# AI/ML - CS 337 Lab - Lab Assignment 1

B.Nikhil 170050099

August 18, 2019

## Task 2: Ordinary Least Squares

### a) Closed form solution for OLS

Closed form solution for w as proved in class is given by,

$$\boldsymbol{w}^* = (\phi^T \phi)^{-1} \phi^T y$$

Therefore, solution for w* exists if $\phi^T \phi$ is invertible.

### b) Solution obtained from gradient descent

Gradient Descent will move towards min of error function. Our ordinary least square error function will have only one minimum. So, gradient descent if converges will reach that minimum. But if the matrix is singular we cannot get its closed form, In that case we have to use gradient descent only.
parameters used are learning rate lr=0.01

## Task 3: Ridge regression

### a) Why Ridge regression?

Ridge regression is used instead of ordinary least squares to avoid overfitting and control norm(w). Even though the error increases when we use ridge regression it avoids overfitting by using extra hyperparameter.

### b) $ridge(X, Y, \lambda)$ gives same weights as $Linear(X_{new}, Y_{new})$

Regularizing ordinary regression decreases the weights which lets the influence of each factor to decrease which inturn decreases overfitting. So if $X_{new}$ and $Y_{new}$ are formed by decreasing the variance of inputs for that features which least affects the output then ordinary regression of $X_{new}$ and $Y_{new}$ will match ridge regression of $X$ and $Y$.

### c)Plot and its explanation

As explained in the previous part, ridge regression is used to decrease norm of w.Increasing $\lambda$ will decrease norm. More is the lambda less is the norm.
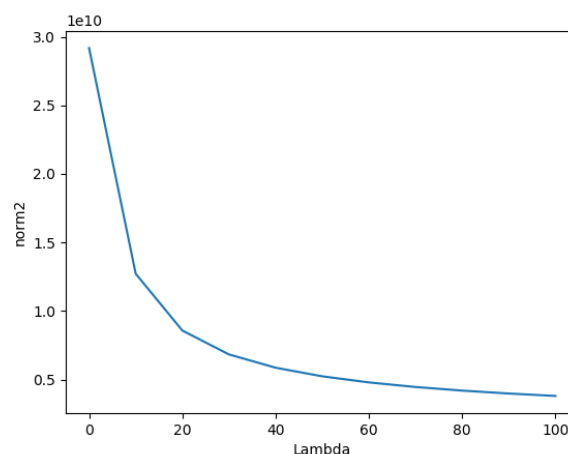parameters used are lambda=100, max-iter=30000, lr=0.005, epsilon = 1e-4



Figure 1: This frog was uploaded via the project menu.