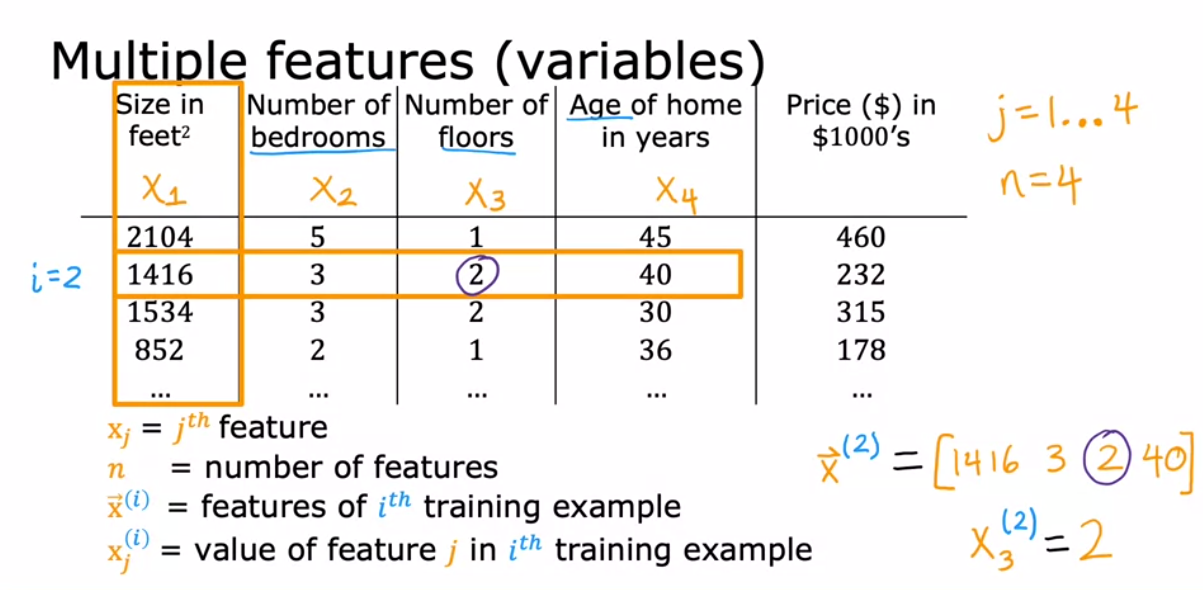
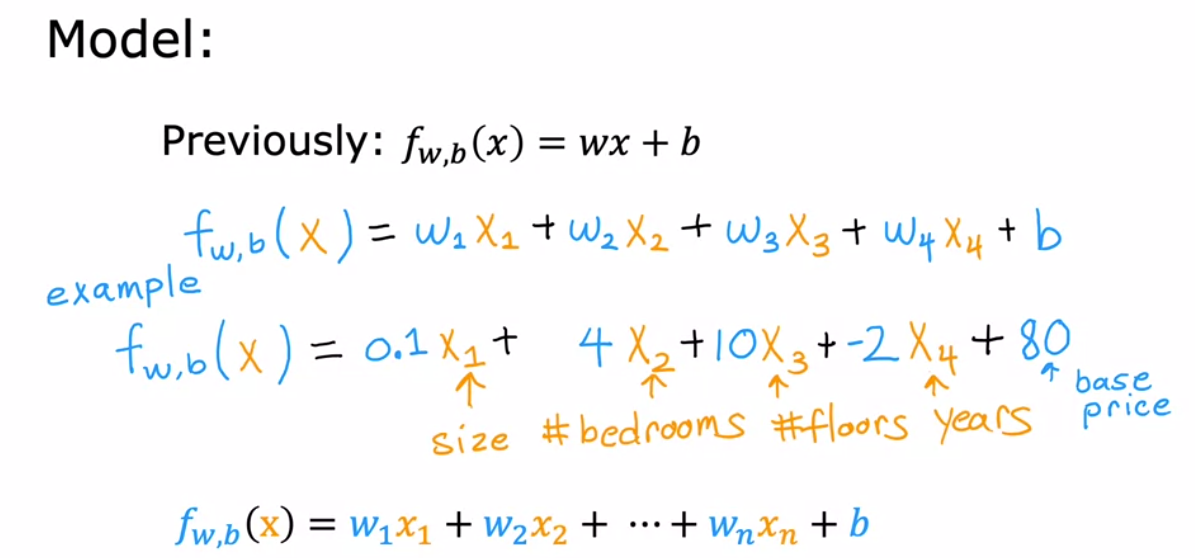
