# Estimation

# what are we tyr to do is to understand a pattern or a slope of a given data

# and the assumption we make is we assume that the dataset has a linear pattern

# y^(y-hat)=wx+b what we try to do that we want to estimate the values which gives the best line that fits most of the dataset

# To calculate the error we use MSE (Mean Equare Error)

# MSE= J(w,b)= 1/N summation for i=1 to i=N (yi-(w\*xi+b))^2

# where MSE calculates the actual value for the data points- the estimate value for the data points all squared

# we do this for all dataset points and divides them by the number of the data points

# to find the best fitted line:

# we need to find the values for the paramters of our model (the weights w and the bias b)that gives the min. MSE

# SO WE NEED TO CALCULATE THE DERivative of j(w,b)-->j'(w,b)

# to do this we use a technique called the Gradient Descent

We calculate at a point we have our parameter value which direction to go used the derivative of the cost function

Updating the parameters:

W = w - alpha \* dw

b = b - alpha \* db

where alpha is the learning rate it tells us how fast or slow to go in the direction the gradient descent tells us to go

A picture containing diagram

Description automatically generated

Steps:

Training:

-initialize weight as zero

- initialize bais as zero

Given a data point:

-predict result by using y|hat= wx+b

-Calculate error

-use gradient descent to figure out new weight and bais values

-repeat n times

Testing:

Given a data point:

-put in the values from the data point into the equation y|hat=wx+b

Calculations

Text, letter

Description automatically generated

Text, letter

Description automatically generated

Text

Description automatically generated with low confidence