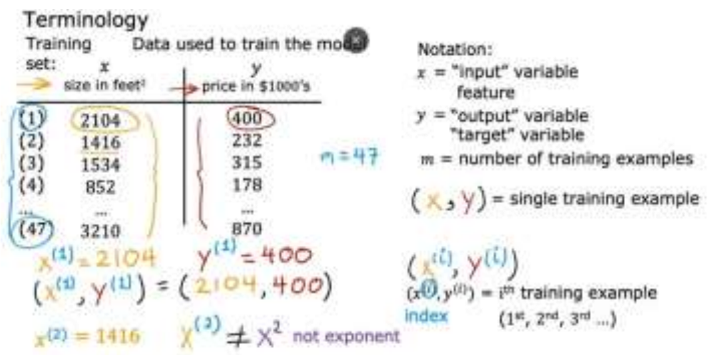
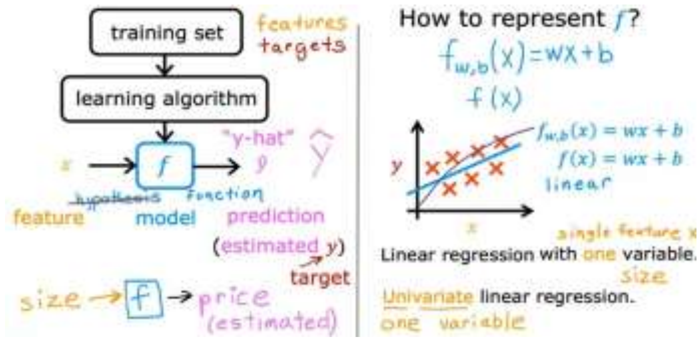
36 – Notation