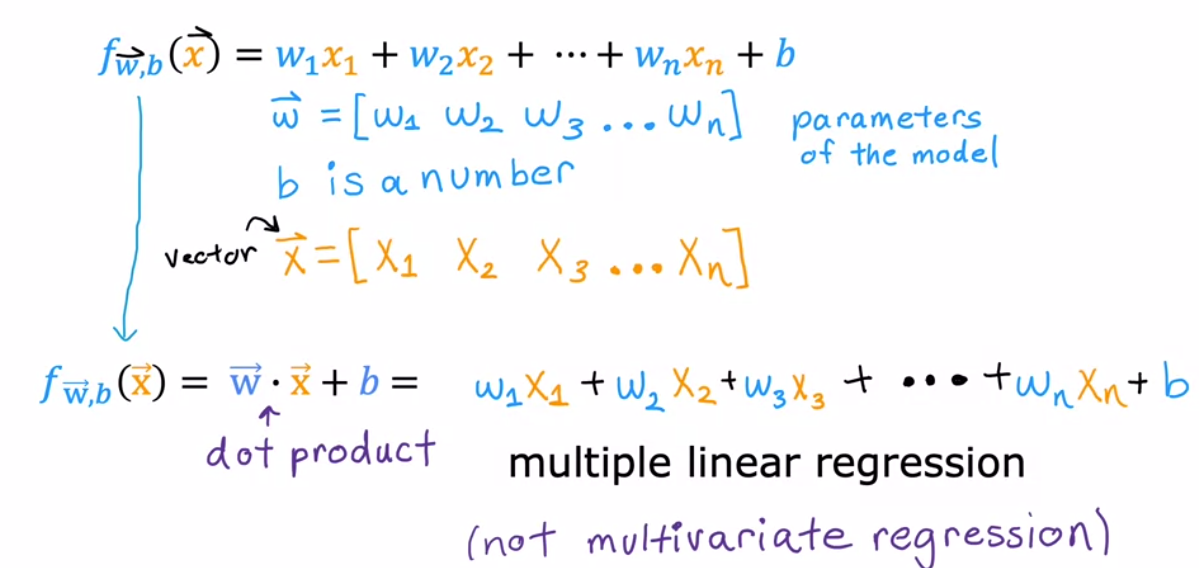
Single feature:  
A picture containing text, screenshot, font, number

