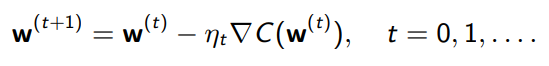
Week 3 Notes

**Gradient Descent**

Gradient is a map which maps vector to other vectors.

Gradient shows how fast the function is changing with respect to the coordinate axes and direction of the change.

It is an iterative algorithm. The algorithm is to start at an initial point and then go in a specific direction with a specific learning rate eta and update the direction according to the final output.



Text, letter

Description automatically generated

**Stochastic Gradient Descent**

Gradient Descent is computationally expensive, Hence SGD.

SGD is used to observe sum structure and optimization problem in ML.

Text

Description automatically generated with medium confidence

SGD is not as smooth as GD.

SGD is .

SGD flies randomly everywhere but the average is still on target while GD moves directly to target.

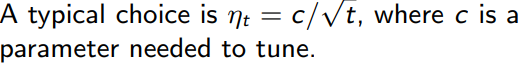
It uses 1 data point per step as opposed to GD which uses all the data points per step.

Computation Complexity 🡪 GD: O(n) SGD: O(1)

Summary: GD has a direct approach and SGD follows a random approach, but SGD is computationally efficient than GD.

**Text

Description automatically generated**

For learning rate in SGD 

**Minibatch SGD**

Minibatch SGD selects a batch of data points per step and performs the SGD operation.

Hence it is not that computationally expensive and not that unpredictable. i.e., the best of both methods.

Text, letter

Description automatically generated

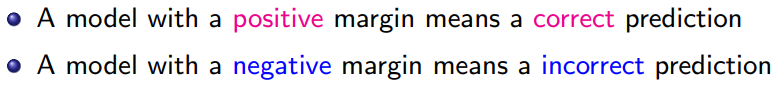
If batch size is 1. SGD = Minibatch SGD.

Typical batch size choices: 32,64,128…

**Linear Models for Classification**

Classify the model and give output as 1 or -1.

Margin = 



To minimize the loss, we maximize the margin.

Text, letter

Description automatically generated

The gradient of the function is A picture containing text, watch, gauge

Description automatically generated

SGD

Text, letter

Description automatically generated

Chart

Description automatically generated with low confidence

To update the weights, we use

Text, letter

Description automatically generated