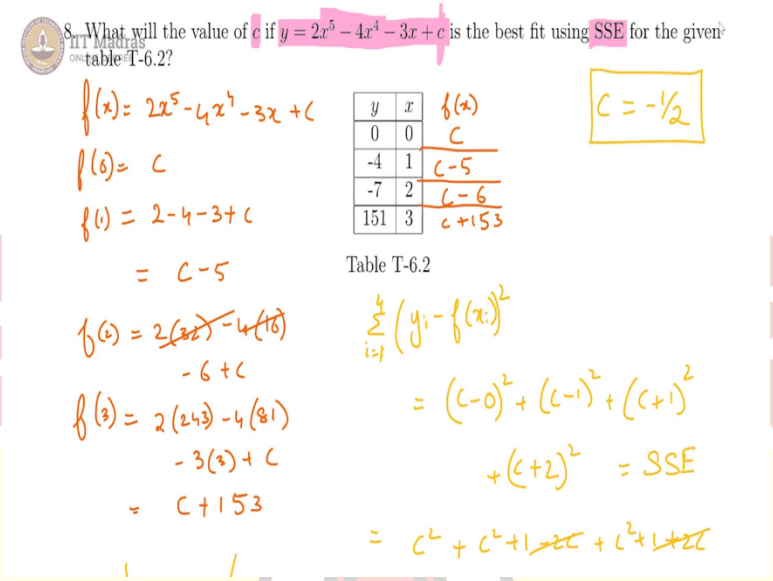
IIT Madras

ONLINE DEGREE

Mathematics for Data Science 1
Week 06 - Tutorial 08

(Refer Slide Time: 0:14)

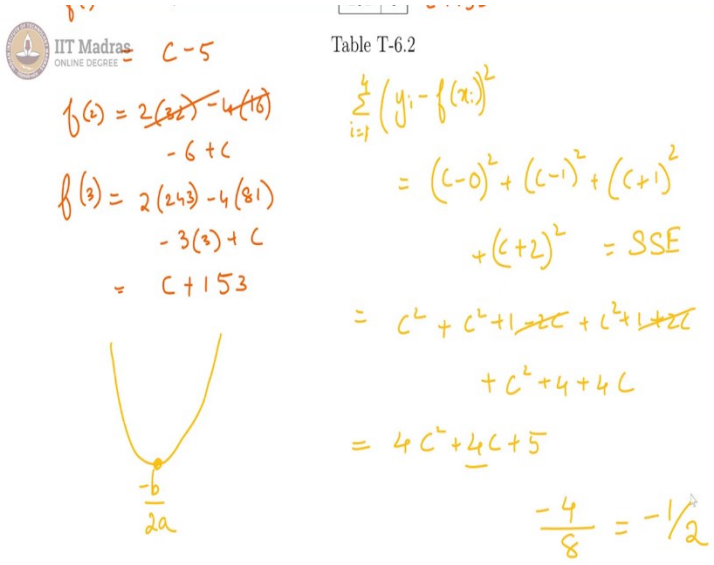


Our last question we are looking at the best fit for some data. So, this is the fit we have obtained a fifth degree polynomial for this data, these 4 points and they are asking what is the value of c, c is the constant term here. What is the value of c if this curve has to be the best fit using sum squared error? So, let us assume this curve is $f(x) = 2x^5 - 4x^4 - 3x + c$. So, we are going to have to also put up the $f(x)$ value, so $f(0)$ is then c because everything else is power of x, so $f(0)$=c, and then we have to look at $f(1)$ which is $2 - 4 - 3 + c$ that is equal to $c - 5$.

So, here this is $c - 5$ and then $f(2)$ is $2 \times 32 - 4 \times 16 - 6 + c$, now 2×32 is 64, 4×16 is 64, so these two cancel off, so you get $6\,c - 6$. And lastly, we have $f(3)$ which is $2 \times 243 - 4 \times 81 - 3 \times 3 + c$ so that gives us $c + 153$, so here it will be $c + 153$. So, for finding SSE we are going to have to do $f(x - y)$ or $(y - f(x))^2$.

So, $(y_i - f(x_i))^2$ and we are going to sum it from $i = 1$ to 4 and that gives us $(c - 0)^2 + (c - 1)^2 + (c - 1)^2 + (c + 1)^2 + (c + 2)^2$ So, this is the sum squared error.

(Refer Slide Time: 3:08)



And we get $c^2 + c^2 + 1 - 2c + c^2 + 1 + 2c + c^2 + 4 + 4c$, so this $-2\,c$ and this $+2\,c$ cancels off and we arrive at $4c^2 + 4c + 5$, this is our sum squared error it is a quadratic in c and for minimum and this is also an upward facing quadratic because the coefficient of $c^2 > 0$, so it will be a parabola like this and the minimum occurs at this point which is the vertex of the parabola and that we know is $\frac{-b}{2a}$, here $-b = -4$ and $a = 4$ so $2a = 8$ so you get $\frac{-1}{2}$, so for $c = \frac{-1}{2}$, we get the minimum sum squared error. Thank you.