Description automatically generated

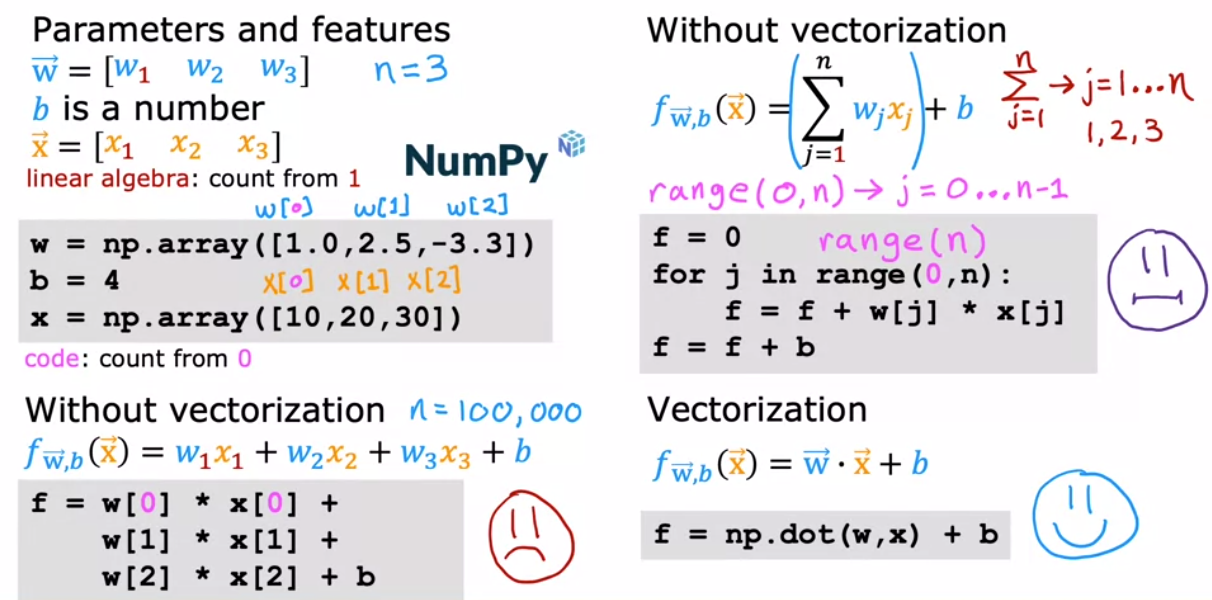


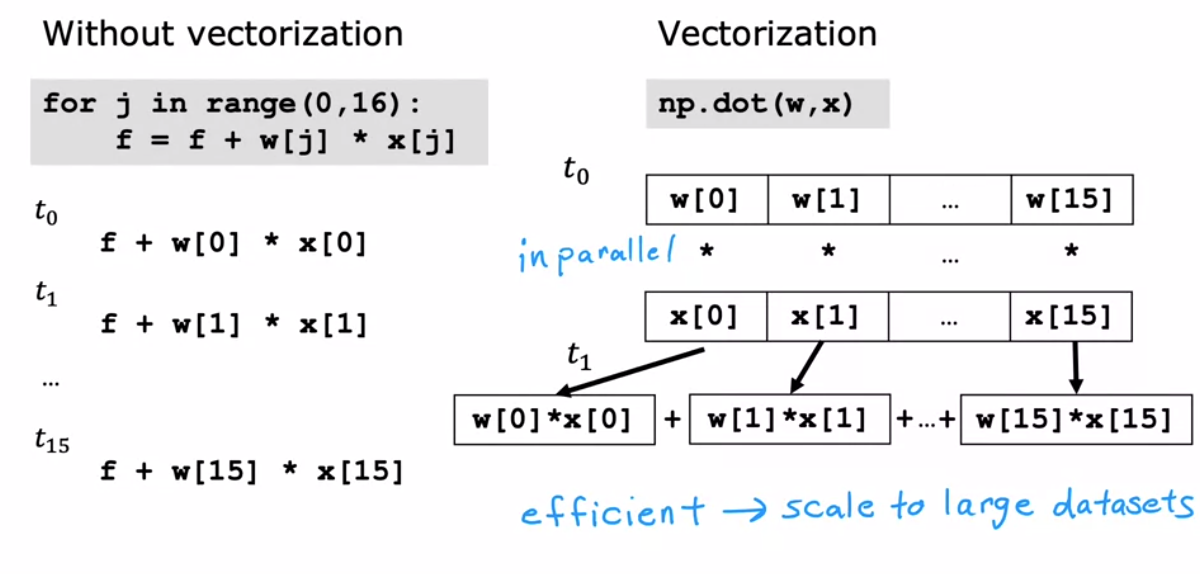


Base price of a house starts at 80 dollars, for each additional floor, + 10…

