Practical Data Science
(Optimization for ML Problems)

Problem 1: Applying gradient descent variations for regression objective function

An educator wants to see how the number of absences for a student in her class affects the students final marks. As we discussed in class, machine learning objective functions are defined based on error or loss induced by train data while learning. Assume the following objective function and train data for learning.

Objective function:

$$E(w_0, w_1) = \frac{1}{n}\sum_{i=1}^{n}\left(y^{(i)} - w_0 - w_1 x^{(i)}\right)^2$$

Train data:

S.no	x(no. of absenses)	y(final marks)
1	10	70
2	12	65
3	2	96
4	0	94
5	8	75
6	5	82

Do the following:

a) Compute the expression for gradient of E.

Applying batch gradient descent Algorithm

b) Write the expressions for weight updates using batch gradient descent algorithm.

c) Simulate(by hand) the batch gradient descent algorithm for 5 iterations. For each iteration, show the error value and updated weights. Assume initial values of (w0,w1) as (0,0) and eta=0.01

Practical Data Science
(Optimization for ML Problems)

Applying stochastic gradient descent Algorithm

d) Write the expressions for weight updates using stochastic gradient descent algorithm.

e) Simulate(by hand) the stochastic gradient descent algorithm for 6 iterations. For each iteration, show the error value and updated weights. Assume initial values of (w0,w1) as (0,0) and random samples picked are 3, 2, 5, 1, 6, 4. Keep eta=0.01

Applying stochastic minibatch gradient descent Algorithm

f) Write the expressions for weight updates using stochastic minibatch gradient descent algorithm.

g) Simulate(by hand) the stochastic mininbatch gradient descent algorithm for 6 iterations. For each iteration, show the error value and updated weights. Assume initial values of (w0,w1) as (0,0) and random batches of size 3 as follows. Keep eta=0.01

(1,4,5), (3,2,6) (2, 3,4), (1,5,6), (1,3,5), (2,4,6)