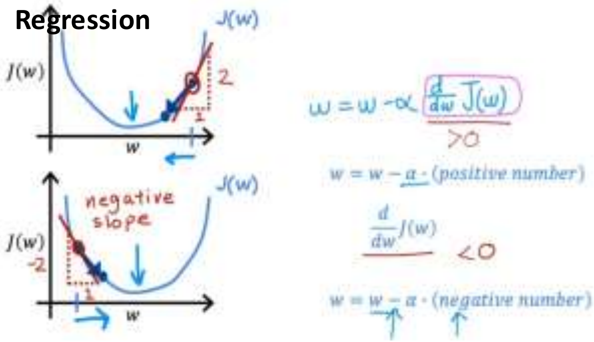


36.1 - Notation

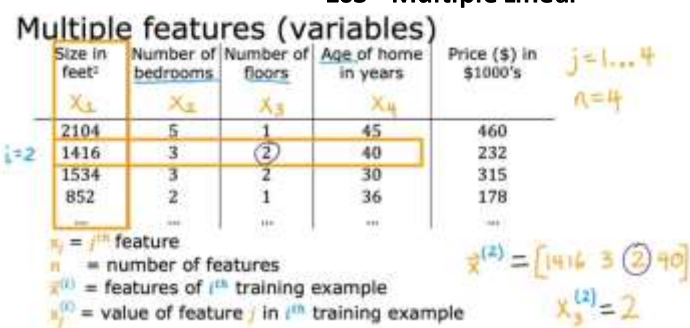


101 - Gradient descent intuition

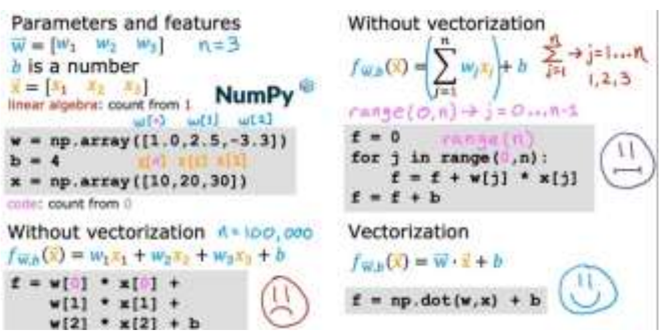
Regression



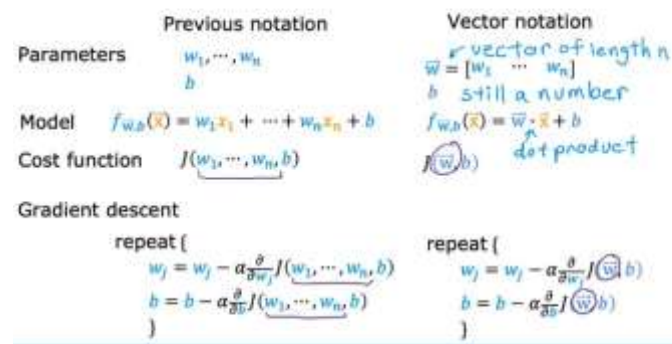
103 - Multiple Linear



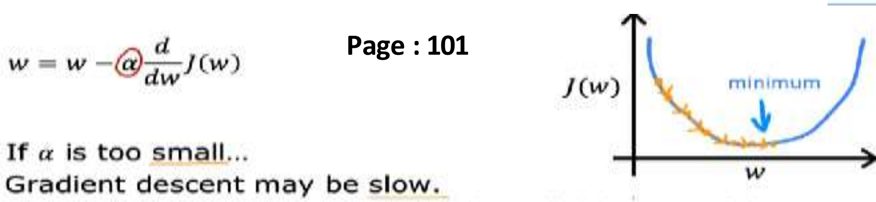
104 - Vector



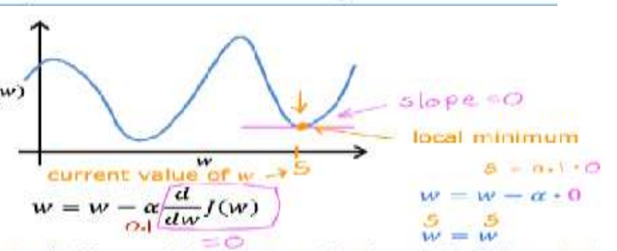
104 – Vector notation cost func



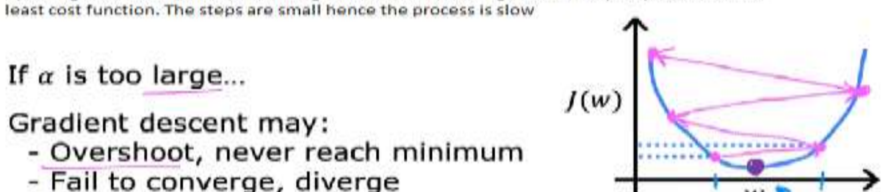
Page : 101



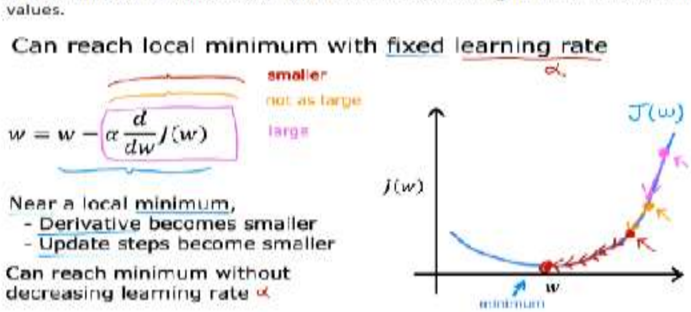
By taking the smallar value of **Learning Rate** we will be taking smallar baby steps towards the least cost function. The steps are small hence the process is slow



Q. what if we reach the minimum cost function point? will the gradient descent algo. update the value of w again?
A. Yes, the algorithm will chamge/update the valuie of w again n again but if the function has already reached the minimun point then it means that the **slope(the derivative part of the formula) will be equal to 0** hence the value of w wont change and be always at local minima.