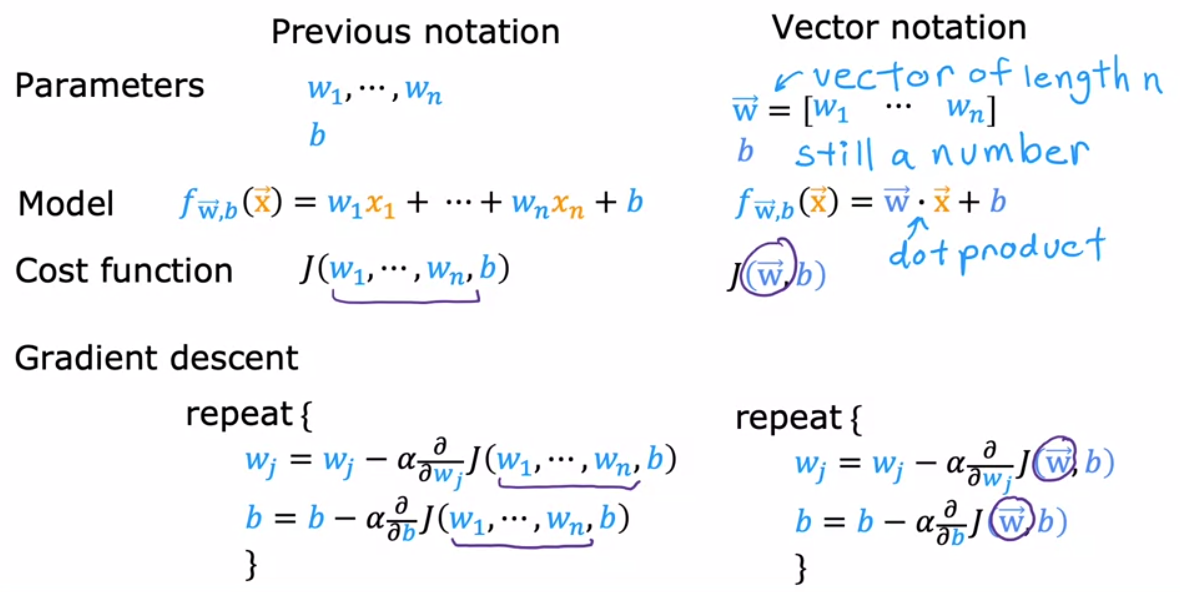


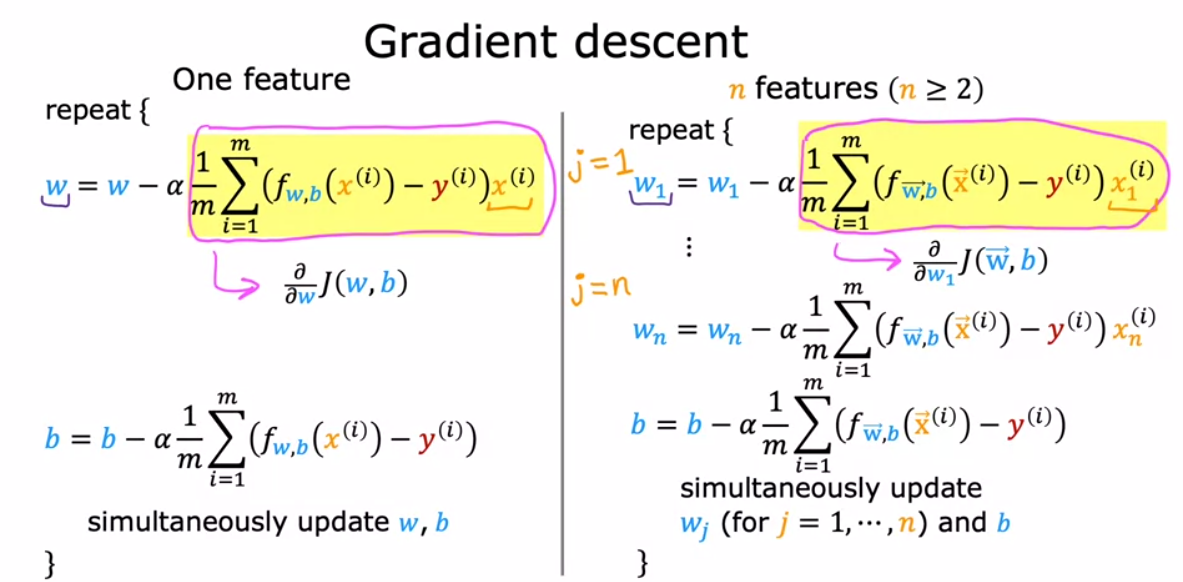
Vectorization



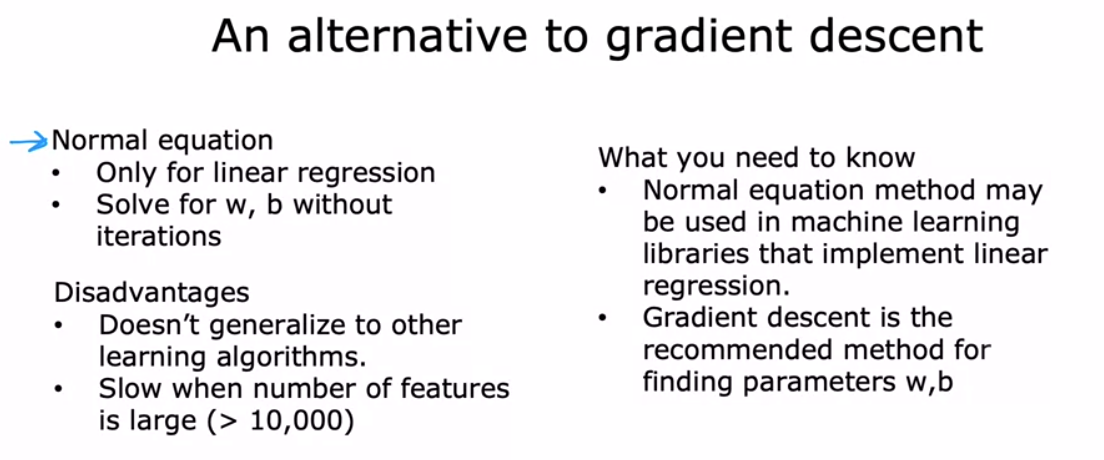






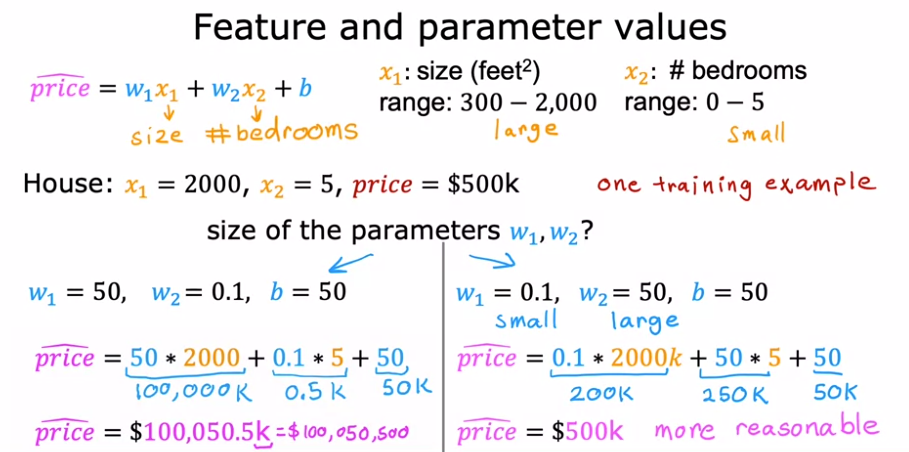


