

Definition: Field of study that gives computer the ability to learn without being explicitly programmed

A program is said to learn from experience E, with respect to task T and some performance measure P if its performance measure on T as measured by P improves with experience E.

DIVISIONS:

- 1. Supervised
 - a. Regression Problem
 - b. Classification Problem
- 2.Learning Theory
- 3. Unsupervised Learning
- 4.Reinforcement Learning

AIM OF THE CLASS (SUPERVISED LEARNING- REGRESSION)

- 1.Linear Regression(Theory)
- 2.Optimising Techniques :
- a. Batch Gradient Decent(Theory)
- b. Stochastic Gradient Decent(Theory)
- c. Normal Equation(Theory)
- 3. Locally Weighted Regression(Theory)
- 4. Newtons Method(Theory)
- 5. Simple Linear Regression model fit without libraries(Python Code)
- 6. Linear Regression With some feature engineering(Python Code)
- 7. Using different optimizing Algorithms for the above problems
- 8. Introduction to Scikit-learn and using it to fit the above data. (Python code)

SUPERVISED LEARNING (PREDICTIVE LEARNING)

TERMINOLOGY:

- 1. Training Set: D = {(x_i,y_i)}_{i=1 to n}
- 2. Fitting Parameter : Θ
- 3. Hypothesis function : h_e(x)
- 4. Cost Function : J(e)
- 5.Learning Rate: a

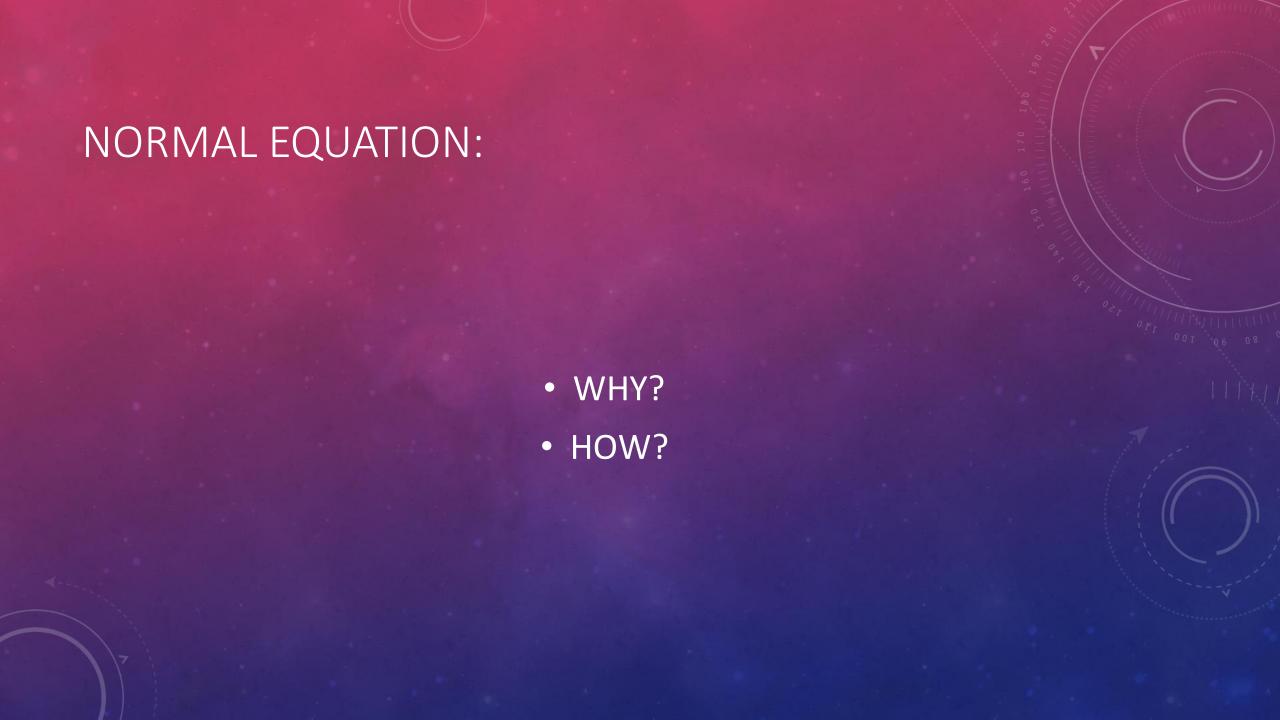
LINEAR REGRESSION Lets start with an example......

MULTI VARIABLE REGRESSION

Let's extend previous example......

OPTIMIZING TECHNIQUES

- 1. Gradient Descent(Batch Gradient Descent)
 - 2. Stochastic Gradient Descent
 - 3. Normal Equation
- 4.Newtons Method(We will discuss tomorrow)



BATCH GRADIENT DESCENT

Why?

How?

Why not?

STOCHASTIC GRADIENT DESCENT

Why?

How?

Why not?

NORMAL EQUATION

- Let's:
- Vectorize all the parameters, and sample data to remove loops
 - Introduce ∇
- We will use the following properties of trace of matrix to find e:
 - tr(AB)=tr(BA)
 - tr(ABC)=tr(BCA)=tr(CAB)
 - ∇_Atr(AB)= B*
 - tr(A) = tr(A*)
 - $\nabla_A tr(ABA^*C) = CAB + C^*AB^*$

NORMAL EQUATION:

$$(X^*X)\theta = (X^*Y)$$

Take Pseudo inverse to get θ .

A LITTLE FEATURE ENGINEERING....

- 1. What is underfitting?
- 2. What is overfitting?
- 3. How can we increase the efficieny of a program?

LOCALLY WEIGHED REGRESSION (LOESS/LOWESS ALGORITHM)

- Why?
- When?
 - How?

CODETIME