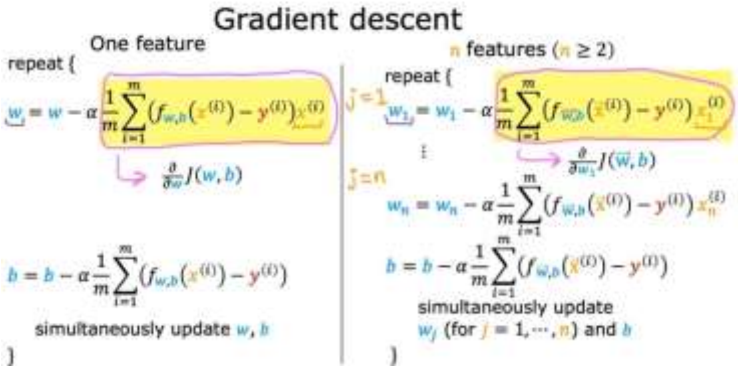


By taking the larger value for **Learning Rate** we will be taking the larger steps towards the least cost function that causes **overshoot** of values and resists reaching the minimum and fails to converge the values.

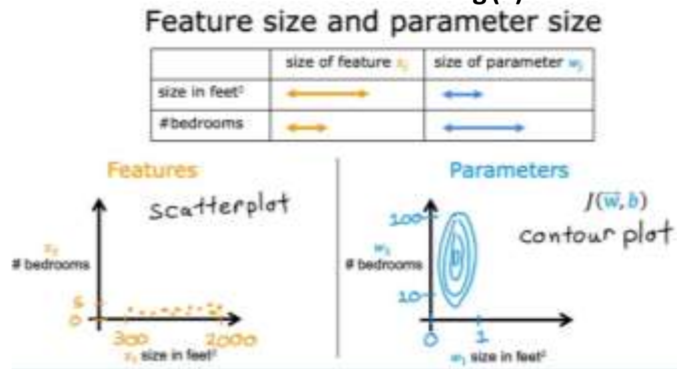


We can reach local minima while fixing the value of **Learning Rate.**
As we get closer to local minima the derivative(slope) part will decrease hence the step we take will be baby step.

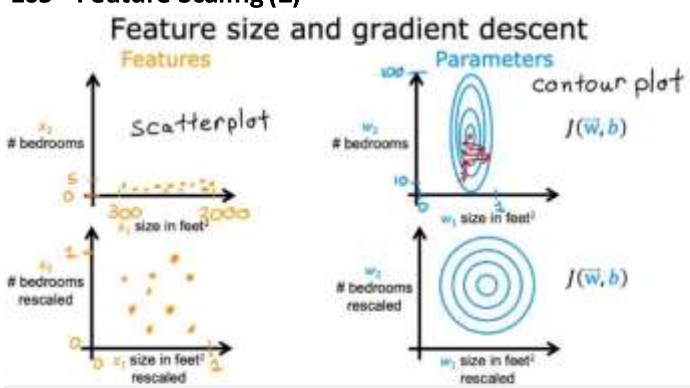
104 - Vector Notation Gradient Descent(2)



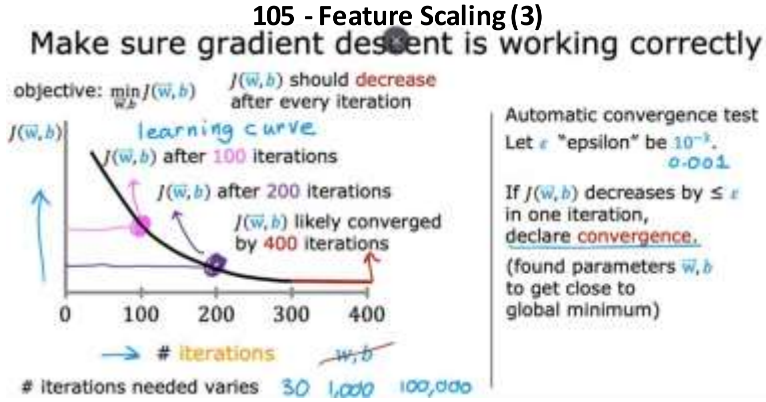
105 - Feature Scaling (1)