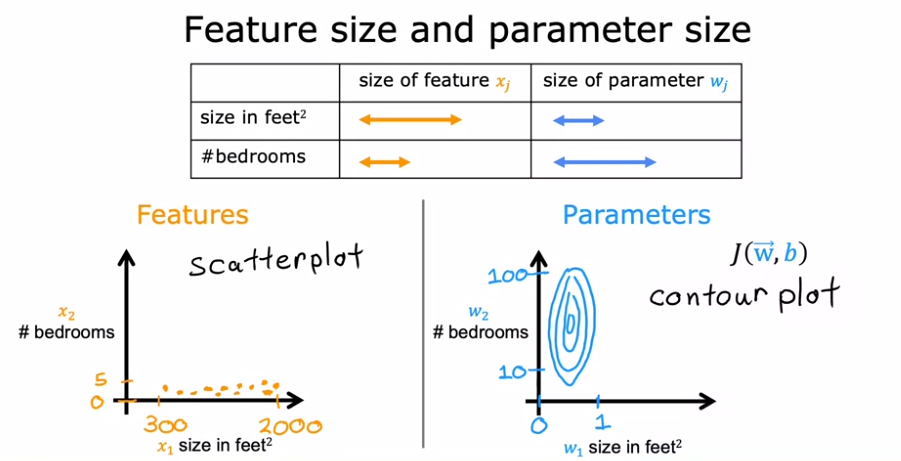
y: output, compared to f



Feature Scaling

When the range is relatively large, it would be good if the model can learn to choose a relatively small weight, vice versa



