26-03-20 notes

Mini batch gradient descent

We use this algorithm when data is clustered around few points and the total number of data points is huge. Gradient descent will be expensive because of large number of datapoints and stochastic gradient descent will be less accurate.

In mini batch gradient descent, we take few points at a time and do the gradient descent algorithm.

Ridge regression

Ridge regression is used when the number of datapoints in training set is less and we have many points in our testing set.

In gradient descent we minimize the sum of square of error terms but in ridge regression we minimize sum of square of error terms + constant*(slope)^2

As we reduce the constant the slope approaches zero.

Lasso regression

In lasso regression we minimize sum of square of error terms +constant*|slope|

Lasso regression is useful when there are many variables and some variables are redundant that is they don't affect the dependent variable.

Ridge and Lasso regression reduce the dependence of dependent variable on the dependent variable.