Machine Learning:

Programming Assignment 1

REPORT - “Linear Regression”

Submitted by- Harjeevan Singh

1. **About the function Linear Regression:**

The function takes the input as X, Y, k and Tao. It learns from the data given, the parameters using the gradient descent approach.

X: An n x m NumPy array of independent variable values

Y: The n x 1 NumPy array of output values

k: the number of iterations (epochs)

Tao: the threshold on change in Cost function value from the previous to current iteration

1. **Motive of the report:**

To know the effect of different values of n (size of data set) and sigma (the variance of undefined error e) on the trained parameters Beta obtained from the function Linear Regression which implement the Gradient Descent algorithm to get the optimal values for Beta.

1. **Observations:**

Default fixed values of hyperparameters:

* K: 1000
* Tao: 0.0001
* Learning Rate: 0.1

The number of random variables was set to 5.

* 1. Case I:

Sigma = 10

n = 10

* Initially, the value of Beta obtained differed from the original Beta’s from 0.01 to 1.
* Cosine similarity was observed to be 0.88.
  1. Case II:

Sigma = 10

n = 100

* As we increased the n from 10 to 100 not much difference in the values of Beta was observed.
* The value of cosine similarity was also 0.88.
  1. Case III:

Sigma = 10

n = 1,00,000

* In this case also, the results were unexpectedly same as the above two cases.
* Here it was seen that extremely big data set did not make any difference, hence it can be said that size of data set does not affect the Beta values much.
  1. Case IV:

Sigma = 20

n = 100

* As the we increased the sigma from 10 to 20, the values of trained Beta start to differ much more than when the sigma was 10.
* The value of cosine similarity went down to 0.80 from 0.88.
  1. Case V:

Sigma = 20

n = 1,00,000

* This increased data set size did not made any visible difference to the values of Beta and also the cosine similarity stayed the same i.e. 0.80.
  1. Case VI:

Sigma = 1

n = 100

* As the value of sigma was significantly lowered down to 1, the similarity in the trained Beta and the original Beta was also very high as compared to any other precious cases.
* The cosine similarity values raised to 0.95 which also implies the high similarity between the two Beta’s.
  1. Case VII:

Sigma = 1

n = 1,00,000

* Again, the increased size of data set did not made any visible difference in the value of Beta from the previous case i.e. Case VI.
  1. Case VIII:

Sigma = 1000

n = 1,00,000

* There was observed a dramatic decline in the cosine similarity value which dropped down to 0.36
* Hence, it is clear now that sigma is the one that affects the values of Beta more as compared to the negligible changes seen due to change in size of data set.

1. **Conclusion:**

* Sigma affects the value of Beta significantly.
* High values of sigma lead to increased difference in the value of original Beta and the trained Beta.
* As the sigma is decreased close to zero, the trained Beta started to look more like the original Beta.
* It also concluded that the size of data does not have any significance effect on the Beta.