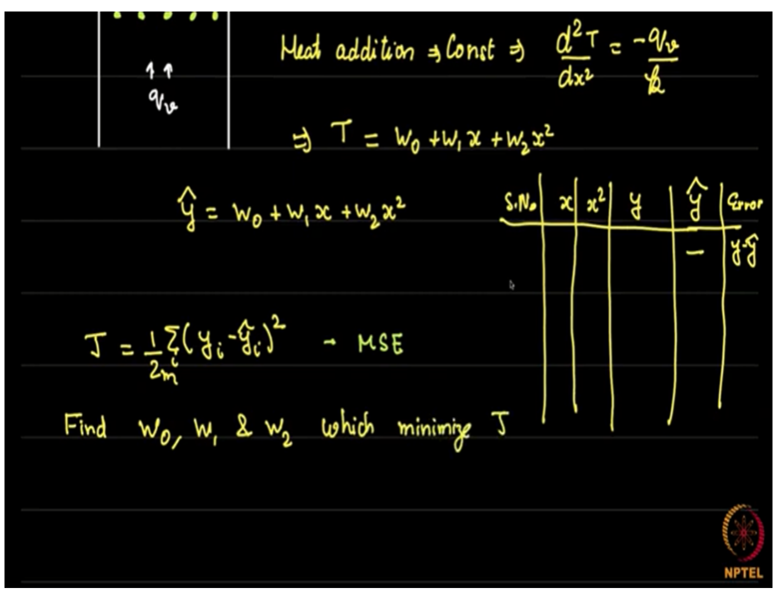
(Refer Slide Time: 02:02)



As it turns out you need what I am going to call right now quadratic regression, more specifically We will see next week that all these are just special cases of linear regression. in fact, you might probably suspect that as I do this problem that this looks just like linear regression. But let us start with this model. let us say we have a case not just the slab but some case where we have,

$$\hat{y} = w_0 + w_1 x + w_2 x^2$$

(Refer Slide Time: 02:36)



The rest of the process remains the same in that, we still make a table. So, we have a serial number, we have x we have y, and then we have $\hat{y}$ and maybe some of you might prefer to even put an extra column for x^2. So that is just calculated. So, once we do that, we basically have to calculate $\hat{y}$ and we have to calculate the corresponding error which is $y - \hat{y}$. And then we calculate J just like before the cost function is still the same thing whatever be the prediction sum up half of $(y - \hat{y})^2$.

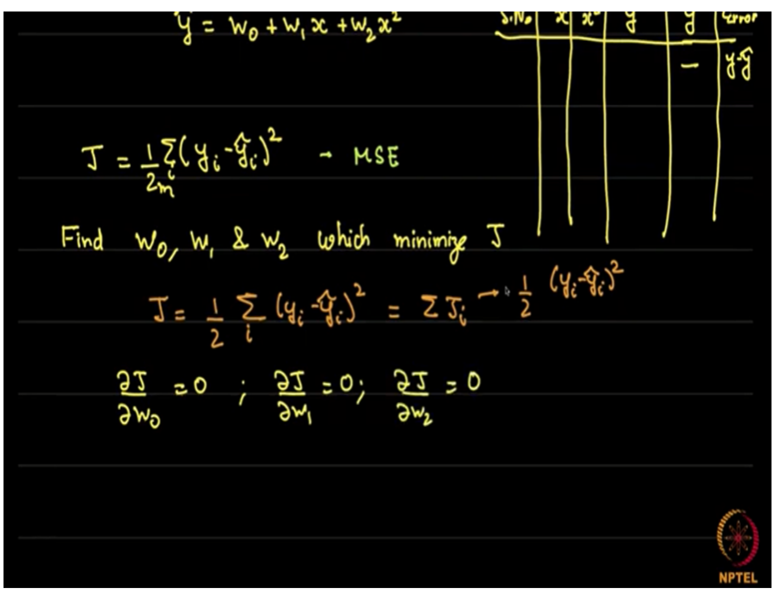
$$J = \frac{1}{2}\sum_i \left(y_i - \hat{y}_i\right)^2$$

You might notice that sometimes I keep m at the bottom and sometimes I do not really speak it is more out of forgetfulness rather than anything else. Generally, it is a good idea to take mean square error.

$$J = \frac{1}{2m}\sum_i \left(y_i - \hat{y}_i\right)^2$$

So, the error still remains the same. So, the mean square error is the error that we are trying to minimize. So, our problem now becomes if you want the best coefficients to find out w_0, w_1 and w_2, which minimize J. so far as you can see there is not much change in the entire process.