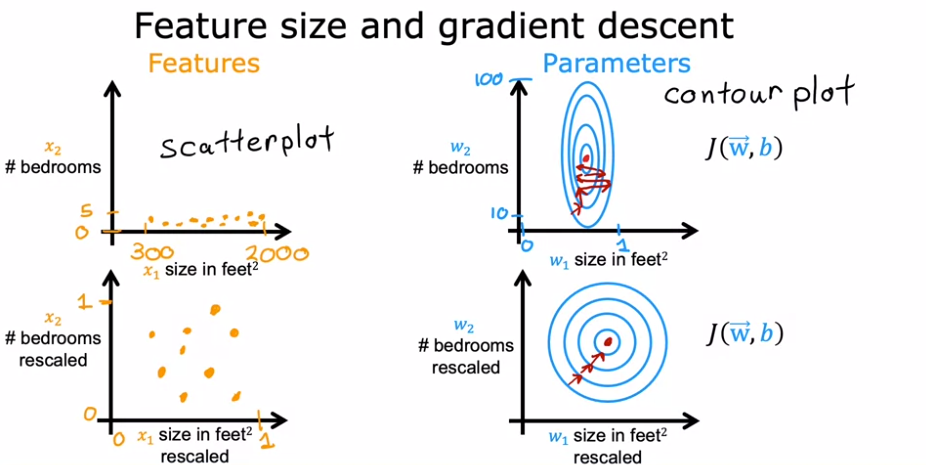


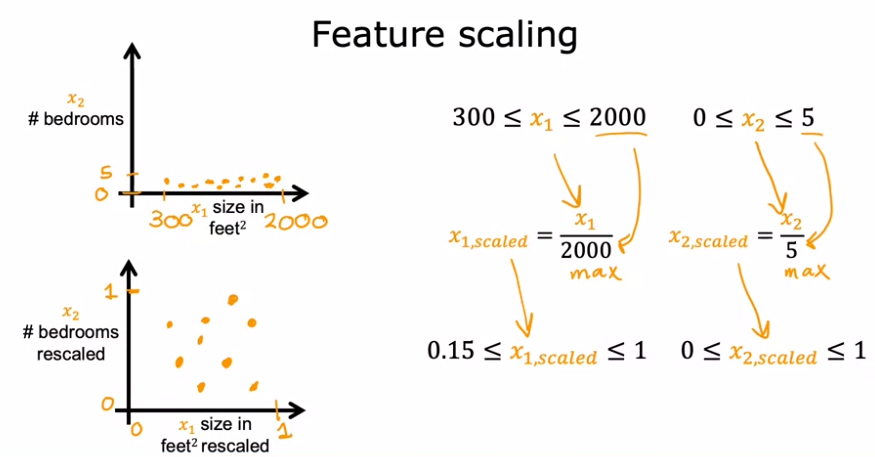
contour plot: the horizontal axis has a much narrower range, say between zero and one, whereas the vertical axis takes on much larger values, say between 10 and 100. So the contours form ovals or ellipses and they're short on one side and longer on the other. And this is because a very small change to w1 can have a very large impact on the estimated price and that's a very large impact on the cost J.

Because w1 tends to be multiplied by a very large number, the size and square feet.

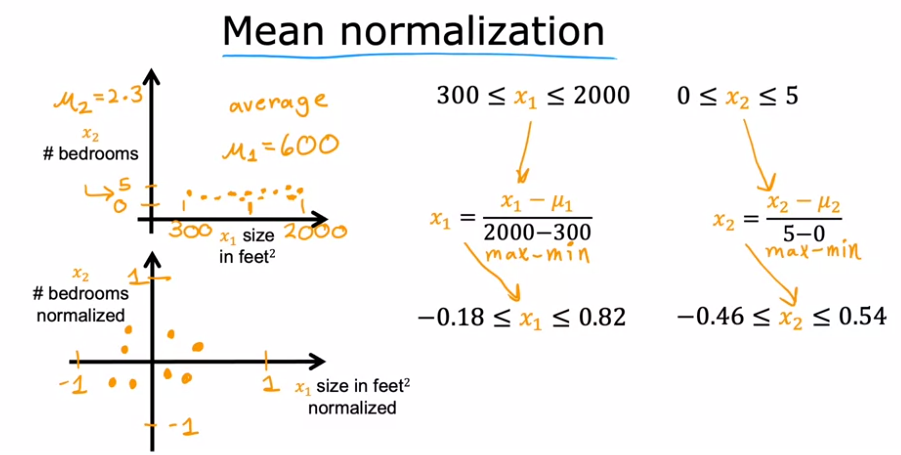
In contrast, it takes a much larger change in w2 in order to change the predictions much.