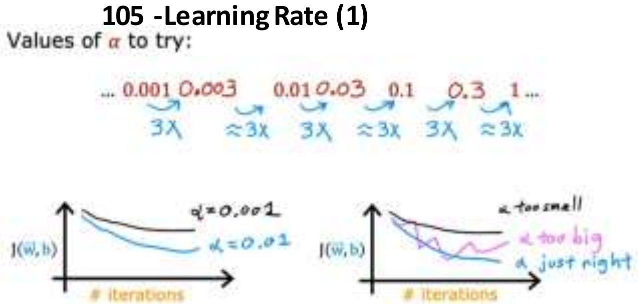
105 - Feature Scaling (2)



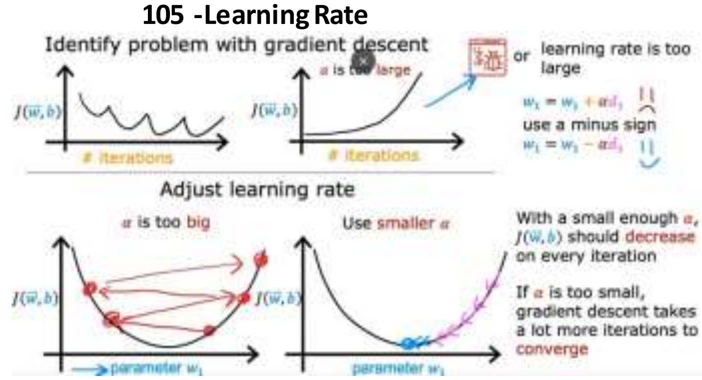
105 - Feature Scaling (3)



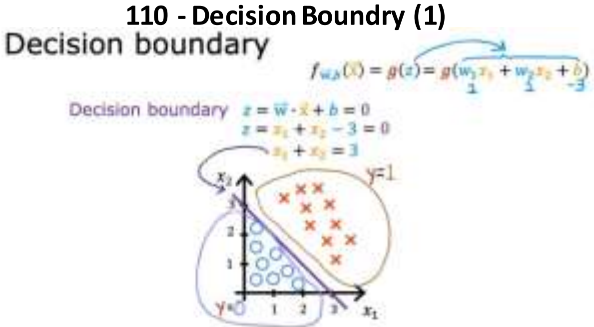
105 -Learning Rate (1)