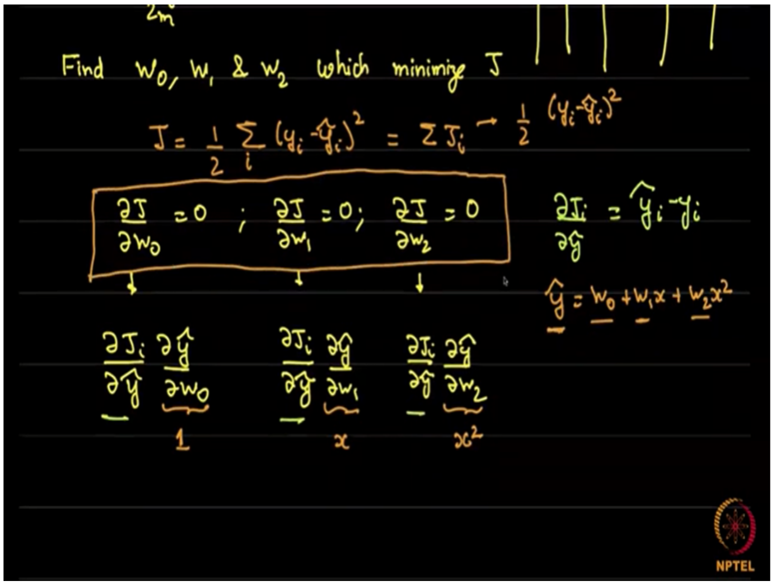
(Refer Slide Time: 04:13)



So now, J is $\frac{1}{2}\sum_i \left(y_i - \hat{y}_i\right)^2$, this we are going to call $\sum J_i$, just like before, where J_i was $\frac{1}{2}\left(y_i - \hat{y}_i\right)^2$, but we are going to set,

$$\frac{\partial J}{\partial w_0} = 0, \frac{\partial J}{\partial w_1} = 0, \frac{\partial J}{\partial w_2} = 0$$

(Refer Slide Time: 04:54)



And you will see that calculating these is not very different from what we did before. Remember $\frac{\partial J}{\partial w_0}$ or let us say If I do,

$$\frac{\partial J}{\partial w_0} = \frac{\partial J_i}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial w_0}$$

This is how we calculated in the first video. So, you can check that out,

$$\frac{\partial J}{\partial w_1} = \frac{\partial J_i}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial w_1}$$

And

$$\frac{\partial J}{\partial w_2} = \frac{\partial J_i}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial w_2}$$

You will notice that these terms the $\frac{\partial J_i}{\partial \hat{y}}$ terms are exactly the same and all these are simply,

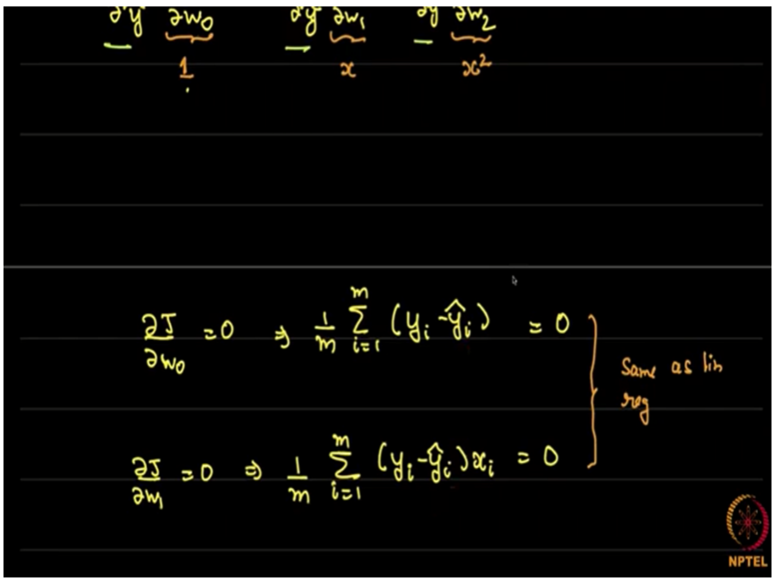
$$\frac{\partial J_i}{\partial \hat{y}} = \hat{y}_i - y_i$$

So, this we had done in the previous videos as well. Now the only thing that changes now is these terms. Now recall $\hat{y}_i$ is $w_0 + w_1 x + w_2 x^2$, so these derivatives now become straight forward,

$$\frac{\partial \hat{y}}{\partial w_0} = 1, \ \frac{\partial \hat{y}}{\partial w_1} = x, \ \frac{\partial \hat{y}}{\partial w_2} = x^2$$

So now if we look at these 3 equations, we can now rewrite them in the following form.

