CSC413: Assignment 2

Farbod Mohammadzadeh
1008360462

19 October 2024

3 Pages

Problem 1 Optimization

1.1 Mini-Batch Stochastic Gradient Descent

1.1.1 Minimum Norm Solution

We begin by examining the update steps for both full-batch gradient descent (GD) and mini-batch SGD.

Gradient Descent (GD)

The gradient of the loss function $L(w) = \frac{1}{2}\|Xw - t\|^2$ with respect to w is given by:

$$\nabla L(w) = X^T(Xw - t)$$

Starting from an initial point $w_0 = 0$, the update rule for gradient descent is:

$$w_{k+1} = w_k - \eta \nabla L(w_k)$$

Substituting the expression for the gradient, we have:

$$w_{k+1} = w_k - \eta X^T(Xw_k - t)$$

Notice that the gradient $X^T(Xw_k - t)$ lies in the row space of X. Since we start with $w_0 = 0$, which is in the span of X, and every update direction is also in the span of X, it follows that every iterate w_k remains in the span of X. As gradient descent converges, it converges to the minimum norm solution w^*, which lies in the span of X.

Mini-Batch SGD

In mini-batch SGD, at each iteration, we compute the gradient using a randomly selected mini-batch of the data. Let B denote the indices of the samples in the mini-batch. The gradient for the mini-batch is:

$$\nabla L_B(w) = X_B^T(X_B w - t_B)$$

where X_B and t_B are the sub-matrix and sub-vector corresponding to the mini-batch B.

The update rule for mini-batch SGD is:

$$w_{k+1} = w_k - \eta \nabla L_B(w_k)$$

Again, since the mini-batch gradient $\nabla L_B(w_k)$ lies in the span of the rows of X, the update step remains in the span of X. Starting from $w_0 = 0$, each update keeps w_k in the span of X, and the algorithm will converge to a solution $\hat{w}$ in the span of X.

Uniqueness of the Minimum Norm Solution

The minimum norm solution $w^* = X^\dagger t$ is unique and lies in the span of X. Since both full-batch gradient descent and mini-batch SGD are confined to the span of X, and the minimum norm solution in this span is unique, it follows that:

$$\hat{w} = w^*$$

1.2 Adaptive Methods

1.2.1 Minimum Norm Solution