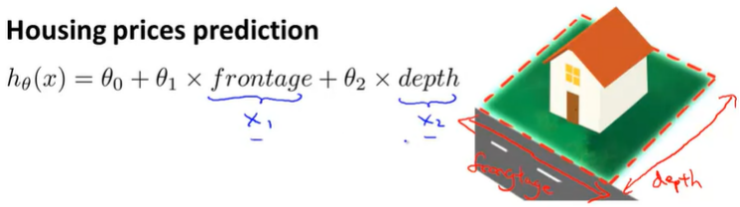
4. Polynomial Regression

August 6, 2021

1 Choosing Features

We are not restricted to what we choose as features for our hypothesis functions. Consider the example of predicting the price of a house. We can have the frontage and depth of a house as 2 features to include in our hypothesis function.



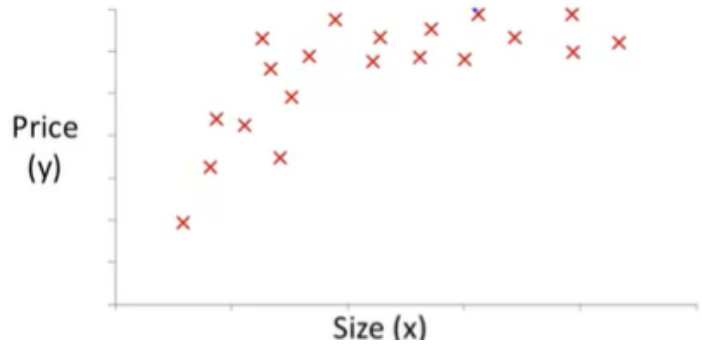
Upon learning our model to predict the price of a house we might find that our model is not as good at predicting housing prices as we hoped. We can then try and revisit the features we defined and conclude that only 1 feature (the area of the house rather than the frontage and depth) might result in a model better at predicting housing prices. This is possible.

The thought experiment above begs the question, "How will we know whether one set of features will allow our model to perform better than another ?". To answer this question we introduce Polynomial regression

2 Polynomial Regression

Our hypothesis function need not be linear!

Consider the example training set below



A number of models can be fit to this training set

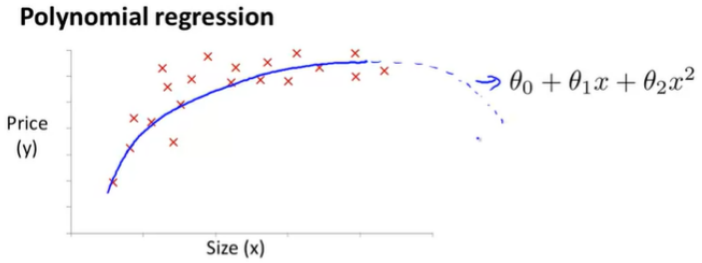


Figure 1: Quadratic model

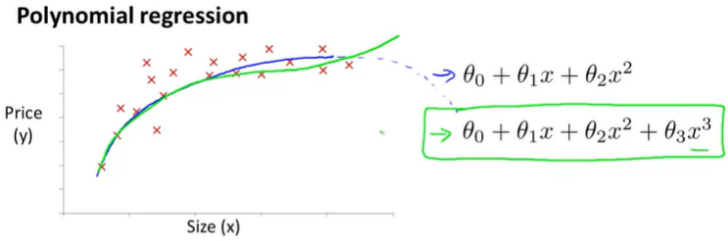


Figure 2: Cubic model

From both the above models we can see that possible "good" features to choose are the size of the house plus the square of the size of the house OR the

size of the house plus the square of the size of the house plus the cube of the size of the house.

Now we have narrowed down the possibility of 2 sets of features to choose from but which one of the 2 will be the best one for our model ? There exist algorithms to determine this.