(Refer Slide Time: 06:35)



$\frac{\partial J}{\partial w_0} = 0$, would mean that,

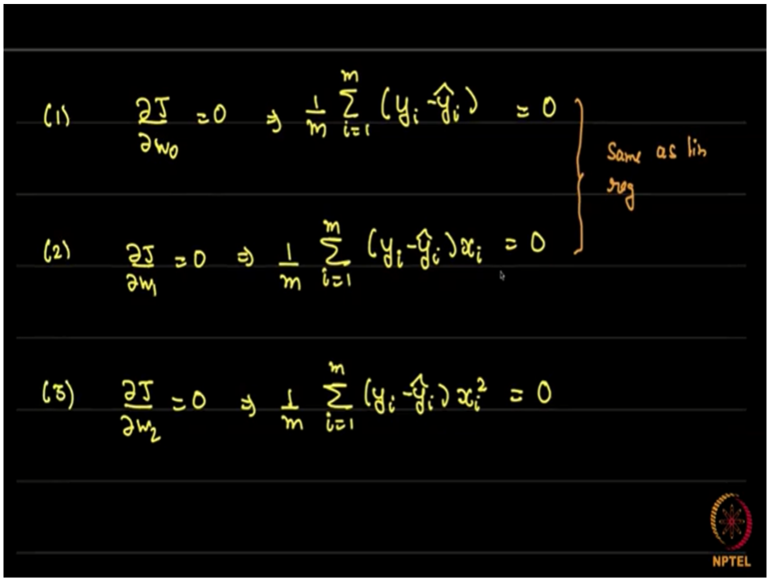
$$\frac{1}{m}\sum_{i=1}^{m}\left(y_i - \hat{y}_i\right) = 0$$

$\sum_{i=1}^{m}\left(y_i - \hat{y}_i\right)$ is multiplied by 1 so this is equal to 0, i = 1 to m we had that $\frac{1}{m}$. So, I will retain it similarly, $\frac{\partial J}{\partial w_1} = 0$, would mean,

$$\frac{1}{m}\sum_{i=1}^{m}\left(y_i - \hat{y}_i\right)x_i = 0$$

$\frac{1}{m}\sum_{i=1}^{m}\left(y_i - \hat{y}_i\right)$, now this is weighted by x_i, you might remember that these 2 equations are the same as the linear regression equations nothing really change. Except now $\hat{y}$ actually has a quadratic term also, as we will see that affects the equation slightly.

(Refer Slide Time: 07:36)

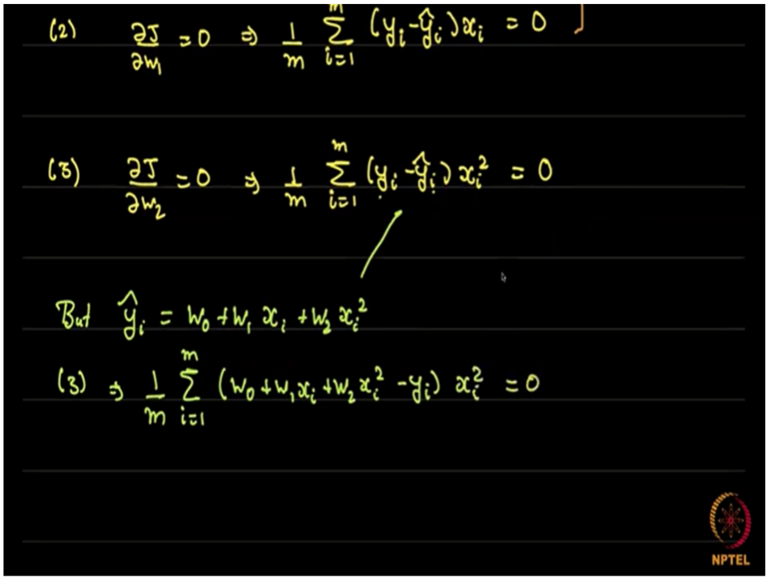


So now the third equation is $\frac{\partial J}{\partial w_2} = 0$, so this will give us,

$$\frac{1}{m}\sum_{i=1}^{m}\left(y_i - \hat{y}_i\right)x_i^2 = 0$$

$\frac{1}{m}\sum_{i=1}^{m}\left(y_i - \hat{y}_i\right)$ now instead of x_i, you are now going to get x_i^2 equals to 0. So let us label these equations as 1, 2 and 3. Now we can expand them.