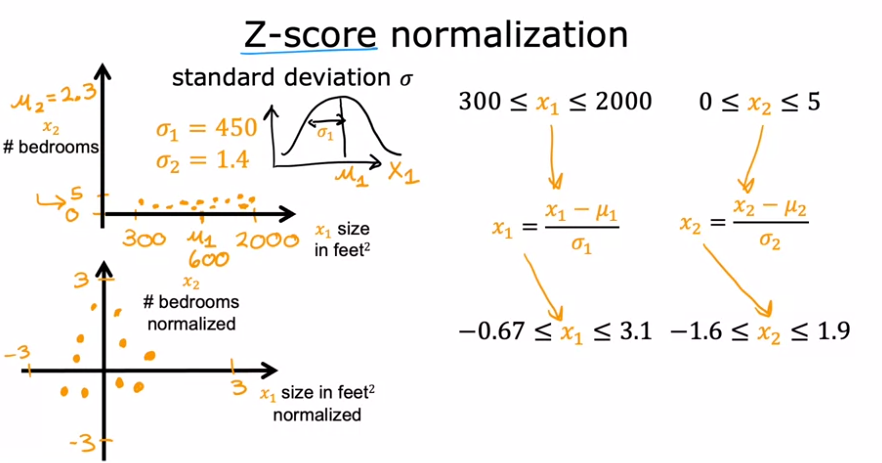
And thus small changes to w2, don't change the cost function nearly as much.