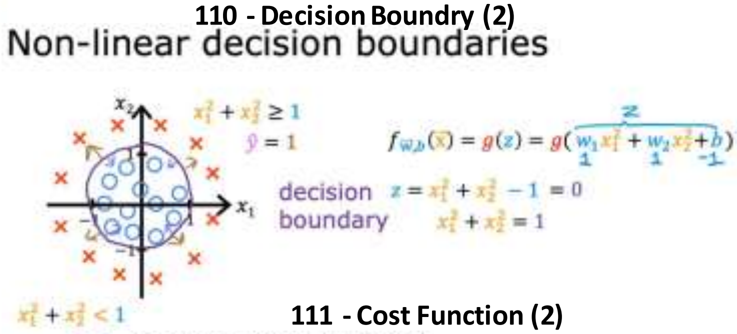
105 -Learning Rate



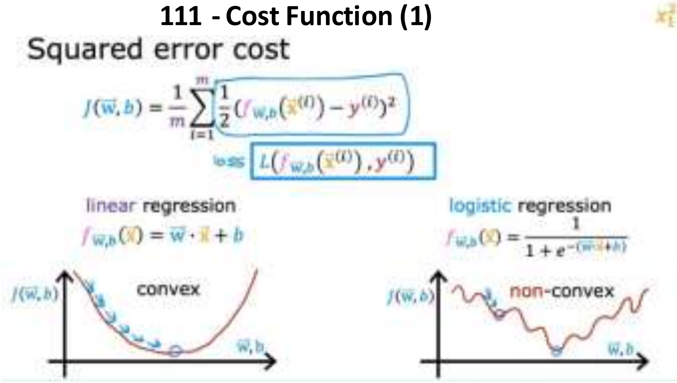
110 - Decision Boundry (1)



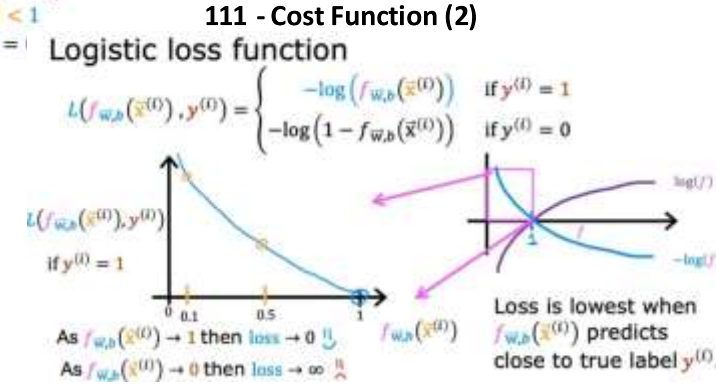
110 - Decision Boundry (2)



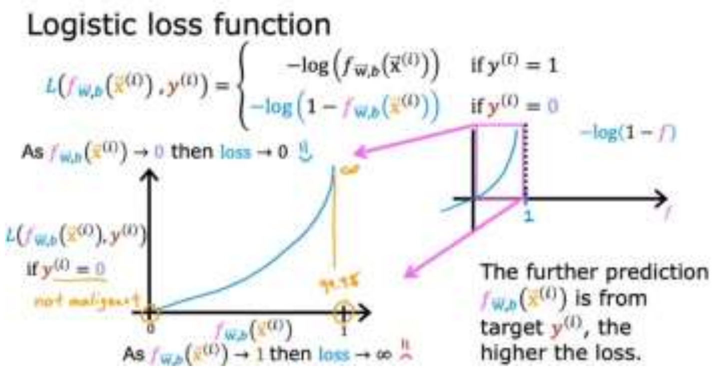
111 - Cost Function (1)



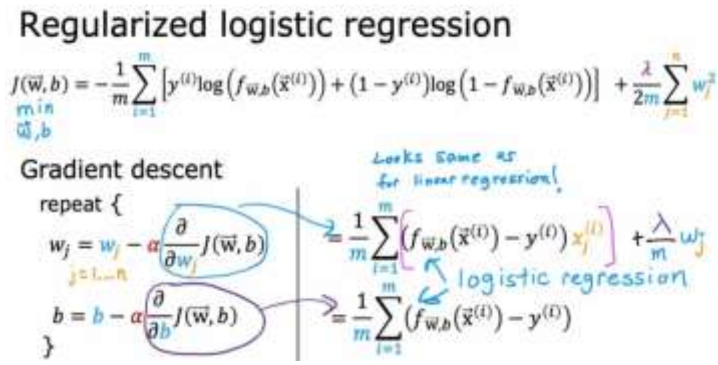
111 - Cost Function (2)