(Refer Slide Time: 08:09)



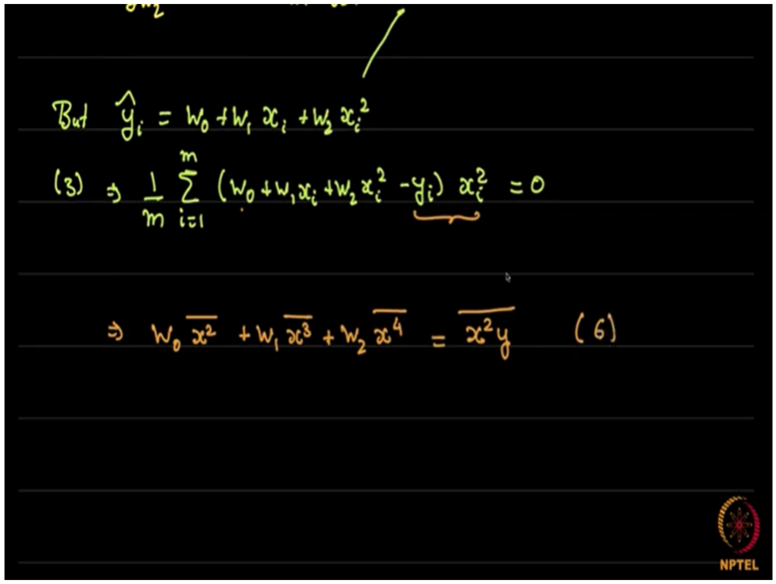
We now know that $\hat{y}$ is $w_0 + w_1 x_i + w_2 x_i^2$. So, if we plug that in, let us say in this equation, I am just going to expand equations 3 and Let as a mild exercise for you to expand equations 1 and 2 so if we expand equation 3,

$$\frac{1}{m}\sum_{i=1}^{m}(w_0 + w_1 x_i + w_2 x_i^2 - y_i)x_i^2 = 0$$

I am going to flip the sign, so that it is easier to write, because I am just taking a negative of this equation and it will still stay the same. So, you are going to have $w_0 + w_1 x_i + w_2 x_i^2 - y_i$ times $x_i^2 = 0$.

In fact, in this whole case, I really should have put because of the derivative in each case the derivative is $\hat{y}_i - y_i$. So, you know that minus is there anyway.

(Refer Slide Time: 09:26)

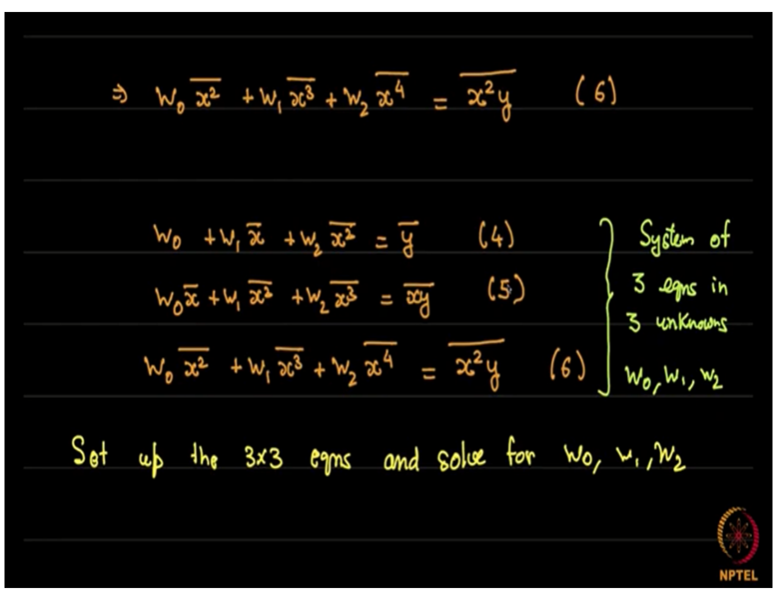


So, if we come here, we see the first term is $\frac{w_0 x_i^2}{m}$. So, I can now write it as before as $w_0\overline{x^2}$, the next term says $w_1\overline{x^3}$, this term is $w_1 x_i \times x_i^2$ and the third term is $w_2\overline{x^4}$. And this whole thing can be taken to the right-hand side and written as $\overline{x^2 y}$.

$$w_0\overline{x^2} + w_1\overline{x^3} + w_2\overline{x^4} = \overline{x^2 y}$$

So, I am going to call this equation 6. Now equation 4 will correspond to this, I am going to write that down directly.