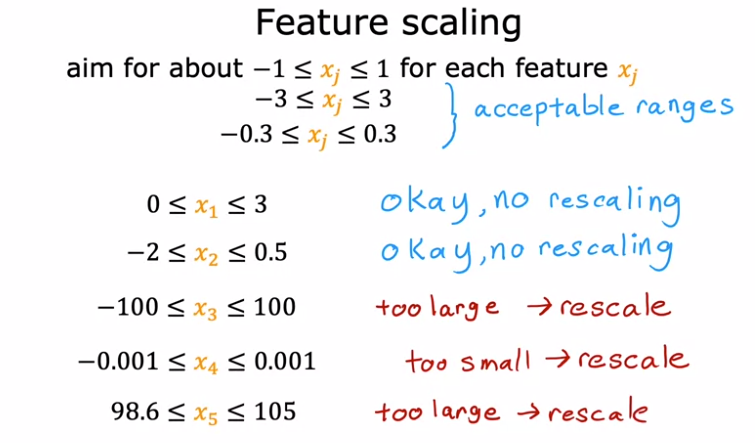




All divided by max

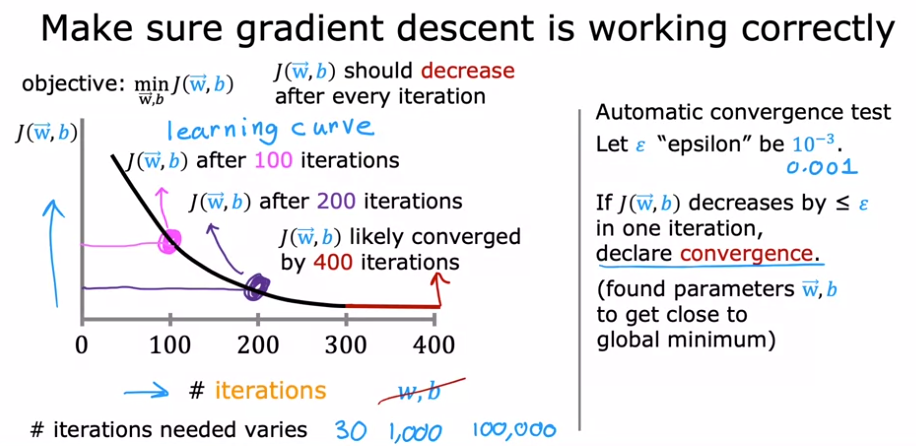


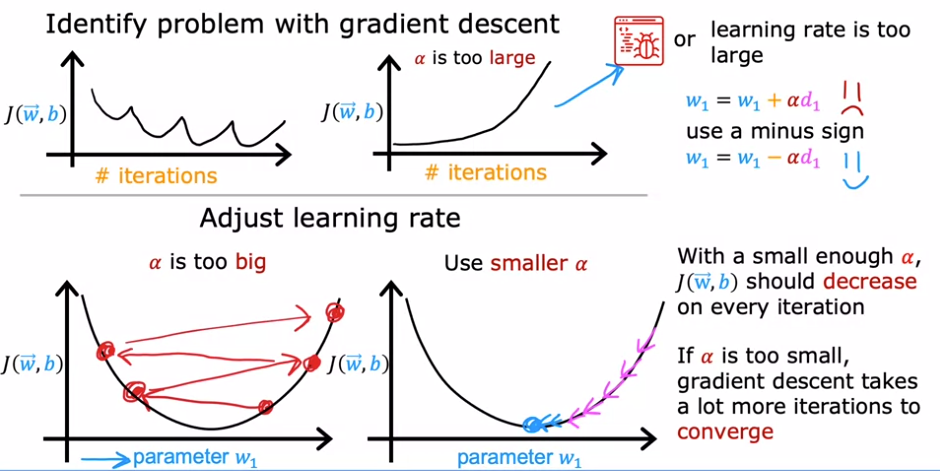






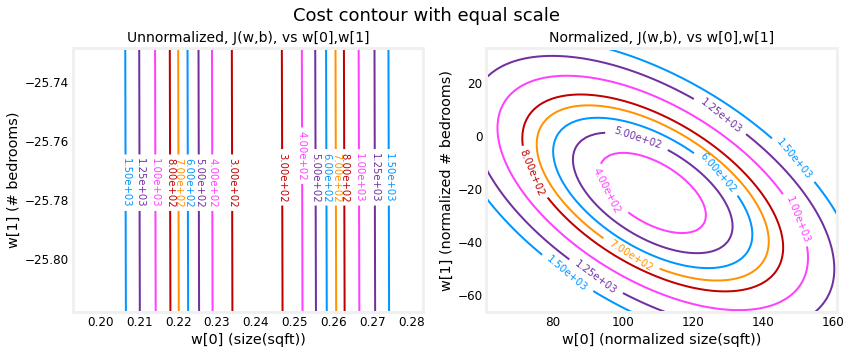
The cost should decrease after every iteration.



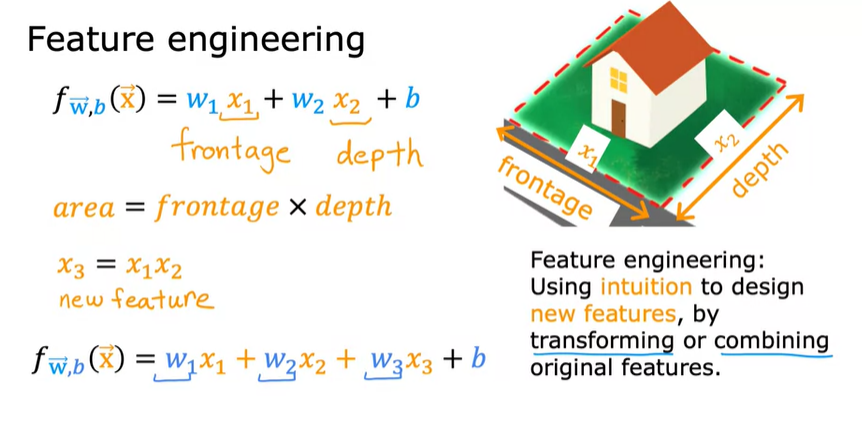




Optional lab3:



Feature engineering



Polynomial regression