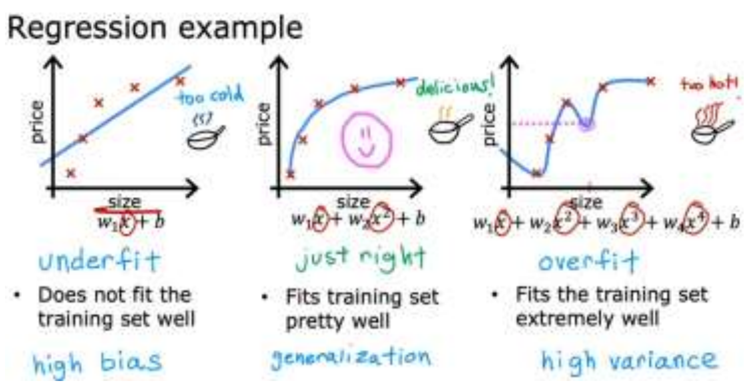
111 - Cost Function (3)



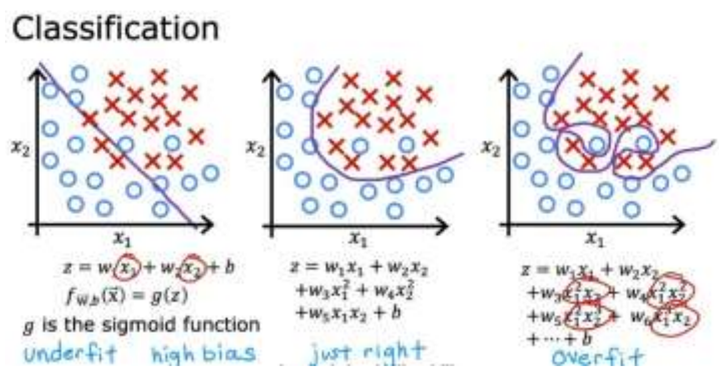
115 - Regularization for Logistics Regression



Overfitting & Underfitting (1)



Overfitting & Underfitting (2)



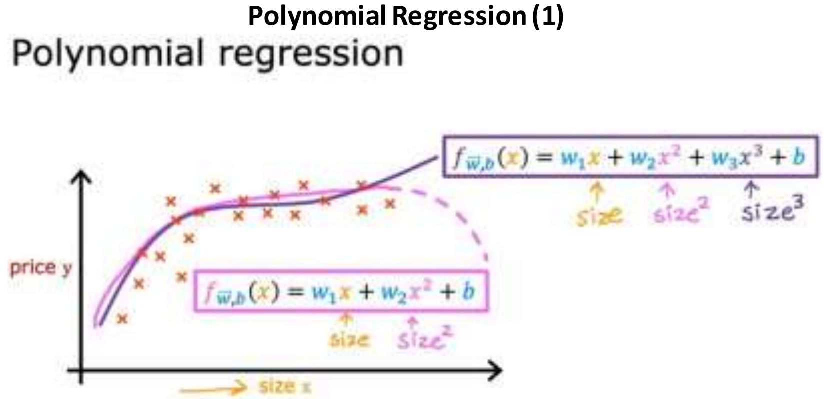
Overfitting & Underfittin (3)

Overfitting: Key definitions

Here are some of the key definitions that'll help you navigate through this guide.

- **Bias:** Bias measures the difference between the model's prediction and the target value. If the model is oversimplified, then the predicted value would be far from the ground truth resulting in more bias.
- **Variance:** Variance is the measure of the inconsistency of different predictions over varied datasets. If the model's performance is tested on different datasets, the closer the prediction, the lesser the variance. Higher variance is an indication of overfitting in which the model loses the ability to generalize.
- **Bias-variance tradeoff:** A simple linear model is expected to have a high bias and low variance due to less complexity of the model and fewer trainable parameters. On the other hand, complex non-linear models tend to observe an opposite behavior. In an ideal scenario, the model would have an optimal balance of bias and variance.
- **Model generalization:** Model generalization means how well the model is trained to extract useful data patterns and classify unseen data samples.
- **Feature selection:** It involves selecting a subset of features from all the extracted features that contribute most towards the model performance. Including all the features unnecessarily increases the model complexity and redundant features can significantly increase the training time.

Polynomial Regression (1)



Polynomial Regression (2)