(Refer Slide Time: 10:20)



So, equation 4 is

$$w_0 + w_1\overline{x} + w_2\overline{x^2} = \overline{y}$$

This is equation 4 and equation 5 is

$$w_0\overline{x} + w_1\overline{x^2} + w_2\overline{x^3} = \overline{xy}$$

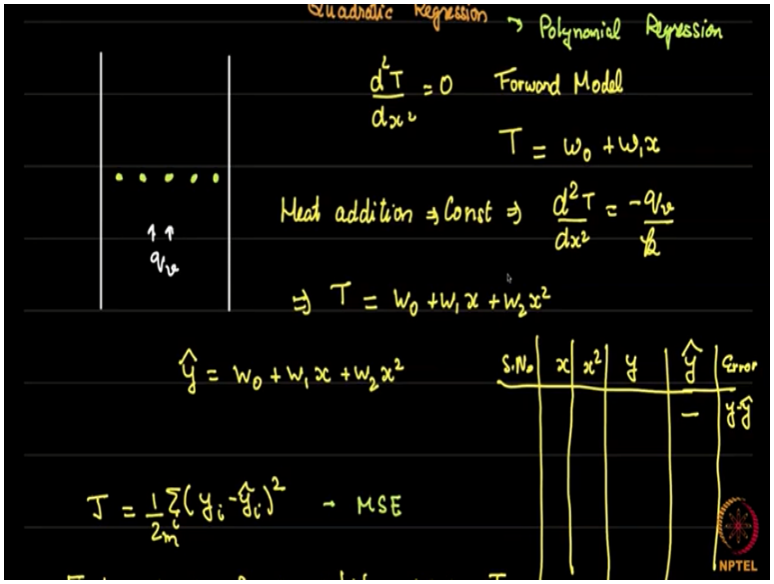
Let us just copy equation 6 and put it here, this is equation 6. Now the way this helps and all of you can probably see the pattern at this point in a straightforward fashion, the first equation says the same thing that it did in linear regression, that is the prediction at the average location should be the average of the predictions.

So, this equation has exactly the same meaning as linear regression. now notice all that is happening is in each equation an extra x is introduced. So, this becomes $\overline{x}$, this becomes $\overline{x^2}$, this instead of x, becomes x^2, this becomes x^3, x^2 becomes x^3, x^4, y becomes xy and then x^2y. So, this now is a system of 3 equations in 3 unknowns. what are the unknowns?

The unknowns are w_0, w_1, and w_2 and just like last time you now have to calculate these extra terms, you we did calculate $\overline{x^2}$, but apart from that you need $\overline{x^3}$ you need $\overline{x^2y}$ you need $\overline{x^4}$ etcetera. So, we leave this as an exercise for you, there is no simple compact formula here. there is a formula that you can derive but it is kind of messy. So, in any quadratic regression problem so set up the equations and solve for w_0, w_1, and w_2 just like, we solved for w_0 and w_1 in linear regression.

So, this is a simple solution. you can do it either by Gauss elimination or by Kramer's rule or whichever method you are familiar with, but nonetheless, it requires the solution of a three-by-three system. now what we have done now is repeat exactly the same process that we did for linear regression and get here. now you should be able to do in case.

(Refer Slide Time: 13:18)



Let us say my right-hand side here was not a constant, but actually, a linear function. If it is a linear function, then your hypothesis function or your forward model would have a cube. So, you would add an additional term and there will be 4 equations in 4 unknowns. So on and so forth so quadratic regression can easily be extended to polynomial regression. But we will see this formally in the next week. so intuitively you can see that the same process that applies to a linear regression actually applies to quadratic, polynomials.

So, there are 2 questions how do we sort of make this without this mess of writing the equations which we will see next week. how to automatically generate that matrix and the second thing is there other problems, which can be solved by the same procedure. So, we will look at a series of linearizable problems in the next week. So please do the assignment if you are taking this course for credit. Thank you.